

# 日本人大学生の語彙知識と 英語母語話者の推定語彙知識との比較

吉井 誠

## 1. はじめに

日本人大学生、特に英文科の学生の英語語彙レベルはどれくらいであろうか。1年間の学びを通してどのような語彙の知識が増えていくのであろうか。また、日本人大学生の語彙力は英語母語話者、特に同年代の大学生と比べてどのような差があるのだろうか。このような疑問からこの研究は始まった。

日本人大学生、特に英語英米文学を専攻とする学生の英語力はどのように変化していくのかについて調査を行ってきた(吉井:2017, Yoshii & Lavin, 2017)。1年生の後期に“Extensive Reading & Listening I”そして2年生の前期に“Extensive Reading & Listening II”を開講しており、これらのクラスの中で実施してきた。吉井(2017)ではこの科目の1年生の授業の中で事前事後テストを実施し、英語力の伸びをTOEICの模擬テスト縮小版(30分50問)を用いて計測した。その結果、1学期間に読解力、聴解力ともに有意な伸びを示しており、英語力の向上が確認できた。また、それに加え、学期当初と学期終了間際に計測した読解速度においても有意な差が検出され、速度も向上したことが分かった。これを受け、Yoshii & Lavin(2017)では、2年生の授業の中でテストを実施し、1年次に受けたテストと比較し、1年間の学習における英語力の変化を観察した。その結果、1年次の後期と2年次の前期の間の春休み期間に英語力が落ちる傾向が見られたが、1年間を通してみた場合、英語力の確かな伸びを確認することができた。

ここまではTOEICの模擬テスト縮小版を用いて英語力全般の変化について観察してきたが、語彙知識はどのように変化しているのか疑問が残った。そこで本研究では、日本人大学生の語彙知識が英文科での1年間の学びを通してどのように変化していくのか調べることにした。

また、本研究では、大学生の語彙レベルをより客観的に知るために、英

語母語話者の大学生の語彙レベルとの比較を行う。ただし、英語母語話者  
に実際に語彙テストを受けてもらうことは困難であり、ここでは、潜在  
意味解析という統計手法を用いて母語話者の語彙レベルを推定し比較を行  
う。

ここで簡潔に潜在意味解析について説明をしておきたい。なお詳細につ  
いては猪原・楠見 (2012)、吉井 (2019) を参照されたい。潜在意味解析 (潜  
在意味分析とも呼ばれる) とは英語で Latent Semantic Analysis であり  
(略して LSA、以後本論文では LSA の略称を使用)、単語・語句・文と他の  
単語・語句・文との関連性を統計的に算出する方法である。関連性は意味  
の類似度として数値で表され、その数値が高い場合は、強い関連性を示唆  
しており、逆に低い場合は関連性が弱いと判断する。そしてこの統計分析  
の材料として、参加者が遭遇するであろうと推定されるテキスト (本、雑誌、  
論文など様々な読書物) の集まりであるコーパスを利用する。分析の手續  
きの詳細については猪原・楠見 (2012 : 102-104)、や吉井 (2019 : 46-47)  
を参照されたい。本研究では、コロラド大学の LSA のウェブサイト (Latent  
Semantic Analysis @ CU Boulder, <http://lsa.colorado.edu/>) を使用  
して分析を行う。このサイトには大学 1 年生の英語母語話者コーパス (大  
学生になるまでに遭遇するとされる読書物、テキストで構成されたもの)  
が含まれており、それに基づいた分析が可能である。なお、このサイトの  
利用方法に関しては Dennis (2007) に詳細が記されている。

