

## 語彙レベルディクテーションテストの日本人学習者への適用

今井由美子

同志社女子大学

三根 浩

同志社女子大学

### ABSTRACT

Fountain and Nation (2000) developed a vocabulary-based graded dictation test (Vocabulary Levels Dictation Test; VLDT) "which focused strongly on vocabulary knowledge and which could be used with learners of widely differing proficiencies." They provided four equivalent forms and showed high reliability and validity. However, there are few reports on the application of VLDT because it is too difficult for EFL learners to apply VLDT such as the Lexical Frequency Profile or the Vocabulary-size test of controlled productive ability. The purpose of this study was to verify the applicability and reliability of the VLDT test for Japanese EFL learners, given to 337 female college students who answered VLDT-Test forms A and B. The results indicated statistically significant correlations and alpha coefficients of 0.86 for both tests, and showed the reliability of the VLDT. On the other hand, the average scores of the VLDT were quite low, which suggests the possibility that the VLDT measured memory spans rather than vocabulary knowledge.

**Key words:** vocabulary, dictation, reliability

### I. はじめに

語学学習の指導において、指導者が学習者の語彙力を把握し、そのレベルに合わせた指導法を取り入れていくことは重要である。語彙学習や語彙サイズ測定のために活用されている学習用語彙リストには A teacher's word book of 30,000 words (Thorndike and Lorge, 1944)、General Service List (West, 1953)、北海道大学英語基本語彙表(園田, 1996)、Academic Word List (Coxhead, 2000)、SLV12000 (アルク英語企画開発部, 2000)、JACET8000 (大学英語教育学会基本語改訂委員会, 2003)などがある。語彙力を測定するテストは、個々の語彙リストからグレード別の単語サンプルを探り、その中から知っている単語の割合を推定することで全体的な語彙サイズを算出している。

語彙テストには、単語の文字形式、すなわち単語の綴りを提示し意味を選択させたり知識をチェックさせたりする再認テスト形式の「受容的語彙テスト」と、単語や文章を書かせる

形式の「発表的語彙テスト」がある。代表的な語彙テストには、受容的語彙力を測定するものとして望月(1998)による語彙サイズテスト、Schmitt (2000)や Nation (2001)による語彙レベルテスト、Meara (1992)のチェックリスト・テストなどが、また発表的語彙力を測定するテストとして Laufer and Nation (1995, 1999)の語彙頻度プロフィール法や統制的発表語彙サイズテストなどがある。一般的に、発表的語彙力のテストは日本人 EFL 学習者にはむずかしそうだと考えられているため、発表的側面を有する語彙テストが使用された研究例は少ない。望月(2003)は、発表語彙サイズを測定するテストの日本における適用例を紹介しているが、その中で、「統制的発表語彙テストは、受験者がある程度の読解力、文法力を持っていることを前提」にすることと、「語彙頻度プロフィールの問題点としては、作文の長さが 400 語以上」であり「比較的長めの作文でないと処理できない」ため初級学習者に向いていないことを指摘している(pp. 191-192)。

Fountain and Nation (2000)は、ヒアリングの要素を含めた語彙知識測定のために、伝統的なディクテーション形式による語彙レベルテスト(Vocabulary Levels Dictation Test; 以下 VLDT)を開発した。ディクテーションテストは特定の言語項目に個別に焦点を当てるのでなく、学習者の言語能力を総合的に測るものであり、ヒアリングという「受容的」スキルと、聞き取った単語や文章を書き取るという「発表的」スキルにより構成されている。このテストの解答者は、一連の音の流れから単語を一つずつ文節化し、順次に記憶し、解答用紙の上にセンテンスを再構成することが求められる。

