

英単語8000の親密度測定の妥当性

Validity of Measuring the Familiarity Levels of 8000 English Words

西 出 公 之

NISHIDE Kimiyuki

本稿は、西出・水本 (2009) で発表した8000語についての親密度測定が妥当なものかどうかについて検証しようとするものである。

「語彙リスト」における語のレベル付けは、主にその語の頻度情報に拠ってきた。よく出合う語は習得され易いという傾向は厳然としてあると思われるが、同じ回数出合ったとしても、覚え易い語・覚え難い語があるというのも実感されるところである。

ALC 社の Standard Vocabulary List (SVL) や JACET8000 は British National Corpus から得た頻度情報に基づいているし、北大語彙表は *Time* と米国エネルギー省の科学論文収録データベースである Department of Energy Corpus から得た頻度情報を基にしている。しかし、巨大コーパスから得られる頻度情報だけでは不十分であることは認識されているようで、いずれも補正が試みられている。JACET8000 では、学習者が接する教科書や試験問題の英文をコーパスにし、そこから得られた頻度情報を併用している。この場合は、メイン・コーパスの頻度情報をいくつかのサブ・コーパスでの頻度情報で補正するというので、補正にも頻度情報が使われている。しかし、何らかの補正が必要であることは認識されていることになる。ALC の SVL は、頻度情報によるレベル分けを教授者 (日本で教えている複数のアメリカ人) の感覚的な判断によって補正している。北大語彙表のように、辞書における見出し語のレベル表示や辞書の定義語彙リストを参考にレベル付けに補正を加えたり、先行の語彙リストのレベル付けを反映させたりすることも行われている。

西出・水本 (2007) は、SVL からレベルごとに30語を抽出して日本語による意味と英単語をマッチングさせる語彙サイズテスト (「水本テスト」) を大学生に実施し、240語の正答率が必ずしも SVL の語彙レベル順になっていないことを明らかにし、学習者が実際にどの程度知っているのかという習得度を頻度情報によるレベル付けの補正に用いることを提案した。また、テストによる正答率で習得度を測るよりも簡便な方法として、「知っているかどうか、知っているとするればどの程度知っているか」について学習者にアンケートを行って親密度を測定することを提案し、その測定を試みてきた。西出 (2007) は、1000語を抽出して10名程度の大学生にアンケートを行った試験的調査である。この調査では、被験者数は限られていたが、各語について差異が観測され、アンケート調査で得られる親密度を習熟度情報と見なすことができるのではないかと示唆した。西出・前田 (2008) は、「水本テスト」で用いられている240語について、同一被験者117人に「水本テ

スト」と「知っているかどうか、知っているとすればどの程度知っているか」のアンケートを行い、正答率と親密度に高い相関が見られたことを報告した。そして、西出・水本 (2009) は、SVL のレベル 8 までの8000語について22名の大学生にアンケートを行い、親密度を測定したものである。これにより、親密度情報を基にした語彙表を作成することが可能になり、頻度情報を基にした語彙表との比較も可能になる。さらに、頻度情報を基にした語彙リストの補正について学習者の習熟度の観点から考察することも可能になる。

しかし、そのためには得られた親密度が十分信頼できるものである必要がある。本稿では、これまでの試験的調査、そしてやや別の観点から行われた横川ら (2006) による調査と比較して、西出・水本 (2009) で得られた8000語の親密度情報についての考察を試みる。また、親密度測定データを群に分けて比較し、内部的な検証も試みる。

1. 西出・水本 (2009) による8000語についての親密度調査の概要

ALC の SVL のレベル 1 からレベル 8 までの8000語について、22名の大学生に、

- 5 この語を使える (文やフレーズを作ることができる)
- 4 知っている (意味を言うことができる)
- 3 見たり聞いたりしているが、意味をはっきり言うことはできない (文脈があれば言えると思う)
- 2 見たり聞いたりしていると思うが、意味は知らない
- 1 知らない (見たり聞いたりしたという記憶がない)

の5段階うちから当てはまるものを数字で答えてもらった。得られた5～1の数値の平均を親密度とし、西出・水本 (2009) では、8000語の親密度と親密度順位を単語のアルファベット順の一覧表として公表した¹。これにより、親密度情報を基にした語彙表を作成することが可能になる。親密度の順位は、被験者の人数が少ないことあって、同位の語がかなり出ている。親密度が5.00であった語は、143語であった。次に高かった親密度は4.95で、これに当たる単語は309語であった。これらは順位としてはすべて144位を与えられた。次に高かった親密度は4.91で、これに当たる単語は229語であり、これらには453位が与えられた。同様に、682位 (評点4.86) の単語が196語、878位 (親密度4.82) の単語が188語、1066位 (親密度4.77) の単語が191語などと続いた。被験者は大学生であり、「やさしい」単語に5.00に近い評点がでるのは当然である²。したがって、得られた親密度だけでは、SVL や JACET8000 のように厳密に1000語ごとのブラケットでくくることはできない³。評点平均に差が出るように被験者数を増やすことが望まれるが、5.00から1.00までの間に8000語を配するわけで、親密度の精度を上げることは、分割を細かくすることに他ならず、親密度の差は大きくはならない。西出・水本 (2009) では、8000語を98段階に分けたことになる。

2. 他の「親密度」調査との比較

西出・水本 (2009) は、8000語について親密度を調査したものであるが、同じようなアンケートを用いた調査に、1000語について少人数を対象にした西出 (2007) や、同一被験者ではないものの3000語についての調査である横川ら (2006) や、240語と調査対象語数は限られているが、被験者は117人と多い西出・前田 (2008) などがある。西出・水本 (2009) が得た8000語の親密度とこれらの調査結果とを比べてみよう。

2. 1 西出 (2007) の試験的な親密度調査

頻度情報によるレベル分けを補正する拠り所として、あらかじめ計測した習熟度を用いることについては、西出・水本・前田 (2007) が示唆し、西出・水本 (2007) がやや具体的な提案をした。しかし、「水本テスト」のような4択の語彙テストをある程度の人数の同一被験者に行って、正答率を求めることは容易なことではない。西出 (2007) は、正答率測定代わりに親密度測定が使えるのではないかと考え、「知っているかどうか、知っているとするほどの程度知っているか」を問うアンケートによる調査を試験的に行ってみた。

SVL を用いた語彙サイズテストである「水本テスト」を実施した西出・水本・前田 (2007) で明らかになっているように、SVL でのレベル1, 2, 3については正答率が高い。これらのレベルの語について問えば、「知っている」との答えが返ってくるのが想定できる。また、高いレベルであるレベル7, 8については「知らない」との答えが多くなることが予想できる。西出 (2007) では、「知っている」と「知らない」が混在すると推定されるレベルの1000語ずつについて2グループで調査した。つまり、事前の「水本テスト」によって受容語彙数が5000語台と推定される受講生には、SVL のレベル5の1000語について、6000語台と推定される受講生にはレベル6の1000語について答えてもらった。

アンケートで与えた指示は、「それぞれの単語について、それぞれの単語について、知っているか、どの程度知っているかを、4段階のボックスにチェック(シ)を記入して示してください」であり、4段階は以下のように示した。