本研究のように学習者のデータを LSA でシミュレーションした代表的  
なものに Landauer & Dumais (1997)、Biemiller, Rosenstein, Sparks,  
Landauer & Foltz (2014) などがある。Landauer & Dumais (1997) では  
英語を外国語として学んでいる留学生に TOEFL を受けてもらい、その点数  
と LSA から推定される得点を比べている。その結果、推定値と実際の学習  
者の数値はほぼ同じであった。Biemiller, et al. (2014) の研究では、  
英語母語話者の子供たちの語彙の発達を LSA を用いてシミュレーション  
し、実際に子供たちに実施した語彙テストの結果と比較している。その結  
果は相関係数で .67 から .74 という高い数値であり LSA が妥当であるこ  
とを証明している。このように LSA を用いて TOEFL のスコアまたは語彙テ  
ストを予測することは可能であり、本研究ではこれを使用し、英語母語話  
者の語彙レベルを推定する。

## 2. 本研究の目的と研究課題

背景のところで述べたように、これまで大学生の英語力の伸びを TOEIC を通して観察してきたが、本研究では語彙知識の変化に焦点を絞って調査を行う。英文科における1年間の学びを通して学生の語彙レベルはどのように変化するのか、また、その語彙レベルは同年代の英語母語話者の語彙レベルと比べてどのように違うのか調べる。具体的な研究課題は次の通りである。

1. 英文専攻の日本人大学生の語彙知識は1年間の学びを通してどのように変化するであろうか。Vocabulary Levels Test においてどのレベルの語彙がどのように増えたのかを検証する。
2. 日本人大学生の語彙知識は英語母語話者の大学生の語彙知識と比べてどのような違いがあるだろうか？日本人大学生の Vocabulary Levels Test の結果（2年次の学期末のテストの結果）とLSAより推定される英語母語話者の大学生の Vocabulary Levels Test の結果とを比較する。

## 3. 研究のデザイン

### 3.1 参加者

本研究の参加者は文学部英語英米文学科の大学2年生であり、1年後期（“Extensive Reading & Listening I”）と2年前期（“Extensive Reading & Listening II”）に多読多聴を促進するクラスを受講していた。この他にも英語運用科目として Writing や Reading のクラス、4技能を統合的に訓練するクラス、文法、音声学、講読や文学史など様々な英語関連のクラスを受講している。前述の二つのクラスを受講者のうち、事前・事後テスト、両方を受験した34名が本研究の対象者となった。

### 3.2 語彙レベル判断テスト

語彙力測定に語彙レベルごとの知識を測定できる Vocabulary Levels Test（略してVLT、今後論文の中でもこの表記を使用）を用いる。VLTはNation（1990, 2001）によって開発され、その後いくつかの版を重ねながら今日に至っている。当初は2000語、3000語、5000語そして10000語の4つのレベルからなり、一つのレベルに18題の問題を含んでいた。本研究ではこの4つのレベルに英語圏の大学でよく使用される単語のレベル、Academic Word Level（Coxhead, 2000）を追加し、レベルの問題数

を30題とした版 (Schmitt, 2000; Schmitt, Schmitt, & Clapham, 2001) を使用する。なお VLT では各レベルに次のような基礎データベースが使用されている。2000 語は英語の基本語と言われている General Service List (West, 1953), 3000 語以上は Thorndike & Lorge (1944) を Kucera & Francis (1967) の頻度情報をもとに修正を加えたものを使用している。そして前述した Academic Word Level は Coxhead (2000) により開発された大学などのアカデミックな場でよく使用される語彙をまとめた Academic Word List が元となっている。1つのレベルには表1のような問題のセットが10セット用意されている。表1に示されているように左側に6つの単語が、右側には3つの語句が提示されており、学習者は3つの語句に当てはまる単語を左側から選ぶようになっている。各レベル30項目、全体では5レベル、すなわち150項目から構成されていた。

表1 VLT2000 語レベルからの問題例

- |              |              |
|--------------|--------------|
| 1. birth     |              |
| 2. dust      | _ game       |
| 3. operation | _ winning    |
| 4. row       | _ being born |
| 5. sport     |              |
| 6. victory   |              |