VLDT は A, B, C, D の 4 つの並行テストが作成されており、各テストは 5 つのパラグラフからなる文章で構成されている。導入パラグラフは基本語彙 1~500 語、第 1 パラグラフは基本語彙 501~1,000 語、第 2 パラグラフは 1,001~2,000 語、第 3 パラグラフは 2,001~3,000 語、第 4 パラグラフは 4,000~6,000 語にグレード分けされた語彙をターゲット語としている。Fountain and Nation (2000)は、これらの 4 並行テストの相互相関は 0.95 以上、語彙レベルテストとの相関は 0.78、Grammar test との相関は 0.70 以上であり、高い信頼性と妥当性を有することを報告している。

## II. 先行研究

ディクテーション課題を行う際、ヒアリングとスペリングの間で作業記憶(Working Memory)が重要な働きをしている。Baddeley (1992)によると作業記憶は、音韻ループ、視空間記録メモ、中央実行系の 3 要素から構成されているが、これらのうち音韻ループが発話に基づく情報の貯蔵とリハーサルを行い、第 1 言語と第 2 言語の両方の獲得に不可欠であるとされている(p. 556)。音韻ループは、音響的情情報を 1~2 秒間保持し、調音制御プロセスに連結する(p. 558)。ディクテーションテストにおける誤答が、ヒアリング、音韻ループ、スペリングのどの段階で生じたのかを検証するためには語聴分析が必要である。

日本人 EFL 学習者のディクテーションにおける誤聴分析を試みた研究には、筧・末延・野間・神崎・山根(1979)、藤永(2002)、東矢(2005)などがある。筧他(1979)は大学生 100 名を被験者とし日記の録音テープ(1 分)を 30 分かけて書き取らせた。その結果、機能語の所で空白が多いこと(p. 6)と「未知語に出くわした場合、context を手がかりにして、意味的に類似し

た語を既知の語の中から選ぼうとする傾向(p. 9)」を指摘した。藤永(2002)は大学生 27 名に対し 122 語と 134 語の会話文をディクテーションさせ間違いの種類と割合を報告した。全ての種類の間違いに共通する原因として、「1) ある音素が聞けない、2) 弱形の発音が聞けない、3) 音変化が聞けない(p. 20)」ことを挙げている。東矢(2005)は大学生 51 名を被験者とし、7 つの英文を各文 4 回ずつ聞かせディクテーションテスト、語彙テスト、リスニングテストを行った。その結果、「知っている語は、知らない語よりも正しく認識できる(p. 498)」ことや初頭効果および親近性効果(p. 499)が見られることなどを報告した。

古莊・今井・三根(2009)は、語彙レベルテストおよび VLDT で測定される語彙力と、CELT や TOEIC などの英語標準テストで測定される語学力、および語彙調べ課題の効果について検証し、「語彙力と語学力は関連していること」、そして「VLDT は語彙力の指標として信頼性があること」を報告した。しかしながら、Fountain and Nation (2000)によって開発された VLDT は、発表的側面を測定する語彙頻度プロフィール法や統制的発表語彙サイズテストなどと同様に EFL 学習者には難しいため、日本において適用されたという研究報告が見受けられない。

### III. 目的

本研究の目的は、日本語を第 1 言語とする日本人 EFL 学習者の発表的語彙知識を測るために、Fountain and Nation (2000) が作成した VLDT の適用可能性とテストの信頼性を検証することである。第 1 実験では、VLDT の Test-A と Test-B を原版のまま実施し両テストの相関係数と平均得点の差異を検討する。第 2 実験では、VLDT の困難性の解決を試み信頼性を検証するとともに、適用困難性の原因を単語レベルで分析し、発表的語彙テストを日本人 EFL 学習者に適用することによる教育指導への貢献を明らかにしたい。

### IV. 実験 I

#### 1. 被験者

被験者は、英語専攻 2 年生で週 1 回 90 分のリスニング科目を受講した女子大学生 198 名である。2007 年の春学期初めに VLDT の Test-A、学期末の 7 月に VLDT の Test-B を実施した。