- A : 知っている (意味を言うことができる)。
- B : 見たり、聞いたりしているが、意味をはっきり言うことはできない (文脈があれば言えるかもしれない)。
- C : 見たり、聞いたりしていると思うが、意味は知らない。
- D : 知らない (見たり、聞いたりしたという記憶がない)。

西出 (2007) では、西出・水本 (2009) にある「この語を使える (文やフレーズを作ることができる)」という選択肢はない。したがって評点のスケールは5段階ではなく、4段階であった。これは、「水本テスト」によって事前に調査された語彙サイズで言えば、「知っているか、知らないか」の臨界領域のレベルについてのみ問うことになり、「使える」という項目がまず当てはまらないと考えられたからである。Aに3点、Bに2点、Cに

表 - 1 SVLのレベル5 の1000語の親密度比較

単語	西出(2007)	西出・水本(2009)
abandon	2.00	3.68
abandoned	1.82	3.45
abolish	2.00	3.45
absurd	1.09	2.41
abundance	1.55	2.86
abundant	1.64	3.14
academy	2.82	4.09
accidental	2.27	3.55
accommodation	1.64	2.86
accomplished	1.82	3.59
accomplishment	1.91	3.50
accordingly	2.00	3.91
accurately	1.45	3.64
accused	1.91	3.45
accustom	2.64	3.41
accustomed	2.18	3.18
acquaint	1.73	2.41
acquisition	1.45	2.73
acute	1.09	2.82
adventurer	2.91	4.14
adventurous	2.64	3.73
advertise	2.45	4.32
advertising	2.55	4.09
aggressive	2.82	4.18
agony	0.73	2.23
agreeable	2.36	3.73
aisle	0.82	2.86
(中 略)		
utter	1.09	3.32
utterly	0.82	3.18
vacuum	2.90	3.77
veil	1.00	2.73
veranda	1.82	4.14
verb	3.00	4.27
victorious	2.18	3.55
viewpoint	2.64	3.95
vigor	0.82	2.18
vine	1.45	3.09
violate	2.36	3.68
violently	2.64	3.59
vital	2.55	4.14
vivid	2.27	4.27
volcano	2.64	4.59
vow	1.91	2.55
waken	2.18	3.86
ward	2.09	2.82
warming	2.73	4.41
warrior	1.64	3.59
wartime	1.55	3.73
wholly	1.00	2.55
widespread	2.73	3.73
widow	1.45	3.41
wineglass	2.73	4.73
wireless	2.73	4.14
wit	1.82	3.73
witch	2.55	3.86
withdraw	1.64	3.41
woodland	1.73	2.82
workshop	2.64	3.64
worldwide	2.45	4.05
zebra	3.00	4.68
zipper	2.64	3.95

表 - 2 SVLのレベル6 の1000語の親密度比較

単語	西出(2007)	西出・水本(2009)
abortion	1.38	2.50
abrupt	1.63	2.73
accessible	2.00	2.86
accidentally	2.88	4.14
accommodate	1.63	3.18
accountant	1.63	2.64
accounting	2.50	3.59
accumulate	2.13	3.09
accuracy	1.50	3.14
accusation	1.00	2.55
ace	0.88	2.50
acknowledge	2.63	4.18
activist	1.38	2.95
adequate	2.00	3.18
adjective	2.63	3.86
administrative	2.00	2.95
administrator	2.50	3.32
admirable	2.50	3.32
admirer	2.00	3.45
adverb	1.88	3.36
advertiser	2.75	3.64
aftercare	2.88	4.18
aha	1.13	3.27
airliner	2.00	3.05
alas	0.50	2.14
alcoholic	3.00	4.18
alley	1.38	2.27
(中 略)		
villa	1.00	2.27
violation	2.25	4.00
virgin	2.38	4.14
virtual	2.63	3.86
virus	2.50	3.82
vitality	2.88	3.64
vocal	2.88	3.86
voltage	2.13	3.27
voluntary	1.88	4.05
voter	1.63	3.14
vowel	1.75	3.23
walnut	0.88	2.86
warrant	0.63	2.59
wasteful	2.50	3.64
watchdog	0.75	2.64
watchful	1.50	3.09
watchman	1.50	3.32
watercolor	1.63	3.05
waterfall	2.38	3.45
waterproof	2.88	3.95
weakly	2.75	4.05
whereas	1.75	3.23
whirl	0.63	1.68
whiteness	2.63	3.55
widen	2.88	3.91
withdrawal	1.88	2.77
woolen	0.88	3.05
workman	2.63	4.32
worldly	2.50	3.73
worship	1.38	3.50
worthless	2.63	4.14
worthwhile	1.50	3.23
wrinkle	1.25	2.68
wrongly	2.75	3.77

1点を与え、Dは0点とし、被験者の平均を求めた。被験者は、レベル5の1000語について11人、レベル6の1000語について8人と少なかった。しかし、レベル5では3点から0点まで、レベル6では3点から0.13までと評点に幅があり、同じレベルに配された語でも親密度の差はかなり見られた⁴。

西出・水本(2009)と西出(2007)で得られた親密度の相関を見るために、両者を対照表にしてみた。表-1、表-2はその一部である⁵。レベル5の1000語についての西出・水本(2009)で得られた親密度と同じ語について西出(2007)で得られた親密度の相関係数を求めてみたところ、0.821であった。また、レベル6の1000語についての西出・水本(2009)での親密度と同じ語について西出(2007)で得られた親密度の相関係数は、0.816であった。どちらも0.8以上で、強い相関を示している。

2.2 横川ら(2006)の3000語についての親密度調査

横川ら(2006)は、関西にある10の大学で学ぶ学生822名を対象にして3000語の親密度を調査している。この調査はおそらく日本で最初の親密度調査であると思われる。第二言語の獲得・処理・学習などに関する研究に取り組もうとすれば、どのような単語を使った英文を用いるかという実験要因を統制しなければならず、そのため日本人英語学習者を対象とした英単語親密度データベースを構築したとのことである⁶。

横川らの調査で対象になっているのは、BNCの出現頻度上位3000語である。これはSVLのおよそレベル1から3までに相当し、西出(2007)では、知っていることが当然であるとの判断で調査から除外した単語群である。被験者は、それぞれの単語について「全く見聞きしない」から「とてもよく見聞きをする」を1~7の7段階の数値を選ぶことによって示すようになっている。1には「全く見聞きしない」とあり、7には「とてもよく見聞きをする」とあるが、1と7の間はスケール上に2~6までの数値が与えられているだけである。7段階から一つを選ぶのは、被験者にとっては難しいこともあるのではなからうかと思うが、5段階スケールよりも評点平均値に差が出やすく、同点の語を少なくする効果があると考えられる。

横川ら(2006)での「親密度」は、「英単語を、英語として見聞きする度合い」であって、「英単語を知っている度合いではない」とされている(p.70, 79)。西出(2007)と西出・水本(2009)では、「知っている(と思っている)かどうか、知っているとすれば、どの程度知っている(と思っているか)」を問うており、被験者への問い方が異なっている。

なお、横川ら(2006)は、西出(2007)および西出・水本(2009)とは異なり、同一の語について同一の被験者による調査ではない。延べ被験者数は810名であるが、3000語を15分し、200語ずつの調査用紙を作り、それぞれについて50名強の被験者に実施している(p.76)。被験者が異なる調査で問題はないのであろうか。