### 3.3 研究の手順

参加者は日本人大学生で1年後期の始めに VLT を用いて事前テストを受験した。図1に示されているように VLT は Google Form 上で実施した。参加者は図1の上の行に示されている単語の中から左に掲載されている語句の意味に一番近いものを選んだ。各レベルに図1のような問題が30問あり、参加者は、5つのレベル(合計150問)に回答した。翌年の2年次前期の授業の最終週、7月末に事前テストと同じものを使用して事後テストを実施した。事前テストと事後テストには1年間という期間があり、問題数も150問と多数であり、テスト効果(同じテストを2度受けることによる学習効果)はほとんど影響がないと判断したうえで、このような形で事前・事後テストを行った。

**Vocabulary Levels Test**

This is a vocabulary test. You must choose the right word to go with each meaning. Match the word or phrase in the column with the word in the row. Here is an example:

**Sample question**

	business	clock	horse	pencil	shoe	wall
part of a house	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
animal with four legs	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
something used for writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Some words are in the test to make it more difficult. You do not have to find a meaning for these words. In the example above, these words are business, clock, and shoe.

If you have no idea about the meaning of a word, do not guess. In that case, you do not need to supply an answer. But if you think you might know the meaning, try to choose the answer.

図1 Formを使用したVLTの問題例

### 3.4 分析方法

事前・事後テストの比較は2000語、AWL、3000語、5000語、そして1万語という各レベルごとに、対応のあるt検定を実施し比較する。英語母語話者と比較するために、最初にLSAを用いて母語話者の大学生の語彙レベルを推定した。

本研究では、英語母語話者の大学1年生の語彙レベルを測定するために、コロラド大学のLSAサイトで、1年生のコーパス（“General\_Reading\_up\_to\_1st\_year\_college”）を利用してVLTの各問題を分析した。表2はVLT2000語レベルの問題をLSA分析した結果を示している。各単語と3つの語句の関連性を数値で表しており、3つの語句の列の中で1番数値が高いものを1番関連性が高いものと判断している。表2では“game”には“sport”を“winning”では“victory”を“being born”では“birth”

をそれぞれ1番高い数値として検出しており、3つとも正解であり、LSAで正しい答えが導き出せている。このようにしてVLTの各問題を分析して行った。各レベルに10セットあり、表2のような分析を合計で50セット実施した。

表2 VLT2000 語レベルの問題例のLSA分析結果(1)

Texts	game	winning	being born
birth	0.02	0.05	0.45
dust	0.03	0.04	0.06
operation	0.01	0.04	0.03
row	0.23	0.17	-0.01
sport	0.73	0.51	0.06
victory	0.10	0.55	-0.06

日本人大学生の事後テストの結果とLSAで推定された英語母語話者の結果の比較にはカイ二乗検定を使用した。事後テストは34名の参加者のデータからなり、そこから平均値を用いることができるが、LSAでは表2にあるようにWebサイトで分析された数値のみを用いていく。この二つの異なる性質の数値を比較するために、本研究では、LSAで分析される数値を正しい答えに導き出された頻度と解釈し、それに対応する形で、日本人学習者の正解の平均値も同様に正解であった頻度と位置付けて分析した。

#### 4. 結果

##### 研究課題1: 日本人学生の1年間の語彙知識の変化

図2に各レベルの事前テストと事後テストの結果が視覚的に表されている。全体的には一つのレベルAWLを除き、事前テストより事後テストで得点が伸びていることが分かる。表3にあるようにt検定をした結果、統計的に有意な差が検出されたのは、2000語、5000語そして10000語レベルであった。全体的に語彙のレベルは上がっており、とくに5000語や10000語など難易度の高い語彙にあきらかな変化がみられた。

