

【原著】

日本人英語学習者における文脈からの意味推測のために 最適な未知語の割合と語彙サイズ

八 島 等

Exploring the Optimal Unknown Word Density and Vocabulary Size for
Japanese EFL Learners to Guess Meanings from Context

Hitoshi Yashima

日本人英語学習者（高校生）にとって文脈からの意味推測のために最適な未知語の割合と語彙サイズを求めて研究を行った。本研究では、「最適な」を「80%以上の推測成功率を収める」と定義した。高校1年生用検定教科書からの4つの未知語の割合（50語、60語、80語、100語にそれぞれ1語）のパッセージを用意し、高校1～3年生205名を対象として調査した。その結果、100語に1語の場合、語彙サイズが3,000 word families程度あれば、平均で60%以上の推測成功率が得られることがわかった。また、80語に1語の場合でも、語彙サイズが4,000 word families程度あれば、推測成功率は平均で80%以上であることがわかった。さらに、4,500 word families程度の語彙サイズがあれば、いずれの未知語の割合でも、平均で80%以上の推測成功率を収めることができることがわかった。その上、語彙サイズと推測成功率の間には.682以上のかなり高い相関が見られることが確認された。

1. 研究の背景と目的

文脈からの未知語の意味推測の重要性については、多くの研究者が指摘している（Liu & Nation, 1985; Nation, 1990, 2001; Nation & Coady, 1988; van Parreren & Schouten-van Parreren, 1981）。例えば、Liu & Nation（1985）は、文脈からの未知語の意味推測は書かれたテキストの中での低頻度語に対処するための最重要方略であると述べている。しかし、その一方で、文脈からの未知語の意味推測の難しさを指摘する研究も多数ある（Dubin & Olshtain, 1993; Haynes, 1993; Laufer, 1997）。例えば、Laufer（1997）は、第二言語または外国語としての英語（以下、L2）における意味推測が大部分の未知語で可能であり、推測の成功は主に学習者の推測方略にかかっているということを当然のこととは受け入れがたいと述べている。語彙カバー率（lexical coverage）や語彙サイズ、読解（reading comprehension）の相互関係についての研究成果は一貫したものではない。

まず、語彙カバー率と読解との関係を初めて調査したLaufer（1989）は、適切な読解には95%の語彙カバー率が必要であると提唱している。それに対して、Hu & Nation（2000）は、95%の語彙カバー率では、適切な読解ができた調査参加者の数は90%の場合より増えたが依然として少数で、98%の語彙カバー率が必要であると提唱している。Schmitt, Jiang, & Grabe（2011）は、60%の読解のためには、98%の語彙カバー率が必要であるとHu & Nation（2000）の提唱を支持している。さらに、Carver（1994）は、小学生及び大学院生の英語母語話者を対象とし

た研究により、題材が比較的易しい場合、未知語はほぼ0%であり、題材が比較的難しい場合、2%以上が未知語であり、題材の難易度が読み手の能力にほぼ一致している場合には、約1%が未知語であると指摘している。このことを語彙カバー率に当てはめてみると、母語の場合でも、99%が適切な語彙カバー率であると考えられる。

次に、語彙サイズと読解の関係について、Laufer (1992) は、56%の読解問題の正答率を収めるためには、3,000 word familiesが必要であると提唱している。

更に、語彙カバー率と語彙サイズとの関係では、Hirsh & Nation (1992) は若い英語母語話者用の簡易化されていない短編小説を楽しく読むために必要な語彙サイズを調査した。97-98%の語彙カバー率に達するには約5,000 word familiesが必要であるということを見出した。Nation (2006) は小説や映画などを調査し、98%の語彙カバー率を理想とするならば、書かれたテキストに対処するためには8,000-9,000 word familiesが必要であると提唱している。

Laufer & Ravenhorst-Kalovski (2010) は「適切(十分)な」(adequate) 読解とは何かという明確な定義はないとした上で、上記の通り、外見上は研究者の提唱は違いがあるように思われるが、研究成果は驚くほどよく収束するように思われると主張している。Laufer (1989) の95%の語彙カバー率は、55%の読解問題の正答率を「適切な」読解と設定し、その最低限の読解に足りるだけのものであり、70%の正答率のためには5,000 word familiesを知っている必要があることを見出していると主張している。さらに、Nation (2006) の、8,000-9,000 word familiesは98%のカバー率のための最高の推定値であり、各語彙レベルの最低カバー率を足すと、95%に達するためには3,000 word familiesが、98%には5,000 word familiesがそれぞれ必要なことになるかと計算している。そして、研究結果として、「適切な」読解の定義に応じて2つの閾値(threshold)を提唱している。すなわち、98%の語彙カバー率をもたらす8,000語を「理想的な(最適な)」(optimal) 閾値、95%の語彙カバー率をもたらす4,000-5,000語を「最低限の」(minimal) 閾値としている。