#### 2. 手続き

Test-A および Test-B の課題音声は、日本人に英語を教えた経験が豊富なアメリカ人女性に依頼し、テスト・スクリプトに示されているボーズ・マーク( / )の位置でフレーズ読みした音声を録音した。テスト用テープには、ボーズごとのフレーズの長さに合わせ、書き取り時間を 9 秒から 17 秒の範囲で挿入した。各録音テープは、Fountain and Nation (2000, p. 30) の指示通り約 12 分の長さにし、1 回だけ提示した。解答用紙は自作し、日本語による簡単な指示文の後に、それぞれのフレーズごとに聞き取った音声を書き取るための解答欄を用意した。全ての書き取りが終了した後、見直しのための時間を 3 分間与えた。採点方法は、Fountain and Nation (2000, pp. 32-33) に準じた。採点の対象となるターゲット語は各パラグラフに 20 語ずつ含まれ、第 1 パラグラフから第 4 パラグラフまで各 20 点ずつ、合計 80 点満点で採点され、導入パラグラフは採点の対象にはならない。採点基準は、スペリングが正確に綴られている

場合だけを正解とした。ただし、名詞の複数形の-s と-es、規則動詞過去形の-d と-ed は、語幹の綴りが正しければ許容される。

### 3. 結果と考察

Test-A と Test-B の相関係数は 0.79 であった。Table 1 は 2007 年の Test-A と Test-B の平均得点と *t* 検定の結果を示した。Test-A の平均得点 20.88 は Test-B の平均点 29.19 に比べて極端に低かった。とくに、第 2 パラグラフにおける差は大きく、Test-A の課題が難しすぎることを示した。この原因としては、課題を 1 回しか提示しないことと、Test-A の課題文ではフレーズの切れ目の数が Test-B に比べて少ないため一つのフレーズに含まれる単語数の多いことが挙げられる。

Table 1. Test-A(原版)と Test-B の平均得点

Paragraph	Test A	Test B	<i>t</i>	<i>p</i>
1	7.46	11.90	-21.91	0.01
2	4.72	9.14	-25.76	0.01
3	4.86	4.60	1.71	0.09
4	3.84	3.56	1.78	0.08
total	20.88	29.19	-21.72	0.01
note, <i>df</i> = 197				

Fountain and Nation (2000)のテスト・デザインによると第 1 パラグラフの平均語数は 5.0 語、第 2 パラグラフが 5.5 語、第 3 パラグラフが 6.0 語、第 4 パラグラフが 7.0 語と指定されている(p. 32)。しかしながら、Test-A の第 1 パラグラフは総語数 36 語が 6 フレーズに分けられているため平均語数は 6.0 語、第 2 パラグラフは総語数 39 語が 7 フレーズに分けられているため平均語数は 5.7 語である。竹蓋(1984)が指摘したように「6 語目以降」急激に誤りが増える(p. 235)ことが Test-A の第 2 パラグラフの得点の低さに影響していると思われる。そこで、実験 II の調査では、ディクテーションテストにおける誤答が、ヒアリング、音韻ループ、スペリングのどの段階で生じたのかを検証するために、Test-A のテキスト中、6 語以上の語数からなるフレーズを発音の切れ目で分割することにより修正を加え、単語レベルでの誤聴分析を試みることにした。

## V. 実験 II

### 1. 被験者

実験 I の結果と比較するために追加した被験者は、英語専攻 2 年生で週 1 回 90 分のリスニング科目を受講した女子大学生 139 名である。2008 年の春学期初めに VLDT の Test-A、学期末の 7 月に VLDT Test-B を実施した。

### 2. 手続き

VLDT のテキストは実験 I と同じである。Test-A だけボーズ箇所を追加した。追加した箇

所は、第 1 パラグラフで 3 か所(それぞれ 8 語文、8 語文、8 語文)、第 2 パラグラフで 3 か所(6 語文、7 語文、8 語文)、第 3 パラグラフで 4 か所(8 語文、7 語文、6 語文、6 語文)、第 4 パラグラフで 1 か所(9 語文)である。Test-B は実験 I から変更無しで実施した。