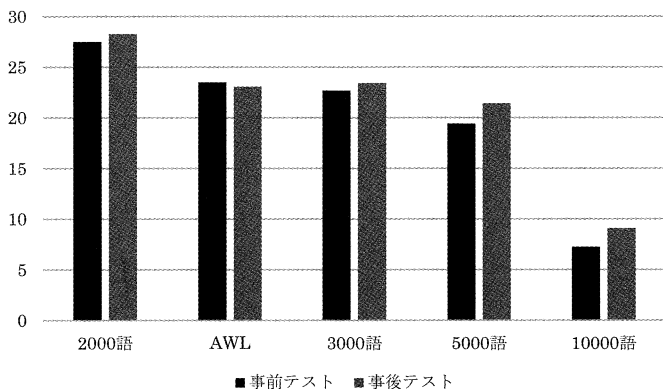


図2 各レベルの事前テストと事後テストの結果

表3 各レベルの事前事後テストの記述統計と *t* 検定の結果

	2000	AWL	3000	5000	10000
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
事前	27.5 (1.96)	23.5 (3.59)	22.68 (3.82)	19.41 (3.97)	7.24 (4.06)
事後	28.26 (1.48)	23.06 (3.46)	23.41 (3.66)	21.41 (3.13)	9.01 (3.40)
<i>t</i> 検定	<i>t</i> = 2.86, <i>df</i> = 33, <i>p</i> = .007,	<i>t</i> = .76, <i>df</i> = 33, <i>p</i> = .45,	<i>t</i> = 1.42, <i>df</i> = 33, <i>p</i> = .17,	<i>t</i> = 3.86, <i>df</i> = 33, <i>p</i> = .000,	<i>t</i> = 2.60, <i>df</i> = 33, <i>p</i> = .01,
効果量	<i>r</i> = .45	<i>r</i> = .13	<i>r</i> = .24	<i>r</i> = .56	<i>r</i> = .41

## 研究課題 2 : LSA で推定される英語母語話者のデータと日本人学生のデータとの比較

英語母語話者の推定値と日本人学習者の実測値を比べた結果が図3ならびに表4で示されている。視覚的に見ると予想と反して基本的な語彙のレベル2000語とAWLにおいて英語母語話者 (Native) の推定値よりも日本人学習者の数値の方が高い結果となった。しかし、語彙レベルが上がるにつれて母語話者の数値が上回っている。

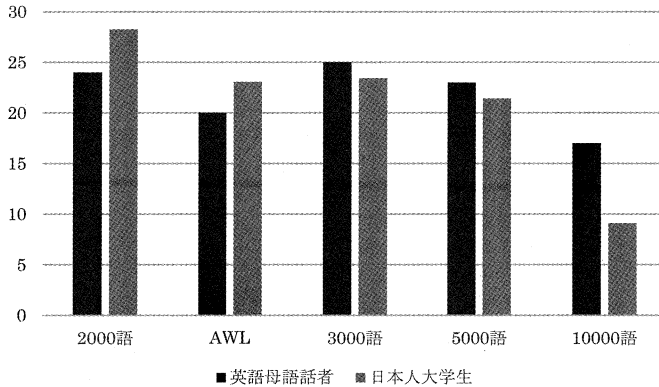


図3 英語母語話者の推定値と日本人学生の実測値との比較

表4 英語母語話者の推定値と日本人学生の実測値

	2000 語	AWL	3000 語	5000 語	10000 語
英語母語話者	24	20	25	23	17
日本人学生	28	23	23	21	9

表4の結果をカイ二乗検定で分析したところ、 $\chi^2 = 3.04$ ,  $df = 4$ , n.s. という結果であり、二つのグループの得点のパターンには有意な差は見られなかった。このことより英語母語話者と日本人学生の語彙知識には大きな差はみられないとの結果となった。

## 5. 考察

### 5.1 日本人大学生の語彙力の変化

表3と図2で示されていたよう日本人大学生の語彙レベルは、AWLを除き、すべてのレベルで伸びていることを確認することができた。最も基本的な2000語レベルにおいては事前テストでも30問中に28問正解とかなり高い正解率を得ており、それだけに事後テストの伸びしろは、あまり残されていなかったが、それにもかかわらず得点が伸びていたことは大きい。また、語彙知識の伸びは特に5000語、10000語で顕著であった。様々な英語関連の科目を受講する中で、難度の高い語彙も含め様々なレベルの単語に出会っていることがうかがえる。意外な現象としてはAWLで伸びが見られなかったことだ。大学でアカデミックな英語に触れAWLも伸びることが