以上のように、語彙カバー率の読解に対する影響に関する議論は多数あるが、実際にその影響を測定した研究はほとんどない(Webb, 2010)。語彙サイズの意味推測に対する影響を測定した研究についても同様である。そこで、日本人英語学習者(高校生)にとっての文脈からの未知語の意味推測のための「適切な」割合(語彙カバー率)について調査してきた。羽鳥(1979)は、20語に1語ぐらいの未知語は十分意味推測ができるはずであると主張している。しかし、八島(1996)では、未知語の割合が20語に1語(95%の語彙カバー率)の場合、意味推測の成功率(以下、推測成功率)の平均は約43%にすぎなかった。未知語の割合が高すぎなければ、第二言語学習者は文脈から60%~80%の未知語の意味推測ができると期待してもよいであろう(Nation & Coady, 1988)という主張から、60%以上の推測成功率が「適切な」ものと考えると、20語に1語は「適切な」未知語の割合とは言えない。また、八島(1999)では、未知語の割合が25語に1語(96%の語彙カバー率)の場合、推測成功率の平均は約47%以下にすぎず、30語に1語(96.67%の語彙カバー率)の場合でも60%以上は1例のみであり、やはり「適切な」未知語の割合とは言えなかった。そこで、未知語の割合が35語に1語(97.14%の語彙カバー率)と40語に1語(97.5%の語彙カバー率)の場合の検証を行った(Yashima, 2001)。その結果、40語に1語の割合では、推測成功率の平均が60%以上であることを見出した。

この結果から、40語に1語の未知語の割合が「適切な」ものであると仮定し、次に、一般化することができるかどうかを検証するために、5校の公立高校の調査参加者(以下、参加者)を対象に、未知語の割合が40語に1語での推測成功率を再確認する調査を行った(Yashima, 2002)。結果として、学校間で推測成功率に有意差が見られ、50.06~67.11%までの差があっ

た。語彙サイズが3,000語以上あればどの学校でも推測成功率の平均が60%以上であることがわかった。このことから、40語に1語の未知語の割合が「適切な」ものであるとは言い切れないことが判明した。さらに、Yashima (2002) までは、題材が中学校3年生用の教科書からのパッケージであったので、高校1年生用の教科書からのパッケージの場合での推測成功率を検証した (Yashima, 2003)。その結果、語彙サイズが3,500語以上あれば、1例を除き、推測成功率の平均は60%以上であることを見出した。このことから、高校1年生用のパッケージにおいて、40語に1語の未知語の割合が「適切な」ものであると言えるためには語彙サイズを考慮する必要があることが判明した。また、Yashima (2002, 2003) では、3,000または3,500語以上を一括りにして推測成功率の平均を算出しているの、実際にその語彙サイズがあれば推測成功率の平均が60%以上になる保証はなかった。

本研究の主な目的は、日本人英語学習者（高校生）が文脈から意味推測をするために「最適な」未知語の割合と語彙サイズを探求することである。前述の通り、Nation & Coady (1988) の主張から、60%以上の推測成功率が「適切な」ものとする。また、80%以上を「最適な」推測成功率と定義する。そこで、本研究では、80%以上の推測成功率を収める未知語の割合と語彙サイズを調査する。さらに、語彙と言語使用との間の強い統計的な関係を示す多くのさまざまな研究があり、語彙知識は読むこと (reading) における重要な構成要素ではあるが、それ以外のさまざまな要素がある (Hu & Nation, 2000) という主張がある。あるパッケージの中の全ての語の意味を知っているからといって、読解問題の正答率が100%になる保証はない。文法や背景知識の有無なども読解に影響を与える可能性があるからである (Hu & Nation, 2000; Webb, 2010)。そこで、語彙サイズと推測成功率との間にどの程度の相関が見られるのかということについての検証も行う。

2. 研究方法

2.1 参加者

本研究の参加者は高校1～3年生、205名である。各学年で2クラスずつ調査をした。学年ごとに、その2クラスの語彙サイズの平均に関してt検定を行った結果、いずれの学年も有意差がないことを確認した (1年生: $t(70) = 1.259$, n.s., 2年生: $t(71) = .312$, n.s., 3年生: $t(58) = 1.421$, n.s.)。

2.2 語彙サイズテスト

語彙サイズを測るテストはさまざまあるが、先行研究において幅広く用いられている Nation (1983, 1990) の改訂版 Vocabulary Levels Test (Schmitt, Schmitt, & Claphman, 2001) (VLT) を用いて、生徒の実態を踏まえて、5,000語レベルまでの語彙サイズを検証した。1語は1 word family (以下、「語レベル」以外は語) を表す。Nation のウェブサイト (<http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>) からダウンロードができ、最も頻度が高い1,000, 2,000語レベルには日本語による bilingual version があるのでそれを用いた。3,000語と5,000語レベルには bilingual version がないので、1,000, 2,000語レベルと形式を揃えるために選択肢を日本語に直した。さらに、4,000語レベルがないので、Laufer & Ravenhorst-Kalovski (2010) では、3,000語と5,000語レベルの数値の平均値を4,000語レベルの語彙サイズとしている。本研究においても、Laufer & Ravenhorst-Kalovski の手法を用いることにした。VLS には各語彙レベル30問あるので、参加者が1問正解すれば、33.33語あるとみなされる。VLS を用いて測った参加者の