横川ら(2006)と西出・水本(2009)で共通している語は、2955語あった。表-3はその一部を示したものである。相関係数を求めてみると、0.848であり、両者には強い相関が見られた。横川ら(2006)が言っている「親密度は、英単語を、英語として見聞きする

表 - 3 横川ら(2006)と西出・水本(2009)で共通する語の親密度対照表

SVL	単語	横川ら (2006)	西出・水本 (2009)
1	a	6.35	5.00
5	abandon	3.55	3.68
2	ability	5.67	4.77
1	able	6.00	4.95
1	about	6.47	5.00
1	above	4.73	4.50
2	abroad	5.88	4.91
3	absence	4.35	4.36
3	absolute	4.08	4.00
4	absolutely	4.20	4.36
4	absorb	3.39	3.77
4	abuse	2.59	2.95
3	academic	4.95	4.36
2	accept	4.71	4.36
4	acceptable	3.04	3.95
4	acceptance	3.11	3.50
3	access	5.63	4.32
2	accident	6.02	4.77
5	accommodation	2.23	2.86
3	accompany	4.06	3.55
3	according	5.02	4.86
5	accordingly	3.17	3.91
2	account	4.95	4.05
4	accurate	3.61	3.68
4	accuse	3.13	3.18
3	achieve	4.36	4.27
4	achievement	3.69	4.32
4	acid	3.29	3.95
6	acknowledge	3.42	4.18
4	acquire	4.16	3.95
5	acquisition	2.23	2.73
(中 略)			
2	worker	5.89	4.73
5	workshop	4.38	3.64
1	world	6.87	4.86
2	worried	4.98	4.77
1	worry	6.16	4.91
2	worth	4.41	4.64
1	would	5.98	4.91
3	wound	3.50	3.82
3	wrap	4.20	4.14
1	write	6.38	4.86
2	writer	6.02	4.86
2	writing	6.27	4.41
3	written	5.55	4.86
1	wrong	5.69	4.95
1	yard	5.09	4.59
2	yeah	5.51	4.68
1	year	6.68	4.95
1	yellow	6.20	4.95
1	yes	6.80	4.95
1	yesterday	6.79	5.00
2	yet	5.11	4.86
4	yield	3.24	3.64
1	you	6.92	5.00
1	young	6.71	4.95
4	youngster	3.36	3.18
1	your	6.50	5.00
1	yours	5.39	4.95
2	yourself	5.91	4.91
2	youth	5.47	4.32
3	zone	5.37	4.59

度合いであって、英単語を知っている度合いではない」ということは当たっていないのではないのではなからうか。つまり、「見聞きする度合い」をたずねても、「知っている度合い」をたずねても、単語間の関係はほぼ同じであると言えそうである。

横川ら (2006) は、各グループ50人強の15の異なるグループの810人を被験者としていたが、同一被験者22人について行った西出・水本 (2009) と相関が高いということは、被験者を統制することができれば、同一被験者について行うのと同様の信頼できる数値が得られることを示していると考えられる。また、西出・水本 (2009) の22人という被験者数も一概に少なすぎるとは言えないということになるのではなからうか。

横川ら (2006) では、Kilgarriff のレマ化されたリストをもとに、全く同じ語形で品詞が異なるものを一つにまとめて合計の頻度を算出し、その上位3000語を調査対象語にしている (p.137)。しかし、「アルファベット順親密度リスト」を見てみると、term と terms、work と works が別語として出ている。これは Kilgarriff のレマ化が不十分だったのを修正できず、引き継いでしまったからのようである。相関係数を求める際には、単数の term と work の数値を採用した。⁷

Kilgarriff List は BNC の頻度情報に基づいているので、イギリス式の綴りになっている。横川ら (2006) では、アメリカ式になおして調査を実施したとのことであるが、afterwards, catalogue, towards が使われている。check の場合、cheque も調査対象語となっている。日本人学習者には、okay よりも OK の方に親しみがあるのではないかとと思うが、okay だけが調査対象語となっている。

横川ら (2006) では3000語を対象としたとのことであるが、campaign という語が2度出現していたので、実際は2999語であったとのことである (p.138)。Kilgarriff List では、campaign は名詞と動詞で出ている。品詞が異なるものを一つにまとめる作業からこの1語が漏れてしまい、2つの問題冊子に掲載されてしまったのであろう。もし、そうであるのなら、campaign については、別の2つのグループの親密度が得られたはずである。同じ語について得られた2つの親密度の異同は貴重なデータであると思われるが、示されていない。残念なことである。

横川ら (2006) は、「教育・研究のための第二言語データベース」とのことであるが、データベースを標榜するには調査語に漏れがありすぎるように思う。SVLでレベル1に配されている145語や⁸、SVLでレベル2に配されている297語が対象語になっていない。⁹ 横川ら (2006) は am, are, is, was, were は be にレマ化され、gone は go にレマ化されている。これらは理解できるが、数詞のほとんどは対象語になっていないし、juice, soap などの生活語彙も含まれていない。これでは、教育・研究のための「データベース」としてはお粗末であると言わざるを得ない。BNCというネイティブスピーカーの使っている英語のコーパスで上位3000語を対象とするのは、日本人学習者にふさわしくないことは、SVL や JACET8000 で補正に苦労しているのを見れば明白である。

なお、横川 (2006) で3000語の対象語となっていて、SVLのレベル8までに出ていない語は、aids, allocation, alright, amongst, aye, borough, cheque, constituency, cos, depending, discourse, established, expected, fewer, firmly, growing, ha, implementation, increased, increasing, Jew, lifespan, long-term, manufacturing, no one, parish, processing, proposed, purchaser, remaining, significantly, so-called, solicitor, squad,

statutory, successfully, terms, thanks, thereby, vat, vendor, videotaped, whilst, working, works がある。-ing 形や-ed 形が目立つ。これは、動詞の変化形ではなく、形容詞とされ、別語として扱われているからである。

2.3 西出・前田 (2008) の240語についての親密度

西出・前田 (2008) は、4 択のテストで得られる正答率による習得度とアンケートで得られる親密度に相関が見られるのか、見られるとすればどの程度の相関が見られるのかを検証するために、同一被験者117人に「水本テスト」と、水本テストで用いられている240語に親密度アンケートを実施した。親密度アンケートの評点段階は西出・水本 (2009) で用いたものと同じく5段階であり、被験者への指示なども同じである。西出・前田 (2008) では、テストで得られた習得度 (正答率) とアンケートで得られた親密度 (評点平均) の相関係数は0.908であった。

西出・前田 (2008) の場合、調査対象語は240と少ないが、被験者は117名と多い。そこで、西出・前田 (2008) で得られた240語の親密度と西出・水本 (2009) での同じ240語の相関を取ってみることにする。表-4はその対照表であり、240語についての2つの親密度の相関は0.971であった。極めて強い相関を示している。なお、西出・前田 (2008) で対象語とした240語と横川ら (2006) での対象語で共通する語は、159語であった。この159語の親密度の相関係数は0.909であった。¹⁰