### 3. 結果

2008 年の Test-A と Test-B の相関係数は 0.82、2007 年と 2008 年を合わせたデータでの信頼性係数  $\alpha$  は Test-A が 0.8594、Test-B が 0.8597 であり、2 つのテストとも十分に高い信頼性を示した。

Table 2. Test-A(修正版)と Test-B の平均得点

Paragraph	Test A	Test B	t	p
1	9.81	12.12	-10.84	0.01
2	6.24	9.47	-13.46	0.01
3	7.09	4.40	14.37	0.01
4	4.96	4.34	3.05	0.01
total	28.11	30.41	-4.83	0.01

note,  $df = 138$

Table 2 は、2008 年の Test-A と Test-B の平均得点を示した。Test-A と Test-B の総得点の差異は有意差があるものの 2.3 点差まで小さくなつた。しかし Test-A では、語彙レベルの低い第 2 パラグラフより第 3 パラグラフで得点が高くなつたことは不合理である。第 3 パラグラフでは 4 か所でボーズを追加したが、内 2 か所は 6 語文へ追加されたボーズであった。その結果、正答率が第 2 パラグラフよりも高くなつてしまつた。

Table 3-1~4 は Test A の各単語における正答率を原版(2007 年)と修正版(2008 年)の間で  $t$  検定した結果である。 $t$  値がマイナスの単語は修正版において正答率が上がつたものである。原版のテスト結果と比較するために、フレーズを分割した箇所を Figure 1~4 に示した。Figure 1 に示されたように、原版では“Some obtain jobs on farms or in industry.”において、最後の industry で 58% の正答率があるものの、jobs, farms ではわずか 4% である。しかし修正版では“Some obtain jobs on farms / or in industry.”と区切ることで、jobs が 12%、ボーズ直前の farms が 34%、そして短く区切られた後の部分の industry は 74% とそれぞれ正答率が向上した。また“A few with skills / in art or music.”においては、ボーズの前後に位置する skills と art で 30% 近く正答率が上がつてゐる。

Table 3-1. Test-Aの第1 パラグラフにおける原版と修正版の正答率比較

	原版	修正版	t	p
OBTAIN	1%	1%	0.29	0.78
JOBS	4%	12%	-2.59	0.01
FARMS	4%	34%	-8.27	0.01
INDUSTRY	58%	74%	-3.16	0.01
OTHERS	67%	57%	1.93	0.05
ACCEPT	17%	14%	0.65	0.52
POSITIONS	21%	24%	-0.59	0.56
GOVERNMENT	42%	59%	-3.01	0.01
SERVICE	55%	65%	-1.91	0.06
MANY	87%	91%	-1.09	0.28
SEEK	6%	26%	-5.39	0.01
POSTS	0%	1%	-0.89	0.37
BUSINESS	61%	69%	-1.54	0.13
TRADE	31%	57%	-4.88	0.01
FEW	91%	92%	-0.55	0.58
SKILLS	43%	77%	-6.78	0.01
ART	27%	55%	-5.59	0.01
MUSIC	94%	99%	-2.68	0.01
APPLY	29%	48%	-3.79	0.01
FIELDS	10%	29%	-4.77	0.01