語彙サイズの平均は3,797.97語、標準偏差 (SD) は589.56であった。

2.3 使用テキスト

参加者が高校生なので、語彙や文法に制限がかかっている検定教科書を題材にするのが望ましいと考えた。2種類の旧課程の高校1年生用教科書で、パッセージ間にレベル差がないように、同じ教科書会社で同じ書名のものを題材とした(平成22年度版:霜崎他 *CROWN English Series I Reading 2*からの397語のパッセージ,平成25年度版:霜崎他 *CROWN English Communication I Reading 2*からの420語のパッセージ)。各教科書から採った2つのパッセージのリーダビリティを小篠・福井の Check Read Ver.2.0 によって測定して、2つのパッセージに読みやすさの違いはほとんどなく、どちらも易しく、ほぼ高校1年生レベルであることを確認した(旧:Flesch Reading Ease(FRE):85.38, Ozasa-Fukui Year Level (YL):3.60, 新:FRE:85.23, YL:3.97)。先行研究から、40語に1語の未知語の割合が「適切な」ものであるとは言い切れないことが判明したので(Yashima, 2002, 2003), Hu & Nation (2000)などが提唱する98%からCarver (1994)が提唱する99%の語彙カバー率になるように、22年度版のパッセージを50語と100語に1語の、25年度版のパッセージを60語と80語に1語の、それぞれ未知語の割合にした。その結果、次のような4つのパッセージとなった。

- (1) Passage A: *CROWN English Series I Reading 2*からの397語のパッセージ (1:50)*
- (2) Passage B: *CROWN English Series I Reading 2*からの397語のパッセージ (1:100)*
- (3) Passage C: *CROWN English Communication I Reading 2*からの420語のパッセージ (1:60)*
- (4) Passage D: *CROWN English Communication I Reading 2*からの420語のパッセージ (1:80)*

*カッコ内の数字は未知語の割合を示す。

2種類のパッセージを用いるということは、未知語の割合以外にも様々な要因が加わることになる。しかし、限られた参加者の中で、各未知語の割合のサンプリング量を確保するために、同一参加者に、読みやすさの違いがほとんどないことを条件に異なった未知語の割合の2種類のパッセージを与えることにした。

2.4 目標語の選択

Passage Aには8語、Passage Bには4語、Passage Cには7語、Passage Dには5語をそれぞれ目標語として選択した。Liu & Nation (1985)に基づいて、目標語の品詞は名詞と動詞のみとした。参加者全員にとっての未知語を見出すことは不可能ではないにしても難しい(Hu & Nation, 2000)ので、目標語は無意味語に置き換えた。Liu & Nation (1985)に基づいて、目標語とほぼ同じ長さにして、参加者が品詞を認識できるように、名詞あるいは動詞の形になるようにして、必要があれば、接尾辞をつけたり、屈折させたりした(例えば, *adventurer*は *tampender*, *knocked*は *nuroved*)。さらに、先行研究(Hu & Nation, 2000; Nation, 1990; West, 1955)から、繰り返しの効果も考慮し、パッセージの中で1度しか出現しない語を目標語とした。その上、手掛かりがすぐにある(local clues)方が、遠くにある(global clues)よりも推測しやすい(八島, 2001)。従って、目標語の数が偶数の場合には、両者を同数にし、奇数の場合には、local clueの方を一つ多くするという統制を行った。置き換えられた目標語に下線部を引いた。参加者、とりわけ、1年生にとって、目標語以外の語で未知語となる可能性がある語の上に注として日本語で意味を添えた。パッセージの後にグローサリーとして置かなかったのは、読みの流れを止めないようにという配慮である(APPENDIX)。下線部や注以外で意味のわからない語があれば、その語に下線を引くように指示を書いていた。実際に下線を引いた

参加者は一人もいなかった。調査実施前に2人の英語母語話者（アメリカ人）に推測をしてもらい、目標語の推測が可能であることを確認し、同時に、他にもあり得る語を指摘してもらった。