3. 西出・水本 (2009) で得られた親密度の内的検証

いくつかのアンケート調査で得られた親密度が同じ数値、あるいはほぼ同じ数値を出していれば、親密度は安定していることになり、アンケートによる親密度調査は信頼できることになると考えられる。西出・水本 (2009) の調査結果を西出 (2007)、横川ら (2006)、西出・前田 (2008) の調査結果と比較した結果、相関が強いことが証明され、この点では西出・水本 (2009) で得られた親密度が信頼できると言えそうである。

西出・水本 (2009) の場合、被験者数19名については、被験者の語彙サイズが事前のテストで、把握されている。もし語彙サイズの小さい被験者の親密度より語彙サイズの大きい被験者の親密度が高ければ、つまり、そのような傾向を示す調査対象語が多ければ多いほど、親密度調査そのものの妥当性が高いと言えるであろう。このほかに、被験者が対象語に与えた評点の合計で、同様の比較をすることも考えられる。テストによって語彙サイズが大きいとされた被験者が「ひかえめ」にアンケートに答える場合もあり、また語彙サイズが小さいとされた被験者が「甘め」に答える場合もあることが考えられるからである。

3.1 語彙サイズと単語の親密度

推定語彙サイズの4000語台、5000語台、6000語台、7000語台の4グループに分けみた。4000語台はA, Bの2名、5000語台はC, D, E, F, G, Hの6名、6000語台はI, J,

表 - 4 240語についての西出・前田(2008)と西出・水本(2009)の比較

SVL	単語	西出・前田 (2008)	西出・水本 (2009)
2	ability	4.74	4.77
4	accurate	3.37	3.68
3	achieve	4.28	4.27
2	actually	4.72	4.82
4	adjust	3.74	4.05
6	administrative	2.41	2.95
8	advocate	2.31	2.68
8	aesthetic	1.65	2.05
2	aim	4.27	4.36
8	allegedly	1.83	2.23
8	allocate	1.65	2.14
2	allow	4.85	4.59
4	alternative	3.43	3.59
8	analogy	2.65	3.18
3	analysis	4.13	4.23
4	ancient	4.48	4.59
4	appropriate	3.66	3.59
5	approximately	2.46	3.59
7	arctic	2.67	2.82
1	arrive	4.79	4.82
6	assess	2.69	3.32
6	asset	2.32	2.50
3	attempt	3.84	4.36
7	attorney	1.70	2.23
6	attribute	2.92	3.14
1	August	4.96	4.95
6	authorize	2.27	3.55
2	aware	4.50	4.59
5	awkward	2.85	3.14
8	barren	2.09	2.68
7	bias	2.12	3.00
(中 略)			
5	substitute	3.33	3.14
3	sufficient	3.93	3.77
3	suggestion	4.56	4.64
3	suitable	3.88	4.00
8	superficial	2.94	2.91
7	supervise	2.44	3.09
8	supposedly	2.54	3.27
1	sure	4.90	4.91
3	surface	4.52	4.32
7	surplus	2.60	3.05
4	survey	4.19	4.59
1	taste	4.70	4.73
8	testimony	1.67	2.14
7	textile	1.92	2.09
3	therefore	4.87	4.95
4	thoroughly	2.87	3.41
7	thrive	2.43	2.73
2	tiny	4.53	4.55
7	tranquil	1.62	2.14
6	transaction	2.07	3.05
6	transmit	2.89	3.18
6	tremendous	2.47	3.45
6	ultimately	3.75	3.77
1	usual	4.79	4.91
6	valid	2.58	3.00
4	vary	3.91	4.41
1	village	4.79	4.82
5	vital	3.96	4.14
8	vulnerable	1.51	2.14
1	wise	4.58	4.77

表 - 5 推定語彙レベルごとの被験者の区分

	被験者	推定語彙数	平均 推定語彙数
4000語台	A	4867	4884
	B	4900	
5000語台	C	5700	5867
	D	5767	
	E	5867	
	F	5933	
	G	5967	
	H	5967	
6000語台	I	6100	6410
	J	6167	
	K	6233	
	L	6433	
	M	6567	
	N	6567	
	O	6800	
7000語台	P	7067	7608
	Q	7700	
	R	7733	
	S	7933	

K, L, M, N, O の 7 名、7000語台は P, Q, R, S の 4 名である。表 - 5 では、各グループの平均語彙サイズも示した。

8000語のそれぞれにそれぞれのグループの評点平均値を求め、上位グループの評点平均値が下位グループの評点平均値よりも小さい場合を「×」でマークした。これを表 - 6 のようにまとめたが、ごく一部を示すに留める。¹¹ 各単語について当該グループの平均評点が直近上位グループの平均評点より低いあるいは同じであれば語彙レベルが上がるにつれて順調に習得される語であると考えられる。しかし、直近上位より高ければ、語彙の習得順にむらがある、あるいは測定そのものに問題があるということになる。「×」の数は表の欄外に示した。4000語台から5000語台へと、6000語台から7000語台へは、983 あるいは875と多いとは言えないが、5000語台から6000語台へは評点平均値が逆転している場合が4981と半数を超え、突出している。これは、グループ間の推定語数サイズの差が、4000語台から5000語台では983語、5000語台から6000語台では543語、6000語台から7000語台では1199語であり、5000語台から6000語台の差が他の2つの差と比べて小さいことが関係している可能性が考えられる。グループ間の被験者数がそろっていないことも問題であろう。

推定語彙サイズでグループ分けをするにしても、被験者数に大きな違いが出ないように3分割してみよう。表-7のように各グループを6～7人とし、各グループごとの平均評点を求め、直近上位グループの評点平均と比較した。表-8は、結果のごく一部である¹²。下位グループが上位グループより高かった場合を「×」でマークし、カウントした。×の数は、表の欄外に示した。評点平均が直近上位群より大きい場合が、1145、3852と依然として多い。

表-7 被験者数をそろえた推定語彙サイズによる3分位群

	被験者	推定語彙サイズ	平均推定語彙サイズ
語彙サイズ 下位群	A	4867	5506
	B	4900	
	C	5700	
	D	5767	
	E	5867	
	F	5933	
語彙サイズ 中位群	G	5967	6145
	H	5967	
	I	6100	
	J	6167	
	K	6233	
	L	6433	
語彙サイズ 上位群	M	6567	7195
	N	6567	
	O	6800	
	P	7067	
	Q	7700	
	R	7733	
	S	7933	

3.2 評点合計と単語の親密度

そもそも被験者は「知っているかどうか」を聞かれているわけで、被験者の語彙サイズより、被験者がアンケートに遠慮がちに答えるか、自信過剰で答えるかの方が影響しているとも考えられる。事前に測った推定語彙数より、各被験者が与えた評点の合計で下位・中位・上位を分けることも考えられる。表-9のように被験者22名を7あるいは8名の3分位群に分けたところ、下位群の最上位Cと中位群の最下位Oの間で1061ポイント、中位群の最上位Fと上位群の最下位Hの間にも1340ポイントというやや大きな差が見られた。