**Paragraph 1****Some obtain jobs on farms or in industry.** /

Original	1%	4%	4%	58%	/
Modified	1%	12%	34% /	74%	/

**Many seek posts in business or a trade.** /

Original	87%	6%	0%	61%	31%	/
Modified	91%	26%	1% /	69%	57%	/

**A few with skills in art or music.** /

Original	91%	43%	27%	94%	/
Modified	92%	77% /	55%	99%	/

Figure 1. 第1 パラグラフにおける正答率の変化

第1 パラグラフ(Table 3-1)において、正答率が 1% と 0% であった語は、原版および修正版で obtain, posts の2語であった。obtain の誤答例を見るとほとんどの被験者は“of”を書き取っており、“of ten”と書いている例が多く見られた。直後に“on farms”と続くにも関わらず “some of”と聞き取ってしまうことにより意味的にも文法的にも意味処理が上手くできていないことが伺える。音声的な視点では、ob-tain の後半は / ei / という二重母音だが、これを「エー」で代用してしまう日本人 EFL 学習者における二重母音の捉え方の特徴を見ることができる。posts の誤答例においても / ou / という二重母音にあたる綴りが“o”であることに気付かず、o-u と発音そのままのスペルを書いている誤りや、pose と綴る誤りが過半数を占めた。この

ように聞き取りの難しい単語ではフレーズ分けの効果が見られない。

第2パラグラフ(Table 3-2)で正答率が上がった単語は、level, education, average, records, achievementであった。フレーズ分けの結果、ボーズ直前に位置することになったeducationとrecordsでは記憶再生が容易になったと思われる。また、level, average, achievementはボーズを入れたことで4語文となり、音韻ループの中で処理できるようになったのであろう。frequently, affect, consider, applicationsでは正答率に有意差は無かった。聞き取りのできない単語ではボーズの有無は正答率に影響していない。

Table 3-2. Test-Aの第2パラグラフにおける原版と修正版の正答率比較

	原版	修正版	t	p
LEVEL	9%	28%	-4.78	0.01
EDUCATION	31%	70%	-7.92	0.01
FREQUENTLY	14%	17%	-0.92	0.36
AFFECT	3%	5%	-0.92	0.36
RANGE	40%	39%	0.26	0.80
POSSIBLE	43%	30%	2.50	0.01
OPENINGS	23%	19%	0.81	0.42
FIRMS	6%	2%	1.72	0.09
INSTANCE	60%	60%	0.00	1.00
SELECT	30%	46%	-3.05	0.01
EXCELLENT	29%	32%	-0.49	0.62
CANDIDATES	11%	7%	1.23	0.22
TRAINING	67%	59%	1.54	0.12
FUTURE	37%	41%	-0.81	0.42
EXECUTIVES	1%	4%	-1.56	0.12
CONSIDER	20%	14%	1.51	0.13
APPLICATIONS	2%	3%	-0.56	0.58
AVERAGE	29%	61%	-6.27	0.01
RECORDS	2%	35%	-9.39	0.01
ACHIEVEMENT	16%	50%	-7.36	0.01

### Paragraph 2

Their level of education frequently affects. /

**Original** 9% 31% 14% 3% /

**Modified** 28% 70% / 17% 5% /

They will not consider applications from people /

**Original** 20% 2% /

**Modified** 14% / 3% /

with only average records of achievement at school. /

**Original** 29% 2% 16% /

**Modified** 61% 35% / 50% /

Figure 2. 第2パラグラフにおける正答率の変化

Table 3-3. Test-Aの第3 パラグラフにおける原版と修正版の正答率比較

	原版	修正版	t	p
FACTORS	33%	41%	-1.54	0.12
INFLUENCE	29%	55%	-5.03	0.01
CHOICE	38%	88%	-10.64	0.01
CAREER	2%	6%	-2.08	0.04
INFORMATION	81%	84%	-0.65	0.52
AVAILABLE	23%	23%	-0.01	0.99
UNCERTAIN	12%	14%	-0.76	0.45
PROBABLE	33%	34%	-0.26	0.80
FINANCE	57%	64%	-1.42	0.16
CONDITIONS	22%	61%	-7.84	0.01
PROSPECTS	1%	15%	-5.07	0.01
IMPROVEMENT	30%	54%	-4.55	0.01
SIGNIFICANT	28%	39%	-2.17	0.03
CONSIDERATIONS	9%	5%	1.56	0.12
APPARENT	0%	2%	-2.08	0.04
ORGANIZATIONS	6%	8%	-0.64	0.52
RETAIN	6%	8%	-0.45	0.65
EMPLOYEES	12%	16%	-0.98	0.33
SATISFACTION	46%	46%	0.15	0.88
RESPECTS	12%	48%	-8.14	0.01