2.5 研究手順及びデータ分析の方法

各学年の一方のクラスには50語と80語に1語のパスセージを与え、もう一方には100語と60語に1語のパスセージを与えた。まず、語彙サイズテストに答えさせた。その後で、2つのパスセージを読ませて、未知語の意味を文脈から推量させ、それを日本語で答えさせた。未知語の割合ごとに、推測成功率を学年別に調べた。その後、語彙サイズ別の推測成功率の平均及び推測成功率別の語彙サイズの平均を調査し、さらに、語彙サイズと推測成功率との相関も調べた。語彙サイズ別の推測成功率の平均を調査するに当たり、以前の研究の問題点を改善するために、3,000語未満から4,500語以上までの5つのグループに分けて調査を行った。

2.6 採点方法

採点は、筆者一人で、目標語の意味と推測された意味との近さと目標語との文脈における置き換えの可能性を基に行った。Liu & Nation (1985) では、後者よりも前者を基に、100%、80%、0%の3段階の採点が行われた。本研究では、両者に重きを置いているため、推測された各々の目標語の意味が教科書通りの語、または、英語母語話者があり得ると判断した語の意味の場合には100%、何とか考えられる場合や品詞・語法間違いの場合は半分50%、それ以外は0%とした。例えば、レストランでの *afford caviar* に対して、「支払う余裕がある」や「手が届く」、また、英語母語話者が *eat* もあり得ると判断したので「食べる」を100%、「買う」や「支払い」、「奢る」を50%、「見る」や「切り分ける」、「提供する」は0%とした。

3. 結果と考察

3.1 未知語の割合が50語に1語（98%の語彙カバー率）の場合

語彙サイズと推測成功率の平均の関係を見ると4,500語以上では80.71%、4,000-4,500語未満（以下、4,000語以上）では68.33%、3,500-4,000語未満（以下、3,500語以上）では57.87%であったが、3,000-3,500語未満（以下、3,000語以上）では35.94%であった（表2）。逆に、80%以上の推測成功率の参加者の語彙サイズの平均は4,455.56語、60-80%未満（以下、60%以上）では3,991.11語、60%未満では3,440.35語であった（表3）。その上、語彙サイズと推測成功率との相関は $r = .682$ であった（表4）。

表1 推測成功率の平均と標準偏差（パスセージ別）

$n = 110$		$n = 95$	
Passage A (1:50)		Passage B (1:100)	
<i>Mean</i>	61.02%	<i>Mean</i>	70.92%
<i>SD</i>	21.63	<i>SD</i>	22.14
Passage D (1:80)		Passage C (1:60)	
<i>Mean</i>	72.27%	<i>Mean</i>	57.82%
<i>SD</i>	22.41	<i>SD</i>	22.81

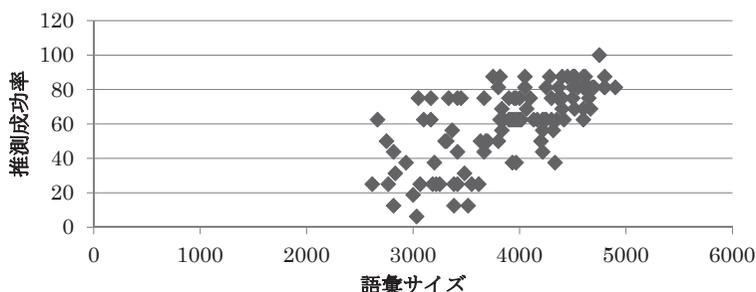


図1 未知語の割合が50語に1語の場合の語彙サイズと推測成功率との関係

語彙サイズの面から見ると、4,500語以上は「最適」、4,000語以上は「適切」と言える。推測成功率の面から見ると、80%以上の「最適な」推測成功率を収めるには、平均で4,455.56語の、60%以上の「適切な」推測成功率を収めるには、平均で3,991.11語のそれぞれ語彙サイズが必要であると言える。その両面を考慮に入れると、4,500語程度が「最適な」ものであると考えられる。また、4,000語以上が「適切な」語彙サイズであると言える。実際、相関係数が.682であることから、語彙サイズと推測成功率との間にはかなり高い相関があることがわかる。しかし、表から、いずれのパスセージにおいても、推測成功率別の語彙サイズの平均のSDの値が大きいがわかる。そこで、散布図で実態を確認することにする。すると、80%以上または60%以上の推測成功率を収めた参加者の語彙サイズにはかなりの幅があることがわかる。しかし、4,000語以上の語彙サイズがありながら、60%未満の推測成功率しか収められなかった参加者は5人と少なく、4,500語以上の場合には全くいない(図1)。このことから、4,500語程度が「最適な」、そして、4,000語程度が「適切な」語彙サイズと言える。

3.2 未知語の割合が：60語に1語（98.33%の語彙カバー率）の場合

語彙サイズと推測成功率の平均の関係を見ると、4,500語以上では86.74%、4,000語以上で69.71%、3,500語以上で61.14%であったが、3,000語以上では47.70%であった(表2)。逆に、80%以上の推測成功率の参加者の語彙サイズの平均は4,335.56語、60%以上では3,948.85語、60%未満では3,350.64語であった。その上、語彙サイズと推測成功率との相関は $r=.684$ であった(表4)。