表 - 8 推定語彙数による3分位群の被験者の評点と評点平均値

S/V	単語	被験者																	下位グループ	判別	中位グループ	判別	上位グループ			
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q						R	S	
1	a	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00		
5	abandon	2	2	2	3	2	3	3	5	3	5	3	5	5	3	5	5	4	5	4	2.33		4.00		4.43	
5	abandoned	1	3	3	3	2	2	4	3	2	4	3	5	4	3	3	4	5	5	4	2.33		3.50		4.00	
7	abandonment	1	2	2	5	1	3	3	3	2	4	3	2	2	3	3	5	4	5	4	2.33		2.83		3.71	
7	abbreviation	3	1	2	1	1	2	2	4	1	3	2	1	1	2	1	3	4	3	5	1.67		2.17		2.71	
2	ability	5	5	5	5	4	5	5	5	5	3	5	5	5	4	5	5	5	5	5	4.83	×	4.67		4.86	
1	able	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5.00		5.00	×	4.86	
7	abnormal	3	4	3	5	4	5	5	5	4	4	4	5	5	4	5	5	4	5	4	4.00		4.50		4.57	
8	abnormality	4	1	3	1	4	4	4	3	3	4	4	3	3	4	4	3	5	4	2.83		3.50		3.71		
4	aboard	5	2	5	5	1	3	2	5	5	5	5	5	2	1	5	1	5	4	3.50		4.50	×	3.29		
5	abolish	2	3	5	3	2	2	5	5	2	4	3	2	2	4	5	3	4	5	4	2.83		3.50		3.86	
7	abolition	2	2	3	2	1	1	3	3	2	4	1	1	2	3	1	2	4	3	4	1.83		2.33		2.71	
8	aboriginal	3	5	4	3	4	1	2	5	1	5	2	1	1	3	3	2	2	4	5	3.33	×	2.67		2.86	
6	abortion	2	1	2	1	1	2	2	2	3	1	1	3	2	1	4	4	5	1.50		1.83		3.43			
8	abound	2	1	5	2	1	2	2	3	2	4	3	1	1	2	1	3	1	3	4	2.17		2.50	×	2.14	
1	about	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	above	5	3	4	5	4	5	5	4	4	5	4	4	5	3	5	5	5	5	4.33		4.33		4.71		
2	abroad	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	4	5.00		5.00	×	4.71		
6	abrupt	1	2	2	5	1	3	1	3	1	2	2	1	2	2	3	5	4	5	4	2.33	×	1.67		3.57	
3	absence	4	2	5	5	4	5	5	5	3	5	5	2	5	4	5	5	4	5	4	4.17		4.17		4.57	
2	absent	5	2	5	5	5	5	4	5	5	5	5	5	5	4	5	5	5	5	4	4.50		4.83	×	4.71	
3	absolute	1	3	3	3	2	5	5	5	3	4	5	3	5	4	5	5	4	5	5	2.83		4.17		4.71	
4	absolutely	2	3	5	3	2	5	5	5	3	5	4	5	4	4	5	5	5	5	5	3.67		4.50		4.71	
4	absorb	3	3	3	5	2	3	2	5	3	5	3	5	2	4	5	3	4	5	4	3.17		3.83		3.86	
8	absorbing	2	2	2	3	2	3	2	3	2	5	2	2	3	4	3	4	5	4	2.33		3.17		3.57		
(中 略)																										
8	wrongdoing	2	3	2	1	1	3	3	4	1	2	2	3	1	3	2	4	1	5	4	2.00		2.50		2.86	
6	wrongly	3	1	3	5	3	4	4	5	5	3	5	3	4	3	3	5	3	5	4	3.17		4.17	×	3.86	
3	yacht	5	4	2	1	1	5	5	2	1	3	1	1	5	1	1	5	4	5	4	3.00	×	2.17		3.57	
1	yard	5	1	5	5	4	5	5	5	5	5	5	5	5	4	3	5	4	5	5	4.17		5.00	×	4.43	
7	yarn	3	1	2	5	2	3	3	3	2	4	4	2	1	2	2	5	3	5	2	2.67		3.00	×	2.86	
4	yawn	3	1	2	3	1	1	2	5	2	3	2	2	1	2	2	4	2	5	4	1.83		2.67		2.86	
2	yeah	5	2	5	5	5	5	5	5	4	5	5	5	5	3	5	5	4	5	5	4.50		4.83	×	4.57	
1	year	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5.00		5.00	×	4.86	
4	yearly	3	3	2	3	2	4	4	5	3	3	5	2	2	3	3	5	4	5	4	2.83		3.67		3.71	
8	yearn	2	1	2	1	1	2	3	4	2	5	5	1	2	2	5	3	3	5	4	1.50		3.33		3.43	
3	yell	5	2	2	5	4	5	5	5	3	5	2	5	5	4	5	5	4	5	5	3.83		4.17		4.71	
1	yellow	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5.00		5.00	×	4.86	
1	yes	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.83		5.00		5.00	
1	yesterday	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
2	yet	5	5	5	5	5	5	5	5	5	5	5	5	4	4	5	5	5	5	5	5.00		5.00	×	4.71	
4	yield	3	4	2	2	4	3	5	5	1	5	3	3	2	3	5	5	4	4	4	3.00		3.67		3.86	
3	yogurt	5	1	5	1	5	5	5	3	3	5	5	5	5	4	1	5	4	5	5	3.67		4.33	×	4.14	
1	you	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	young	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5.00		5.00	×	4.86	
4	youngster	5	1	3	3	2	4	5	3	2	3	3	5	2	3	1	4	3	4	4	3.00		3.50	×	3.00	
1	your	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	yours	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5.00		5.00	×	4.86	
2	yourself	5	4	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	4.83		5.00	×	4.86	
2	yourselves	5	4	5	5	5	5	4	5	5	5	5	5	5	4	5	5	5	5	5	4.83		4.83		4.86	
2	youth	4	4	5	5	2	4	4	5	2	5	4	5	3	5	5	4	5	5	4.00		4.17		4.57		
4	youthful	3	3	3	5	1	3	4	5	2	5	4	4	3	3	4	5	4	5	4	3.00		4.00		4.00	
7	zeal	1	1	2	1	1	1	1	2	1	1	2	1	1	2	2	3	2	2	4	1.17		1.33		2.29	
7	zealous	1	2	1	1	1	1	2	3	1	3	2	1	4	1	2	2	4	1	4	1.17		2.00		2.57	
5	zebra	5	4	5	5	5	3	5	5	5	5	5	5	4	5	5	4	5	4	4.50		5.00	×	4.57		
2	zero	5	4	5	5	4	5	5	5	5	5	5	5	4	5	5	4	5	5	4.67		5.00	×	4.71		
7	zip	3	3	2	5	2	3	5	3	2	3	3	5	3	4	2	3	2	3	4	3.00		3.50	×	3.00	
5	zipper	3	4	4	5	3	4	5	3	2	4	4	5	5	2	2	4	4	5	5	3.83		3.83		3.86	
3	zone	5	4	5	5	4	5	4	5	3	5	5	5	5	3	5	5	4	5	5	4.67	×	4.50		4.57	
1	zoo	5	4	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	4.83		5.00	×	4.86	
8	zoom	4	3	4	5	2	5	5	5	3	5	4	5	4	3	4	5	4	5	4	3.83		4.50	×	4.14	
																					1145		3852			