## Paragraph 3

What factors influence the choice of a career? /

Original	33%	29%	38%	2%
Modified	41%	55%	/	88%

The information available on this is uncertain /

Original	81%	23%	12%	/
Modified	84%	23%	/	14%

working conditions and prospects of improvement /

Original	22%	1%	30%	/
Modified	61%	/	15%	54%

give them satisfaction in these respects. /

Original	46%	12%	/
Modified	46%	/	48%

Figure 3. 第3 パラグラフにおける正答率

第3 パラグラフ(Table 3-3)において、正答率が上がった単語の中では、influence, choice, conditions, respectsにおいてフレーズが分割されたことによる効果が出ている。とくに choiceは、先頭に位置することで初頭効果が生じ音韻ループで処理しやすくなつたと推測できる。正答率が比較的高い単語は information と finance であるが、これらはカタカナ語として定着しているうえに、カタカナ発音も本来の英語の発音と大して変わらないためと思われる。こ

のパラグラフで正答率が特に低かった単語は、career と apparent の 2 語であった。career は、フレーズ分けの効果を若干受けているものの修正版でも正答率は低かった。カタカナ語では career は「キャリア」と読まれてしまうことが多いが、正しい発音はかなり異なるため、誤答例のほとんどは韓国の“Korea”と書かれ、次に多かったのが発音そのままにスペルした“Koria”であった。apparent は 3 語からなるフレーズの最後の単語だが、誤答例としては the や a のついた parent の、“the parent”と “a parent”というパターンが多く見られた。“parent”は、直前に a がついた場合、形容詞の apparent とよく似た発音となる。音声的には処理できているが、“It seems apparent that...”と続く文の一部であり、意味的処理は上手くいっていない。

Table 3-4. Test-A の第 4 パラグラフにおける原版と修正版の正答率比較

	原版	修正版	t	p
THOROUGH	0%	1%	-0.25	0.81
INVESTIGATION	31%	32%	-0.33	0.74
MOTIVES	7%	6%	0.50	0.62
OPERATE	16%	17%	-0.41	0.68
SELECTION	13%	14%	-0.30	0.76
EMPLOYMENT	9%	12%	-1.10	0.27
PROFITABLE	0%	1%	-0.25	0.81
TOPIC	29%	16%	2.93	0.01
RESEARCH	40%	50%	-1.89	0.06
EMPLOYERS	9%	23%	-3.71	0.01
APPRECIATE	14%	25%	-2.68	0.01
ASSISTANCE	2%	24%	-6.65	0.01
FINDINGS	33%	28%	0.88	0.38
ENLIST	0%	0%		
MAINTAIN	36%	27%	1.82	0.07
STABLE	6%	7%	-0.39	0.69
STAFF	55%	59%	-0.87	0.39
INDUCED	2%	8%	-2.52	0.01
INVEST	10%	46%	-8.52	0.01
PROJECT	76%	92%	-3.97	0.01

## Paragraph 4

Employers who would appreciate the assistance of the findings. /

Original	9%	14%	2%	33%	/
Modified	23%	25%	/	24%	28%

Figure 4. 第 4 パラグラフにおける正答率の変化

第 4 パラグラフ(Table 3-4)において、正答率 0% の単語は enlist、1% の単語は thorough と profitable であった。兵役のない日本で enlist が書けないのは仕方がないといえよう。thorough の解答はほとんどが空白であり、何か書いている場合でも th だけで終わっていた。音の聞き取りも綴りも難しい上に、被験者において、意味の知識が充分にあるとは期待できない単語