予想されたが、それほど伸びはみられなかった。これは AWL のセクションに登場する単語や語句に特異なものがあり、それが正解率に影響したのか、それとも現存のカリキュラムの中で遭遇している読み物、テキスト類では AWL の単語に十分に遭遇していないのか調べる必要がある。

## 5.2 日本人大学生と母語話者との比較

2000 語レベル、AWL においては英語母語話者のデータより非母語話者の日本人学生のデータの方が上回っていた。幾つかの理由が考えられるが、その一つとして LSA は実際の参加者のデータを過小評価する可能性がある事が考えられる。実際の参加者は VLT のような多岐選択問題に対して、単語や語句の意味を推測したり、選択肢を俯瞰したり、既知の単語を手がかりに消去法などを駆使したりしながら答えを探している。しかし、LSA では純粋に数値で表された関連度のみで判断を行う。そこには推測や消去法などが入り込む余地はない。表 5 には VLT2000 語レベルの問題を LSA で分析した結果が示されている。表 2 の時のように、LSA が示す 1 番高い数値がそのまま正解であるパターンから逸脱した結果となっている。表 5 で示されているように “pleasing quality” という語句に対して LSA では “lack” が 1 番高い数値となりそれを正解として判断している。しかし実際の正解は “charm” であり、これは LSA では 2 番目に高い数値であるが LSA では不正解である。実際の参加者の場合、1 番目と 2 番目の候補がある場合は、どちらがより妥当であるか、これまでの知識を総動員して判断するであろう。また “not having something” という語句に対しては、正解は “lack” であるが、LSA ではその正解にはたどり着かず、その代わり “charm” と “treasure” が同じ数値で 1 位に並んでいる。大学 1 年生までのコーパスはそれまでの英語母語話者の読書経験を代表するものであり、その豊富な種類と量から推測して、なぜこのような結果になるのかはまだ不明である。

表5 VLT2000 語レベルの問題例のLSA分析結果(2)

Texts	gold and silver	pleasing quality	not having something
attack	-0.02	-0.01	0.01
charm	0.07	0.07	0.07 ×
lack	-0.07	0.14 ×	-0.01
pen	-0.05	0.10	-0.02
shadow	0.02	0.04	-0.02
treasure	0.60	0.01	0.07 ×

また、以下の表6のようなケースもある。VLT2000 語レベルの別の問題をLSAで分析した結果が示されている。ここでは、“commonly done”という語句に対して、“brave”と“usual”が同値で並んでいる。このような場合、実際の学習者の場合は、上述したように二つの選択肢を比べ、どちらがより該当する語句の内容に近いかを模索する。それに加え、他の問題との兼ね合いも考慮に入れるであろう。すなわち、“having no fear”という語句に対してはLSAの数値も物語っているように、“brave”という単語を選ぶ可能性が高い。そうすると、先程の“commonly done”という語句に対しては、“brave”という選択肢が除かれ、残った“usual”を正解として選択する事が予想される。このように実際の参加者は他の回答との兼ね合いも考慮に入れながら消去法も用い答える事が想像される。それに比べてLSAでは、単に単語と語句の数値のみに基づいて判断していくために、正解に行き着く術が限られており、この実験の結果のように、実際の参加者よりも回答率が低くなってしまふ。

表 6 VLT2000 語レベルの問題例の LSA 分析結果 (3)

Texts	commonly done	wanting food	having no fear
brave	0.02 $\Delta$	-0.00	0.14
electric	-0.00	0.01	0.00
firm	0.01	0.02	-0.04
hungry	-0.15	0.41	0.11
local	0.00	-0.00	-0.04
usual	0.02 $\Delta$	-0.04	0.11

このような LSA の限界を踏まえながらも、それでも母語話者の強みははっきりと表れたのは 3000 語以上のレベルであった。このあたりのレベルになると日本人大学生のそもそもの語彙知識がおぼつかなくなり、先に述べたような総合的な知識を総動員しても、また消去法などのテスト受験方略を用いても、LSA で示される母語話者の単語知識との差が明らかになっていく。特に 10000 語ではそれが顕著である。