表 - 10は、評点合計によって被験者を3分し、下位群・中位群・上位群の各単語の評点平均値の推移を見たものである。表 - 6、表 - 8と同様、結果は歴大なものになるので、ごく一部を示すにとどめる。¹³

下位群の評点平均値が中位群の評点平均値より高い場合が595、中位群の評点平均値が上位群の評点平均値より高い場合が1376であった。予想されたとおり、送転現象は大きく減少した。それぞれ8000のうちであるから、その比率は約7%、約14%である。表 - 9でわかるように、下位群の評点合計の平均と中位群の評点合計の平均の差は5318ポイントであり、中位群の評点合計の平均と上位群の評点合計の平均の差が3825ポイントである。下位群と中位群のポイント差が下位群と中位群のポイント差の約2倍ほどである。これは、各単語の群ごとの評点平均の逆転現象が下位群と中位群で少なく、中位群と上位群での逆転現象の約半分であることと符号している。

表 - 9 評点合計による3分位群

	被験者	推定語彙 サイズ	評点合計	評点合計の 平均
評点合計 下位群	B	4900	22728	25206
	E	5867	23365	
	I	6100	24367	
	A	4867	24947	
	M	6567	25208	
	N	6567	27887	
	C	5700	27939	
評点合計 中位群	O	6800	29000	30524
	L	6433	29562	
	Q	7700	29894	
	G	5967	30744	
	K	6233	31007	
	D	5767	31712	
	F	5933	31749	
評点合計 上位群	H	5967	33089	34349
	T		33135	
	U		33393	
	J	6167	33425	
	S	7933	34415	
	P	7067	35036	
	V		35561	
	R	7733	36737	

表 - 10 評点合計による3分位群の評点平均の比較

T/S	単語	被験者																		下位群	判別	中位群	判別	下位群				
		B	E	I	A	M	N	C	O	L	Q	G	K	D	F	H	T	U	J						S	P	V	R
1	a	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
5	abandon	2	2	3	2	3	5	2	5	5	4	3	3	3	4	5	3	5	4	5	5	5	2.71		3.71		4.50	
5	abandoned	3	2	2	1	3	4	3	3	5	5	4	3	3	2	4	3	5	4	4	4	4	2.57		3.57		4.13	
7	abandonment	2	1	2	1	3	2	2	3	2	4	3	3	5	3	4	3	3	4	4	4	4	1.86		3.29		4.00	
7	abbreviation	1	1	1	3	2	1	2	1	1	4	2	2	1	2	5	4	1	3	5	5	3	1.57		1.86		3.63	
2	ability	5	4	5	5	4	5	5	5	5	5	5	5	5	5	4	5	5	3	5	5	5	4.71		5.00	×	4.63	
1	able	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
7	abnormal	4	4	4	3	4	5	3	5	5	4	5	4	5	5	4	5	5	4	4	5	5	3.86		4.71	×	4.63	
8	abnormality	1	4	3	4	3	3	3	4	3	3	4	4	1	4	4	3	5	4	4	5	4	3.00		3.29		4.25	
4	aboard	2	1	5	5	2	5	5	1	5	1	2	5	5	3	5	5	5	5	4	5	5	3.57	×	3.14		4.88	
5	abolish	3	2	2	2	4	2	5	5	2	4	5	3	3	2	4	5	2	4	4	5	3	2.86		3.43		4.00	
7	abolition	2	1	2	2	3	2	3	1	1	4	3	1	2	1	4	3	2	4	4	5	2	2.14	×	1.86		3.38	
8	aboriginal	5	4	1	3	3	1	4	3	1	2	2	2	3	1	4	5	3	5	5	3	2	4.00	×	2.00		3.88	
6	abortion	1	1	2	2	2	3	2	1	1	4	2	1	1	2	4	2	3	5	5	4	5	1.86		3.00		3.75	
8	abound	1	1	2	2	2	1	5	1	1	1	2	3	2	2	4	3	3	4	4	3	3	2.00	×	1.71		3.38	
1	about	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	above	3	4	4	5	3	5	4	5	4	5	5	4	5	5	4	5	5	5	5	5	5	4.00		4.71		4.75	
2	abroad	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	4	5	5	5	4.86		5.00	×	4.88	
6	abrupt	2	1	1	1	2	2	2	3	1	4	1	2	5	3	3	3	3	2	4	5	5	1.57		2.71		3.75	
3	absence	2	4	3	4	4	5	5	5	2	4	5	5	5	5	4	5	5	5	4	5	5	3.86		4.43		4.75	
2	absent	2	5	5	5	4	5	5	5	5	5	4	5	5	5	4	5	5	4	5	5	5	4.43		4.86	×	4.75	
3	absolute	3	2	3	1	4	5	3	5	3	4	5	5	3	5	5	5	3	4	5	5	5	3.00		4.29		4.63	
4	absolutely	3	2	3	2	4	4	5	5	5	5	4	5	5	4	5	5	5	5	5	5	5	3.29		4.86		4.88	
4	absorb	3	2	3	3	4	2	3	5	5	4	2	3	5	3	4	5	5	4	5	3	5	2.86		3.86		4.50	
8	absorbing	2	2	2	2	3	2	2	4	5	4	2	2	3	3	4	3	5	4	3	3	5	2.14		3.29		4.00	
(中 略)																												
8	wrongdoing	3	1	1	2	3	1	2	2	3	1	3	2	1	3	4	4	5	2	4	3	4	5	1.86		2.14		3.88
6	wrongly	1	3	5	3	3	4	3	3	3	3	4	5	5	4	4	5	3	3	4	5	5	3.14		3.86		4.25	
3	yacht	4	1	1	5	1	5	2	1	1	4	5	1	1	5	4	2	5	3	4	5	5	2.71	×	2.57		4.13	
1	yard	1	4	5	5	4	5	5	3	5	4	5	5	5	5	5	5	5	5	5	5	5	4.14		4.57		5.00	
7	yarn	1	2	3	2	1	2	2	2	3	3	4	5	3	3	3	2	4	2	5	5	5	1.86		3.14		3.63	
4	yawn	1	1	2	3	2	1	2	2	2	2	2	3	1	4	5	5	3	4	3	4	5	1.71		2.00		4.13	
2	yeah	2	5	4	5	3	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	4.14		4.86		5.00	
1	year	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
4	yearly	3	2	3	3	3	2	2	3	2	4	4	5	3	4	4	5	5	3	4	5	5	2.57		3.57		4.50	
8	yearn	1	1	2	2	2	2	2	5	1	3	3	5	1	2	4	4	5	4	5	3	5	1.71		2.86		4.38	
3	yell	2	4	3	5	4	5	2	5	5	4	5	2	5	5	4	5	5	5	5	5	5	3.57		4.43		4.88	
1	yellow	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
1	yes	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
1	yesterday	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
2	yet	5	5	5	5	4	4	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	4.71		5.00	×	4.88	
4	yield	4	4	1	3	3	2	2	5	3	4	5	3	2	3	4	5	5	5	4	4	5	2.71		3.57		4.50	
3	yogurt	1	5	3	5	4	5	5	1	5	4	5	5	1	5	5	3	5	5	5	5	5	4.00	×	3.71		4.75	
1	you	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	young	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
4	youngster	1	2	2	5	3	2	3	1	5	3	5	3	3	4	4	3	3	3	4	3	4	2.57		3.43		3.50	
1	your	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5.00		5.00		5.00	
1	yours	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.86		5.00		5.00	
2	yourself	4	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.71		5.00		5.00	
2	yourselves	4	5	5	5	4	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	4.71		4.86		5.00	
2	youth	4	2	2	4	3	5	5	5	4	4	4	4	5	4	4	5	5	5	5	5	5	3.57		4.43		4.88	
4	youthful	3	1	2	3	3	3	4	4	4	4	4	5	3	4	5	5	4	4	5	5	5	2.57		4.00		4.63	
7	zeal	1	1	1	1	2	1	2	2	1	2	1	2	1	1	3	2	1	1	4	3	3	1.29		1.43		2.38	
7	zealous	2	1	1	1	1	4	1	2	1	4	2	2	1	1	4	3	1	3	4	3	2	1.57		1.86		2.63	
5	zebra	4	5	5	5	4	5	5	5	5	4	5	5	5	3	4	5	5	4	5	5	5	4.71	×	4.57		4.75	
2	zero	4	4	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.57		4.86		5.00	
7	zip	3	2	2	3	4	3	2	2	5	2	5	3	5	3	4	3	3	3	4	5	3	2.71		3.57	×	3.50	
5	zipper	4	3	2	3	2	5	4	2	5	4	5	4	5	4	4	3	5	4	5	5	4	3.29		4.14		4.38	
3	zone	4	4	3	5	3	5	5	5	5	4	4	5	5	5	4	5	5	5	5	5	5	4.14		4.71		4.88	
1	zoo	4	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4.71		5.00		5.00	
8	zoom	3	2	3	4	3	4	4	4	5	4	5	4	5	5	4	5	5	4	5	5	5	3.29		4.57		4.75	
																									595		1376	