だと推測される。*profitable* は 7 語文の中央である 4 語目に位置していた単語であり、空白のままの被験者がほとんどであった。直後に“topic for research”と記憶に残りやすい単語が複数続くことで音韻ループから脱落していた可能性が高いと考えられる。一方、修正版で分割した 9 語文(Figure 4)では、文末の *findings* を除く *employers*, *appreciate*, *assistance* で有意な正答率の上昇が見られる。*findings* は未知語と考えられないため *the* や *or* を含む文章においても 5 語目の単語の処理が難しいことと、*assistance* のスペリングを考えている間に音韻ループから脱落した可能性とが考えられる。

第 4 パラグラフは 4,000~6,000 語レベルの語彙力を測定するため、全体の正答率は他のパラグラフと比較してさらに低いものになった。しかしながら、被験者が 50% 前後まで正答している単語は *research*, *staff*, *project* などいくつか見受けられた。これらの単語は Longman の辞書では S1 または S2, W1 または W2 といった 1,000 語~2,000 語にランクされている語であり、「知っている語は、知らない語より正しく認識できる」ことを裏付けた(東矢, 2005, p. 498)。このような単語が 4,000~6,000 語レベルのターゲット語とされたのは、Thorndike and Lorge (1944) の語彙リストが用いられたためである。当時は 5,000 語レベルにリストされていた単語も、時代の推移とともに日常的ななじみのある単語となってきていたため、被験者にとっては容易に正解に結びつけられたのであろう。

Test-B は原版と修正版では同じものを使用したため、正答率に大きな差異は見られなかつた。第 2 パラグラフで正答率が 2% と 3% であった *scarce* の誤答例を見た場合、被験者の多くは *scares*, *scared*, *scare* という発音のよく似た既知語を書いていた。語末の /s/ を -ce ではなく -se と書いている誤答例も数例見受けられた。文末に位置している単語であるが、被験者には未知語である可能性が高いため正答に結びつけられなかったと思われる。第 3 パラグラフでは *failures* の正答率が 0% であった。誤答例としては *failing* や *fail* がわずかに見られただけでほとんどは空白のままであった。しかし、被験者の多くはこの単語の直前の *previous* およびチャンク末の *starving* を正解に至らないまでも聞き取ることができていた。このことから、7 単語からなる “in the previous failures to prevent starving” の中央に位置する “failures” が音韻ループから脱落したことが原因と考えられる。第 4 パラグラフでは、*regions*, *rebelling*, *devised* の 3 語が 0%~1% の正答率であった。*regions* の誤答例は、少数であるが *reasons* が最も多く、-gions の有声破擦音部分と -sons の有声摩擦音を聞き間違えていた。*rebelling*においては、語頭の re- と語末の -ing の組み合わせた解答例が多く見受けられたが、被験者にとって未知語であった可能性が高く、その組み合わせのほとんどは正しく綴られていなかった。*devised* の正答率は 0% であったが、最も多い誤答例は正しいスペリングではない “devide(d)” であり、次いで意味が異なる “device” であった。この単語も被験者にとって未知語であったと思われる。

## VI. 考察

Test-A と Test-B の相関係数は 0.79~0.82、また信頼性の指標とされている  $\alpha$  係数は、Test-A が 0.8594、Test-B が 0.8597 であり十分に高い信頼性を示した。修正版のデータでは、Test-A で 6 語以上の語数からなるフレーズを分割した結果、Test-A と Test-B の総得点の差異は小さくなつた。原版のテスト形式では、長いフレーズを聞き取る際に前半部分の単語が音韻ル

の容量を超えることでセンテンスの再構築ができなかった。長い文を聞き取る際に、中央に位置する単語が音韻ループに入りきらないことで、その部分を再構築することが難しくなる「中だるみ効果」は、観他(1979, p. 4)により指摘されている。修正版では、長いフレーズを分割した結果、ボーズの前後の単語における正答率が向上した。分割されたことでターゲット語を音韻ループに残すことが可能になったと推測できる。