表2 語彙サイズ別の推測成功率の平均と標準偏差(パスセージ別)

	4,500語以上	4,000-4,500未満	3,500-4,000未満	3,000-3,500未満	3,000未満
PA&PD	$n=23$	$n=30$	$n=27$	$n=22$	$n=8$
PC&PB	$n=7$	$n=25$	$n=25$	$n=28$	$n=10$
Passage A (1:50)					
Mean	80.71%	68.33%	57.87%	43.47%	35.94%
SD	8.20	12.81	18.88	23.02	15.93
Passage D (1:80)					
Mean	89.13%	83.00%	72.59%	50.46%	42.50%
SD	9.00	13.17	16.55	23.60	14.88
Passage C (1:60)					
Mean	86.74%	69.71%	61.14%	47.70%	27.86%
SD	9.61	15.74	20.21	17.61	17.32
Passage B (1:100)					
Mean	92.86%	84.00%	75.50%	61.61%	37.50%
SD	6.68	12.25	16.72	18.61	22.05

表3 推測成功率別の語彙サイズの平均と標準偏差 (パッセージ別)

	80%以上	60-80未満	60未満		80%以上	60-80未満	60未満
PA	<i>n</i> = 27	<i>n</i> = 45	<i>n</i> = 38	PB	<i>n</i> = 36	<i>n</i> = 40	<i>n</i> = 19
PC	<i>n</i> = 14	<i>n</i> = 30	<i>n</i> = 51	PD	<i>n</i> = 60	<i>n</i> = 26	<i>n</i> = 24
Passage A (1:50)				Passage B (1:100)			
<i>Mean</i>	4455.56	3991.11	3440.35	<i>Mean</i>	4116.20	3539.58	3062.28
<i>SD</i>	319.02	466.85	470.39	<i>SD</i>	416.34	410.42	433.43
Passage C (1:60)				Passage D (1:80)			
<i>Mean</i>	4335.56	3948.85	3350.64	<i>Mean</i>	4208.57	3805.13	3231.94
<i>SD</i>	430.24	461.08	455.84	<i>SD</i>	412.41	513.77	351.32

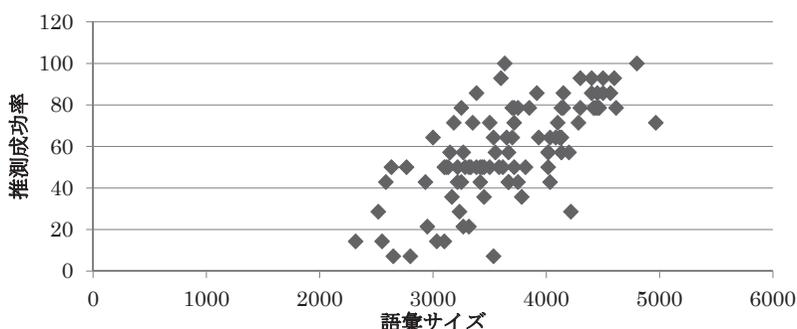


図2 未知語の割合が60語に1語の場合の語彙サイズと推測成功率との関係

以上の結果は50語に1語の場合と似たものであると言える。すなわち、語彙サイズの面から見ると、4,500語以上は「最適」、4,000語以上は「適切」、3,500語以上も「適切」と考えられる。推測成功率から見ると、「最適な」推測成功率を収めるには、平均で4,335.56語の、「適切な」推測成功率を収めるには、平均で3,948.85語のそれぞれ語彙サイズが必要であると言える。その上、相関係数が.684以上であることから、語彙サイズと推測成功率の間にはかなり高い相関があると言える。従って、表2、3から、「最適な」語彙サイズは、50語に1語の場合より、200語下げて、4,300語程度と考えてよからう。50語に1語の場合とは違い、3,500語以上でも「適切な」推測成功率を収めることがわかった。しかし、上記の通り、「適切な」推測成功率を取めたグループの語彙サイズの平均は3,948.85語である。そこで、散布図で実態を確認すると、50語に1語の場合と同様に、4,000語以上の語彙サイズがありながら、60%未満の推測成功率しか収められなかった参加者は7人と少ない。また、4,300語以上の場合には全くいない(図2)。このことから、4,300語程度が「最適な」、そして、4,000語程度が「適切な」語彙サイズと考えられる。

3.3 未知語の割合が80語に1語 (98.75%の語彙カバー率) の場合

語彙サイズと推測成功率の平均の関係を見ると、4,500語以上では89.13%、4,000語以上で83.00%、3,500語以上で72.59%であったが、3,000語以上では50.46%であった(表2)。逆に、80%以上の推測成功率の参加者の語彙サイズの平均は4,208.57語、60%以上では3,805.13語、60%未満では3,231.94語であった(表3)。その上、語彙サイズと推測成功率との相関は $r = .689$ であった(表4)。