4. まとめと今後の課題

西出・水本（2009）で発表した8000語についての親密度測定が妥当なものかどうかについて検証を試みた。

西出（2007）での試験的な親密度調査、やや別の観点から行われた横川ら（2006）による親密度調査、親密度と語彙テストの正答率の相関を見た西出・前田（2008）で得られた親密度と比較した。比較対象になった調査は、被験者数が少ない調査であったり、被験者の知っているかどうかの臨界域にある語に限った調査であったり、被験者が異なった調査であったり、語数は限られているが被験者数が、語数は限られているが被験者数が多かった調査であったりと、いくつかの点で西出・水本（2009）とは異なっていたが、共通して調査された語について親密度の相関係数を求めると、いずれも0.8以上であった。被験者にスケール上の数値を選ばせるというアンケートという調査方法を用いたのであるが、安定した数値が得られていると言えるのではないだろうか。

さらに、水本テストで推定された語彙数や、親密度調査のアンケートで被験者が与えた評点総計などで被験者をグループ分けし、下位グループから上位グループで親密度が上がっているかどうかを見た。残念ながら、下位グループの親密度が上位グループの親密度より高いというケースも見られた。しかし、語彙サイズごとのグループ分けより、評点合計によるグループ分けの方が下位グループの親密度と上位グループの親密度の逆転のケースは少ないことがわかった。

西出・水本（2009）での被験者は22名であった。これはやはり少ないのではないだろうか。どれぐらいの被験者が必要なのかを見極めるためにも同様の調査をし続ける必要がある。と同時に、西出・水本（2009）は安定した数値を得ているとも考えられる。少なくとも、今回の考察では、信頼できない数値であるということにはならなかった。8000語の親密度を用いて新たな語彙リストを作成したり、これまでに提案されている語彙リストを親密度の観点から検討したりすることが望まれる。

注

- 1 この表は、エクセルファイルの形でも提供している。nishide@tsuru.ac.jp まで請求いただきたい。
- 2 下位になると同位の語数は少なくなる。7990位（親密度1.59）の単語が5語、7995位（親密度1.55）の単語が3語、7998位（親密度1.50）の単語が2語、8000位（親密度1.41）の単語は hitherto 一語であった。
- 3 厳密な1000語ずつのレベル付けはできないが、以下のように約1000語のレベル付けはできる。

段階	評点	順位	語数	レベル	語数	段階	評点	順位	語数	レベル	語数		
1	5.00	1	143	1	1065	50	3.29	5571	1	6	975		
2	4.95	144	309			51	3.27	5572	101				
3	4.91	453	229			52	3.24	5673	2				
4	4.86	682	196			53	3.23	5675	102				
5	4.82	878	188			54	3.18	5777	119				
6	4.77	1066	191	2	1002	55	3.14	5896	94				
7	4.73	1257	165			56	3.09	5990	95				
8	4.68	1422	181			57	3.05	6085	113				
9	4.64	1603	154			58	3.00	6198	93				
10	4.59	1757	153			59	2.95	6291	99				
11	4.55	1910	158			60	2.91	6390	92				
12	4.50	2068	125	3	958	61	2.86	6482	95				
13	4.45	2193	144			62	2.82	6577	101				
14	4.41	2337	135			63	2.81	6678	1				
15	4.38	2472	2			64	2.77	6679	88				
16	4.36	2474	131			65	2.76	6767	1				
17	4.33	2605	1			66	2.73	6768	78				
18	4.32	2606	146			67	2.71	6846	2				
19	4.27	2752	130			68	2.68	6848	107				
20	4.24	2882	1			69	2.67	6955	3				
21	4.23	2883	143			70	2.64	6958	83				
22	4.18	3026	154			4	1129	71	2.62	7041	1		
23	4.14	3180	167	72	2.59			7042	92				
24	4.09	3347	143	73	2.55			7134	85				
25	4.05	3490	130	74	2.50			7219	68				
26	4.00	3620	120	75	2.45			7287	78				
27	3.95	3740	133	76	2.41			7365	68				
28	3.91	3873	143	77	2.36			7433	70				
29	3.86	4016	139	78	2.32			7503	50				
30	3.82	4155	118	79	2.29			7553	1				
31	3.81	4273	1	80	2.27			7554	51				
32	3.77	4274	99	81	2.23			7605	52				
33	3.73	4373	116	82	2.18			7657	57				
34	3.71	4489	1	83	2.14			7714	54				
35	3.68	4490	125	84	2.09			7768	44				
36	3.67	4615	1	85	2.05	7812	35						
37	3.64	4616	139	86	2.00	7847	28						
38	3.62	4755	1	87	1.95	7875	22						
39	3.59	4756	128	88	1.91	7897	23						
40	3.57	4884	1	89	1.86	7920	14						
41	3.55	4885	119	90	1.82	7934	11						
42	3.52	5004	1	91	1.77	7945	21						
43	3.50	5005	105	92	1.73	7966	14						
44	3.48	5110	1	5	955	93	1.68	7980	4				
45	3.45	5111	117			94	1.64	7984	6				
46	3.41	5228	112			95	1.59	7990	5				
47	3.38	5340	2			96	1.55	7995	3				
48	3.36	5342	118			97	1.50	7998	2				
49	3.32	5460	111			98	1.41	8000	1				
				6	975					8	960		

4 西出 (2007) の付表 4 - 2 (p.100) に欠落があった。親密度2.88の語群と親密度2.63の語群の間に、以下の親密度2.75の65語を入れていただきたい。(なお、エクセルファイルで提供した表には欠落はない。)