VLDT は  $\alpha$  係数の値が示したように十分に高い信頼性を有するが、被験者の平均正答率が非常に低いことは、語彙力よりも音韻ループの容量に制約を受けていた可能性を示唆しており、とくに Test-A の妥当性には疑問が残る。VLDT は、フレーズに含まれる単語数が多いことにより日本人 EFL 学習者にとって難しい課題になった可能性があり、やさしい単語でも記憶スパンから外れると書き取ることができなかつた。また、7 語文以上のフレーズ中央の単語は脱落しやすかつた。すなわち、VLDT では語彙力よりも記憶スパンを測定している比重が大きいと考えられる。

正答率が 0%~1% であった単語では、音声的には処理ができているものの、意味的・文法的な処理ができないことによる誤答例が見受けられた。カタカナ読みが英語の発音に近い単語の正答率は高いが、英語本来の発音と異なる単語の場合にはローマ字表記に近い綴りでの解答が顕著であり、二重母音とスペリングの結びつきが薄い。また子音においても（例えば /s/ において -se, -ce, -s など複数の表記パターンがある場合）音声的に聞き取れているのに正しいスペリングにつながらない誤答が見受けられる。

これらの結果から、今後の指導法として、1) 記憶スパンを長くするスキル(語彙化や句構造の把握など)の指導、2) カタカナ語と英語の発音の違いについての指導、3) 二重母音や子音部分のスペリングにおける正確性を高めるための指導に重点をおく必要性があることを提案できると考える。

注：この論文は、第 48 回（2009 年度）JACET 全国大会において口頭発表したものに加筆修正を加えたものである。また今回の実験にご協力いただいた同志社女子大学の枝澤康代先生、古庄智子先生、平岩葉子先生に感謝する。

## REFERENCES

- アルク英語企画開発部 (2000)『英検、TOEIC スコア・アップ 英会話上達のための最強のボキャブラリー』東京：アルク  
大学英語教育学会基本語改訂委員会（編）(2003)『大学英語教育学会基本語リスト (JACET8000)』大学英語教育学会  
藤永真理子 (2002) なぜ英語が聞き取れないか：学生のディクテーションの分析から『経済理論(和歌山大学経済学会)』306, 1-22  
古庄智子・今井由美子・三根浩 (2009) 大学生における語彙力と英語標準テストの関連性(3) —ディクテーションテストと語彙力—『第 49 回外国語教育メディア学会全国研究大会』

146-147

- 寛寿雄・末延岑生・野間司郎・神崎和男・山根繁 (1979) 誤聴分析『大学英語教育学会紀要』10, 1-9
- 望月正道 (1998) 日本人学習者のための英語語彙サイズテスト『語学教育研究所紀要』12, 27-53
- 望月正道 (2003) 語彙はどうテストするのか 望月正道・相澤一美・投野由紀夫 (編)『英語語彙の指導マニュアル』東京: 大修館書店 181-209
- 園田勝英 (1996) 大学生用英語語彙表のための基礎的研究『言語文化部研究報告叢書』7 北海道大学言語文化部
- 竹蓋幸生(1984)『ヒアリングの行動科学』東京: 研究社出版
- 東矢光代 (2005) 英文リスニングにおける日本人学習者の語句認識『日本大学大学院総合社会情報研究科紀要』6, 495-506
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Fountain, R. L., & Nation, I. S. P. (2000). A vocabulary-based graded dictation test. *RELC Journal: Guidelines*, 1, 76-80.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16, 33-51.
- Meara, P. (1992). *EFL vocabulary tests*. Centre for Applied Language Studies, University College Swansea.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, UK: Cambridge University Press.
- Thorndike, E. L., & Lorge, I. (1944). *A teacher's word book of 30,000 words*. Teachers College, Columbia University.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co..