表4 語彙サイズと推測成功率との相関（パッセージ別）

Passage A (1:50)	Passage B (1:100)	Passage C (1:60)	Passage D (1:80)
.682**	.715**	.684**	.689**

注：** $p < .01$

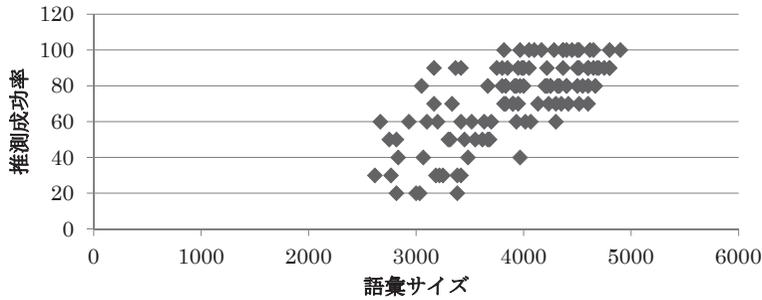


図3 未知語の割合が80語に1語の場合の語彙サイズと推測成功率との関係

以上のことから、80語に1語の未知語の割合では、上記2つの割合とは異なる結果が得られることがわかった。すなわち、語彙サイズの面から見ると、4,000語以上からは「最適」、3,500語以上は「適切」と言える。推測成功率から見ると、「最適な」推測成功率を収めるには、平均で4,208.57語の、「適切な」推測成功率を収めるには、平均で3,805.13語のそれぞれ語彙サイズが必要であると言える。さらに、語彙サイズと推測成功率との相関は上記2つの未知語の割合と同様に.689とかなり高いものであることがわかった。それゆえ、表2、3から、「最適な」語彙サイズは、60語に1語の場合より、100語下げて、4,200語程度が「最適」なものと考えられる。しかし、4,000語程度とも考えられる。「適切な」語彙サイズに関しても、3,500語以上とも、3,800語程度とも考えられる。そこで、散布図で実態を確認すると、3,500語以上の語彙サイズがありながら、60%未満の推測成功率しか収められなかった参加者の数は5人と少ない。また、4,000語以上の場合には全くいない(図3)。このことから、4,000語程度が「最適な」、そして、3,500語程度が「適切な」語彙サイズと考えられる。

3.4 未知語の割合が100語に1語（99%の語彙カバー率）の場合

語彙サイズと推測成功率の平均の関係を見ると、4,500語以上では92.86%、4,000語以上で84.00%、3,500語以上で75.50%、3,000語以上でも61.61%であった(表2)。逆に、80%以上の推測成功率の参加者の語彙サイズの平均は4,116.20語、60%以上では3,539.58、60%未満では3,062.28語であった(表3)。その上、語彙サイズと推測成功率との相関は $r = .715$ であった(表4)。

以上のことから、100語に1語の未知語の割合は、80語に1語の場合と似たような結果であるが、更に「適切な」ものであることがわかった。すなわち、語彙サイズの面から見ると、4,000語以上からは「最適」、3,000語以上からは「適切」と言える。推測成功率から見ると、「最適な」推測成功率を収めるには、平均で4,116.20語の、「適切な」推測成功率を収めるには、平均で3,539.58語のそれぞれ語彙サイズが必要であると言える。その上、語彙サイズと推測成功率との相関は.715と高いものであることがわかった。従って、表2、3から、4,100語程度が「最

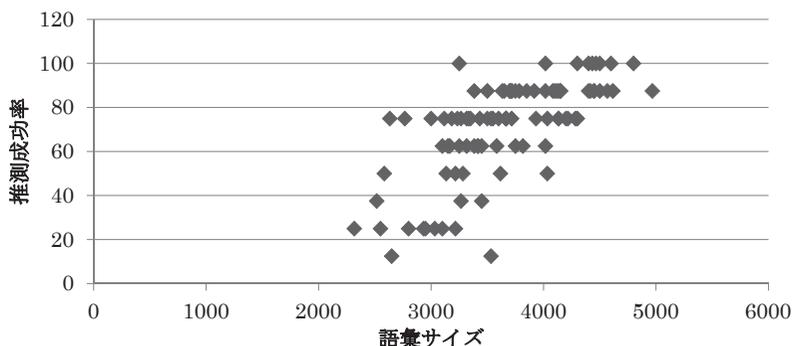


図4 未知語の割合が100語に1語の場合の語彙サイズと推測成功率との関係

適」なものであると考えられる。しかし、4,000語程度とも考えられる。「適切な」語彙サイズに関して言うと、3,000語以上とも、3,500語程度とも考えられる。そこで、散布図で実態を確認すると、3,000語以上の語彙サイズがありながら、60%未満の推測成功率しか取められなかった参加者の数は少なく、4,000語以上の場合には一人だけである(図4)。このことから、4,000語程度が「最適な」、そして、3,000語程度が「適切な」語彙サイズと言える。