2.75 [23/8] advertiser, architectural, athletics, brightness, cheapness, clearness, closing, colored, compass, comprehend, coordinate, cosmetic, database, defender, demonstrator, dental, depress, descent, embarrassment, entertaining, excellence, feminist, gene, importantly, impossibility, incapable, indication, indifference, instrumental, Islam, isolate, loser, loudspeaker, mammal, mathematician, mini, motorcar, murderer, nationalist, navigation, opener, orient, overwhelm, participation, perfection, plantation, poetic, popularly, qualification, recreation, refreshment, regularity, representation, respectful, sensitivity, sharpen, succession, trash, unreasonable, unwise, usefulness, vacant, viewer, weakly, wrongly

- 5 紙数の都合から一部しか示すことができない。都合2000語のデータは、西出・水本（2009）と西出（2007）から作成は可能であるが、エクセルファイルで提供する。
- 6 横川ら（2009:ii）。
- 7 term の平均評点は4.89、terms の評点平均は4.54であり、work の平均評点は6.49、works の評点平均は6.14となっている。
- 8 レベル1の145語とは、airplane, am, apartment, are, aunt, bake, baker, banana, baseball, basket, basketball, bat, beef, been, bicycle, blackboard, butter, bye, cafe, calendar, candy, chalk, children, Christmas, classmate, cloudy, cookie, dad, daddy, delicious, did, does, doll, eight, eighteen, eighty, elephant, eleven, fifteen, fifty, five, fork, forty, four, fourteen, fox, frog, glove, gone, goodbye, grandfather, grandmother, happiness, has, hi, hobby, homework, honey, hundred, hungry, is, jam, juice, lion, men, midnight, Miss, monkey, Mr., Mrs., Ms., musician, nail, nine, nineteen, ninety, noon, notebook, o'clock, oh, ours, pants, pencil, pet, piano, picnic, pig, pork, rabbit, rainy, rice, robot, rocket, rope, salad, sandwich, seven, seventeen, seventy, shower, singer, six, sixteen, sixty, ski, skiing, sleepy, smoker, smoking, snake, soap, soccer, soda, soup, spoon, subway, supermarket, taxi, ten, tent, textbook, theirs, third, thirteen, thirty, thousand, three, tiger, toe, toilet, tomato, towel, toy, twelve, twenty two, umbrella, uncle, vacation, videotape, violin, was, were, worst, zoo である。
- 9 レベル2の297語とは、absent, accent, actress, admire, adventure, ah, airline, almond, alphabet, angel, ant, apron, automobile, awake, backward, bacon, balloon, bay, bean, beard, beautifully, bee, blanket, boil, bookshop, bookstore, boring, boyfriend, brake, brave, broadcast, bush, businessman, butterfly, cab, cable, cage, campus, candle, carpenter, cent, centimeter, cheer, cherry, cigar, cinema, cleaning, clerk, cloth, cocoa, colorful, comic, cooking, corn, cotton, courage, cowboy, crazy, crowded, dancer, dancing, dentist, dial, diamond, dislike, downstairs, downtown, dozen, drill, drinking, drown, drunk, earring, educate, eighteenth, eighth, eightieth, elbow, elder,

elevator, eleventh, ending, envelope, eraser, everyday, exact, exit, fare, fax, fifteenth, fifth, fiftieth, fisherman, fishing, flag, float, foolish, foreigner, forever, forgive, fortieth, fortunate, fourteenth, fourth, ghost, girlfriend, gorilla, granddaughter, grandma, grandpa, grandson, grape, grapefruit, greatness, greeting, gym, hairbrush, haircut, ham, hamburger, hammer, handkerchief, handsome, happily, harbor, headache, helicopter, hey, hiking, homeroom, hook, ink, insect, intelligent, jazz, jeans, jet, jungle, kilo, kilogram, kilometer, lamp, lazy, lemon, locker, lonely, mama, mankind, mate, mathematics, medal, melon, melt, menu, merry, mild, millimeter, million, monster, monthly, mosquito, mud, navy, neat, necklace, needle, nephew, nest, niece, nineteenth, ninetieth, ninth, noble, noisy, nut, ocean, oneself, onion, oven, painful, painter, pajamas, pan, panda, papa, parade, pardon, passport, pea, peaceful, peach, peanut, pearl, pepper, photographer, pie, pillow, pineapple, pizza, poison, polite, pond, powder, precious, princess, professor, punish, purple, purse, puzzle, quiz, racket, railroad, rainbow, rat, reporter, ribbon, ruler, rumor, sailing, sailor, sauce, saying, schoolboy, schoolgirl, seventeenth, seventh, seventieth, shade, shampoo, shore, shrimp, sixteenth, sixth, sixtieth, skate, skating, skirt, sock, sofa, spaghetti, speaking, spelling, spider, spoken, steak, stereo, straw, strawberry, suitcase, sundae, sunlight, sunset, sunshine, super, supper, swallow, swan, sweat, sweater, teen, teenage, teenager, temple, tenth, thief, thirsty, thirteenth, thirtieth, thumb, tire, tobacco, toothbrush, toothpaste, trick, truck, trumpet, tulip, tummy, turkey, twelfth, twentieth, underwear, unhappy, violet, vocabulary, volleyball, waiter, waitress, walker, weekly, whiskey, wolf, wool, worse, yourselves, zero **である。**

- 10 **西出・前田 (2008: 75-6)。**
- 11 **すべてをエクセルファイルで提供するので、nishide@tsuru.ac.jpまで請求いただきたい。**
- 12 **エクセルファイルですべてを提供する。**
- 13 **エクセルファイルですべてを提供する。**

参考文献

- ALC (2001) "Standard Vocabulary List 12000," [http://www.alc.co.jp/goi/PW_top_all.htm]
Kilgarriff, A. "BNC database and word frequency lists." [http://www.kilgarriff.co.uk/bnc-readme.html]
- 西出公之 (2007) 「『親密度』に基づく語彙リストの可能性」『都留文科大学研究紀要』第66集 pp.85-106.
- 西出公之・水本篤 (2007) 「『語彙サイズテスト』の語彙表とサンプリング」『都留文科大学大学院紀要』第11集 pp.21-31.
- 西出公之・水本篤 (2008) 「『習得特性』に基づく語彙リストの可能性」『都留文科大学大学院紀要』第12集 pp.31-46.

- 西出公之・水本篤 (2009) 「英単語8000語についての親密度測定の試み」『都留文科大学大学院紀要』第13集 pp.57-92.
- 西出公之・水本篤・前田浩 (2007) 「『語彙サイズテスト』による語彙数推定の試み」『都留文科大学研究紀要』第65集 pp.61-77.
- 西出公之・水本篤・前田浩 (2008) 「『語彙サイズテスト』による経年的語彙数推定の試み」『都留文科大学研究紀要』第67集 pp.47-58.
- 西出公之・前田浩 (2008) 「語彙リストのための親密度測定と習熟度測定」『都留文科大学研究紀要』第68集 pp.63-79.
- 横川博一編 (2006) 『日本人英語学習者の英語親密度 文字編 教育・研究のための第二言語データベース』くろしお出版.