## 6. 結論と今後の研究の課題

この研究から大学 1 年生は日ごろの英文科での学びを通して語彙レベル、AWL を除いて全てのレベルの語彙知識を変化させていたことが分かった。気になったのは、もっと伸びるであろうと期待されていた AWL の数値がそれほど上がっていなかったことである。考察のところでも述べたが、これは AWL の単語や語句の問題なのか、現存のカリキュラムの問題なのか今後も検証していかなくてはならない。基本的な 2000 語レベルにおいては高得点を得ており、基本的な語彙の知識を身に着けていることが分かった。

英語母語話者の語彙知識を LSA という手法を用いて推定し、日本人大学生の結果とを比較したが、この二つのグループ間には有意な差はないことが分かった。うれしい半面、この結果は注意を持って受け止めなければいけない。LSA を用いたシミュレーションでは、考察の所で見たように実際の参加者のデータよりも低い可能性がある。人間の場合、他の問題との

兼ね合い、消去法、単語に関するこれまでの既存の知識から推測など学習方略を駆使して、一番ふさわしい回答へと辿り着ける。しかし、LSAのSimulationの場合は、ただ単に対象のコーパスのデータから単語と語句の間の関連性を表す数値のみで判断していかなければならない。また、LSAでは同じ数値が二つあり、どちらか判断がつかないような場合もあった。今回のLSA分析を通して、そもそも語彙知識とはどういうものかという問いを改めて考える機会となった。語彙知識とはLSAがそうであるように、単語と語句の関連性、意味の類似性、即ち、単語そのものの知識とするのか、消去方法や推測なども含め、他の単語の知識も含め、総合的なものとするのか、VLTの妥当性も含め、語彙知識をどう捉えるか今後も検討していかなければいけない。

その他の今後の研究の課題としては、母語話者のデータをLSAによる推定ではなく、参加者を募り、実際の母語話者からデータを取り分析することが挙げられる。そして、今回のLSAのシミュレーションの推定値と実際の英語母語話者の実測値を比較することが必要となる。今回のテストでは基本語2000語の知識は高いことが分かったが、これはあくまでも単語と語句の繋がりが認識できるという受容的な知識を指しており、産出的な知識、即ちその単語が使用できるのかについても調べていく必要がある。

## 謝辞

この研究はJSPS科研費基礎研究(c)課題番号18K00748によって助成を受けている。

## 参考文献

- Biemiller, A., Rosenstein, M., Sparks, R., Landauer, T. K., & Foltz, P. W. (2014). Models of vocabulary acquisition: Direct tests and text-derived simulations of vocabulary growth, *Scientific Studies of Reading*, 18(2), 130-154.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34, 213-238.
- Dennis, S. (2007). How to use the LSA website. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.). *Handbook of latent semantic analysis* (pp. 57-70). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kucera, B. & Francis, W. N. (1967). *A Computational Analysis of Present-Day American English*. Providence, Rhode Island: Brown University Press.

- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.
- Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. Boston, MA: Heinle & Heinle.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88
- Thorndike, E. & Lorge, I. (1944). *The Teacher's Wordbook of 30,000 Words*. NY: Columbia University Teacher's College.
- West, M. (1953). *A General Service List of English Words*. London: Longmans, Green and Co.
- Yoshii, M., & Lavin, R. (2017). A new extensive reading and listening class: How effective is it? 『九州英語教育学会紀要』, 45号, 29-38.
- 猪原敬介・楠見孝 (2012) 読書習慣が語彙知識に及ぼす影響—潜在意味解析による 検討— *Cognitive Studies*, 19(1), 100-121
- 吉井誠. (2019). 「潜在意味解析を用いた語彙習得研究の展望について」. 『熊本県立大学文学部紀要』, 78号, 45-57.
- 吉井誠. (2017). 「多量のインプットが英語学習にもたらす効果：多読多聴クラスにおける学習者の変化について」. 『熊本県立大学文学部紀要』, 76号, 59-71.