4. 全般的な考察

4つの未知語の割合についての結果から、50語に1語の場合と60語に1語の場合、80語に1語の場合と100語に1語の場合がそれぞれ似た結果となることがわかった。語彙カバー率に換算すると、60語に1語の場合は98.33%、80語に1語の場合は98.75%とそれぞれなり、その差はわずか.42%にすぎないが、大きな違いをもたらすと言える。また、語彙サイズと推測成功率との相関は、いずれの未知語の割合でも.682以上とかなり高く、未知語の割合が下がるにつれて高くなり、100語に1語の場合では.715という高い値であることがわかった。さらに、4,000語程度の語彙サイズがあれば、80語に1語の場合でも、「最適な」ものと言える。100語に1語の場合では、3,000語程度の語彙サイズがあれば「適切な」ものと言える。逆に、50語に1語の場合でも、4,500語程度の語彙サイズがあれば、「最適な」ものと考えられる。しかし、「最適な」語彙サイズがあれば、「最適な」推測成功率を取められる訳では必ずしもない。逆に、「適切な」語彙サイズでも、「最適な」推測成功率を取められる場合もあるということが散布図からも確認できる。

5. 結 論

本研究の結果、50語に1語の割合で80%以上の推測成功率を取めるためには、4,500語程度が必要であることがわかった。これは、Hirsh & Nation (1992) の約5,000語に近い値である。80語に1語になると、80%以上の推測成功率を取めるためには、4,000語程度が必要であることがわかった。100語に1語になると、さらに、3,000語程度の語彙サイズがあれば、60%以上の推測成功率を取めることができることがわかった。本研究からは、「最適な」未知語の割合についての結論は出せない。しかし、散布図から、100語に1語の未知語の割合が調査した4つの中では最も「適切な」ものであると言える。100語に1語の未知語の割合を語彙カバー率に換算する

と99%になる。Carver (1994) が提唱する99%の語彙カバー率を支持する結果となった。どの割合でも、語彙サイズと推測成功率との相関は .682以上とかなり高いものであることが確認された。しかし、同時に、語彙サイズだけでは推測成功率は決定されないこともわかった。このことは、Hu & Nation (2000) の主張を支持する結果となった。

以上のことから、題材が高校1年生用の検定教科書の場合、100語に1語の未知語の割合が調査した4つの中では最も「適切な」ものであり、語彙サイズが4,000語程度ある学習者には、80語に1語の割合でも「最適な」ものであることがわかった。さらに、4,500語程度の語彙サイズがある学習者であれば、50語に1語の割合でも、平均で80%以上の推測成功率を収めることができることがわかった。但し、未知語の割合や語彙サイズは、推測成功率に対して非常に大きな影響を与える要因であろうが、あくまでも一要因であることを忘れてはいけない。

6. 教育上の示唆と今後の課題

本研究により、題材が高校1年生用の教科書の場合、100語に1語の未知語の割合では、3,000語程度の語彙サイズがあれば、推測成功率の平均は60%以上であるという教育上の示唆を得た。教師は、学習者に未知語の推測をさせる場合には、題材のレベル、未知語の品詞や割合、学習者の語彙サイズなどを十分に考慮すべきである。その上で、推測をする練習をさせれば、覚えるべき語彙数を減らすことができる可能性がある。

今後の課題としては、本研究の参加者が少なかったので、この結果が一般化できるのかどうかをもっと多くの参加者を得て再調査する必要がある。また、同一のパッセージを使っても同じ結果となるのかどうかの調査も必要である。

追記：本研究の参加者からは、調査の同意が得られている。

引用・参考文献

- Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, 26(4), 413-437.
- Clarke, D. F., & Nation, I. S. P. (1980). Guessing the meaning of words from context: Strategy and techniques. *System* 8(3), 211-220.
- Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence from L1 readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181-202). Norwood, NJ: Ablex.
- Haynes, M. (1993). Patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes, & J. Coady J. (Eds.), *Second language reading and vocabulary learning* (pp. 46-62). Norwood, NJ: Ablex.
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Basingstoke & London: Macmillan.
- Laufer, B. (1997). The lexical plight in second language reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). New York: Cambridge University Press.

- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Liu, Na, & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33-42.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines* 5(1), 12-25.
- 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.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59-82.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43.
- 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.
- van Parreren, C. F., & Schouten-van Parreren, M. C. (1981). Contextual guessing: a trainable reader strategy. *System* 9(3), 235-241.
- Webb, S. (2010). Pre-learning low-frequency vocabulary in second language television programmes. *Language Teaching Research*, 14(4), 501-515.
- West, M. (1955). *Learning to read a foreign language* (2nd ed.). London: Longman.
- Yashima, H. (2001). Word inference ability in Japanese EFL learners. *ARELE*, 12, 101-110.
- Yashima, H. (2002). Factors affecting guessing unknown word meanings from context. *ARELE*, 13, 151-160.
- Yashima, H. (2003). Analysis of senior high school students' ability to guessing unknown word meanings from context. *ARELE*, 14, 161-170.
- 羽鳥博愛. (1979). 「リーディング・テスト作成上の留意点」. 『英語指導法ハンドブック④評価』 (pp. 107-121). 東京：大修館.
- 八島 等. (1996). 「検定教科書での未知語の推測」. 『語学教育研究所紀要』, 第10号, 35-42.
- 八島 等. (1999). 「未知語の割合と文脈からの意味の推測の関係及び推測に影響を与える他の要因」. 『語学教育研究所紀要』, 第13号, 31-49.
- 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』, 16, 29-42.

APPENDIX

調査で用いた Passages A (無意味語の後に元の語を提示)

In the big city of New York, adventure is always waiting. At every corner, eyes look around us, or fingers are pointed at us. Adventure is there. But few of us ever see it.

Rudolf Steiner was a true (1) tampender (adventurer). There were few evenings when he did not go out looking for something different. He was always interested in what might be waiting around the next corner.

One evening Rudolf was walking slowly along a street in the older part of the city. Many people were walking along the street that night. Some were going home. Others were going to have dinner at some restaurant.

He passed a busy restaurant and saw next to it an open door. Above the door was a (2) hesument (sign) for a doctor's office. A very large man stood at the door. He was (3) santaring (offering) cards to those who passed by. Rudolf had seen such people before. The card would show the name of the doctor in the office on the third floor. Rudolf took it quickly without looking at it.

When he had walked on further, he looked at the card. He was surprised. He turned it over and looked again with interest. On one side there was (4) modendation (nothing). On the other side were three words: "The Green door."

Three steps beyond, another man threw down the card he had received. Rudolf (5) finemined (picked) it up. There was the doctor's name, with the street and number. This was what Rudolf had (6) kindemed (expected) to find on his own card. Rudolf turned around and walked past the doctor's office again. The man

handed him another card. Once again the card read, "The Green Door." Three or four other cards were (7) yo leaking (lying) in the street. People had thrown them away. He looked at them. Every one had the doctor's name on it.

Rudolf went back to where he had received the card and looked up at the building. He believed that adventure was calling him. There was a small restaurant on the first floor. On the second floor were rooms where people lived. Above that was the doctor's office.

Rudolf entered the building and walked up to the second floor. He looked around and saw a green door. He walked straight to the green door and (8) nuro lved (knocked) loudly. He heard a soft sound and then the door opened slowly. A girl not yet twenty stood there.

(397 words from *CROWN English Series I*)

調査で用いた Passages C (無意味語の後に元の語を提示)

It was 20 years ago and I was living in Paris. I was earning just enough money to live. She had read a book of mine and had written to me about it. I answered, thanking her, and then I received from her another (1) temitionality (letter) saying she would like to see me; but the only free moment she had was on Thursday. She asked if I would take her to a little luncheon at Foyot's. Foyot's was such an expensive restaurant that I had never even thought of going there. But I was (2) thmanged (flattered) and I couldn't say no to a woman. I had 80 francs to last me the rest of the month and a luncheon should not be more than 15.

I answered that I would meet her at Foyot's on Thursday at half past twelve. She was not young or pretty but she liked to talk. Since she was talking about me, I was happy to (3) defolive (listen).

I was surprised when the menu was brought, for the prices were very high.

"I never eat anything for luncheon," she said.

"Oh, don't say that!" I answered.

"I never eat more than one thing. I wonder if they have any salmon."

Well, it was early in the year for salmon, but I asked the waiter if there was any. Yes, a beautiful salmon had just (4) nescovered (come) in. I ordered it for her. The waiter asked her if she would have something else.

"No," she answered, "I never eat more than one thing. Unless you have a little caviar."

I knew I could not (5) kugduce (afford) caviar, but I could not very well tell her that. I told the waiter to bring caviar. For myself, I chose the cheapest dish on the menu and that was a mutton chop.

"I think it's not a good idea to eat meat," she said.

"I don't know how you can work after eating heavy things like chops."

Then there was a question of drink.

"I never drink anything for luncheon," she said.

"I don't, either," I answered.

"Except white wine," she said. "These French white wines are wonderful."

"What would you like?" I asked.

"My doctor won't let me drink anything but champagne."

I turned pale. I ordered half a bottle. I mentioned that my doctor told me not to drink champagne.

"What are you going to drink, then?"

"(6) Medolment (Water)."

She ate the caviar and the salmon. She talked about art and literature and music. But I wondered what the (7) centration (bill) would come to.

(420 words from *CROWN English Communication I*)

—2020年 9月24日 受理—