

# 語彙知識とそのテキスト理解との関係：中国語・韓国語を母語とするL2言語学習者と日本語母語話者の比較研究

著者名(日)	堀場 裕紀江
雑誌名	言語科学研究：神田外語大学大学院紀要
巻	21
ページ	23-46
発行年	2015-03
URL	<a href="http://id.nii.ac.jp/1092/00001169/">http://id.nii.ac.jp/1092/00001169/</a>

# 語彙知識とそのテキスト理解との関係： 中国語・韓国語を母語とするL2学習者と 日本語母語話者の比較研究

堀場 裕紀江

## 要旨

本研究では、中国語（50名）または韓国語（20名）を母語とする第二言語（L2）学習者と日本語母語話者（40名）を対象に、語彙知識とそのテキスト理解との関係を調べた。語彙知識の広さと深さは語義組み合わせテストと語連想選択テストによってそれぞれ測定した。テキスト理解は再生テストと要約完成テストによって測定し、語彙知識との関係を回帰分析によって調べた。主な結果は以下のとおりである。語彙知識の構成要素の相対的な強度について、中国語話者と韓国語話者の間およびL2学習者と母語話者の間に違いがある。L2学習者のテキスト理解は語彙知識の広さ・深さの両方と中程度の相関があり、広さが深さに比べて強い説明要因であるが、母語話者の場合はそうではない。語彙知識の要素のうちsyntagmatic associationが中国語話者にとって特に重要であるが、韓国語話者と母語話者にとってはそうではない。よって、L2学習者の語彙知識とそのテキスト理解との関係には母語に関する要因が複雑な影響を与えていることが推察される。

キーワード：語彙知識の広さ、語彙知識の深さ、読解、第2言語としての日本語、母語背景

## 1. はじめに

語彙知識はテキスト理解を含めた言語運用の中で最も重要な要因の一つである（Anderson & Freebody, 1981; Huckin, Hayes, & Coady, 1993; Nagy & Anderson, 1984; Stanovich, 1986）。テキストは、句、節、文に構成された語の集合体である。そのため、語彙知識が豊富であればあるほど、テキスト内でより多くの既知語に出会う可能性が高まり、テキストに記されているアイデアや関係を理解することが容易になる。一方で、テキスト処理を行うことによって、テキスト内の語や語と語の関係に関する知識を獲得し発達させる機会を得

ることができる。このように、語彙知識とテキスト理解の関係は処理と発達という観点において相互依存の関係にある。しかし、このような一般的なこと以外、非母語（L2）話者について語彙知識がテキスト理解にどのように関わるかは明らかでない。L2語彙知識の多面性については益々多くの研究（Greidanus, Beks, & Wakely, 2005; Greidanus & Niehuis, 2001; Henriksen, 2008; Laufer & Paribakt, 1998; Nissen & Henriksen, 2006; Schmitt & Meara, 1997; Vermeer, 2001; Zareva, 2007; Zareva, Schwanenflugel, & Nikolova, 2005）が行われているが、それらは、L2研究一般と同様、主に英語を対象としており、特に非印欧祖語に関する研究は殆ど行われていない。さらに、語彙知識の多面性がL2テキスト理解にどのように関わるかを調べた研究（Qian, 1999, 2002）は極めて少ない。本研究はL2としての日本語の語彙知識を、異なるL1背景を持つL2学習者と母語話者の比較を通して調べ、日本語における語彙知識とテキスト理解の関係を探る、という2つの目的を達成すべく行われた。

## 2. 語彙知識の多面性

語彙知識は多面的でその中に異なる種類の情報を含んでいる（Meara, 1996; Nagy & Scott, 2000; Nation, 1990, 2001）。語彙知識についてよく用いられる識別として語彙知識の広さと深さという識別がある（Henriksen, 1999; Read, 2004; Wesche & Paribakt, 1996）。語彙知識の広さは一般に、人がある程度まで知っている語がいくつあるかという観点から量的に測定される。例えば、英語の語彙テストとしてよく知られているVocabulary Levels Test (VLT) (Nation, 1990, 2001) は、学習者が異なる頻度レベル（2,000, 3,000, 5,000, 10,000, およびUniversity Word List）の語の基本的な意味を知っているか否かを査定するように設計されている。具体的には、受験者は語とその定義（例 *wine - a drink*）を6つの語と3つの定義からそれぞれ選んでマッチングするように求められる。テスト得点は学習者の語彙サイズの目安として解釈され使用される。文の空所を補充させる形式の伝統的な語彙テストもまた、語彙知識を表面的に査定するという点で語彙知識の広さを問題にしていると考えられる（Wesch & Paribakht, 1996）。

これに対し、語彙知識の深さは知識の質に関するものである。研究者によって異なるアプローチが用いられている。語彙知識の質を語の意味に関する知識

の正確さとしてとらえる研究者 (the Vocabulary Knowledge Scale) (Wesche & Paribakht, 1996) もいるが、表記、音韻、形態、統語、意味、連語、語用などの様々な種類の情報について語の知識がどのくらい包括的であるかという観点からとらえる研究者もいる。さらにまた、語彙知識を既知語のネットワーク (メンタルレキシコン) (Aitchison, 1994) として概念化するアプローチもある。このアプローチによると、深さは語をメンタルレキシコンのネットワークに結合することであり、関連語と結びつけたり区別したりする能力と関係している (Haastrup & Henriksen, 2000; Read, 2000, 2004)。このアプローチは対象語について最初に思いつく語を産出させる語連想タスクから始まったものである。L1研究では、子どもは思春期までに応答のパターンがsyntagmatic associationから成人のようなparadigmatic associationへと変化することや、成人の母語話者の連想パターンにはsyntagmatic associationへの回帰がみられることが示されている (Aitchison, 1994; Namei, 2004)。また、L2学習者については、成人の母語話者と比べてより多様で不安定な連想を産出することや、連想パターンには学習者の言語能力レベル、項目 (語) の頻度や品詞などの要因が関係していることが示されている (Henriksen, 2008; Meara, 1984; Nissen & Henriksen, 2006; Wolter, 2001; Zareva, 2007; Zareva et al., 2005)。

語連想としての語彙知識という概念は受容テストでも採用されている。例えば、Read (1993, 1998) によって開発された the Word Associates Test (WAT) では、受験者は対象語と関係のある語 (連想語と呼ばれる) を多肢選択式で選ぶように求められる。対象語と連想語の関係は3種類ある。すなわち、paradigmatic (対象語と連想語が同義語であるか意味的に上位語・下位語の関係になっている；例 *table - furniture*)、syntagmatic (対象語と連想語が一文の中で連語や共起語として現れる；例 *erase - data*) およびanalytic (連想語が対象語の意味の重要な要素で辞書の定義で使われる可能性が高い；例 *edit - publishing*) である。この語連想テストは語彙知識の質的側面を測定するのに有効であることが示されており、他言語でも採用され調査されている (例 Greidanus et al., 2005; Greidanus & Nienhuis, 2001)。例えば、Greidanus et al. (2001, 2005) では、受容式の語連想テストを用いてフランス語のL2学習者と母語話者を比較分析した結果、頻度レベルがL2学習者と母語話者の両方に影響を与えることが分かった。また、この研究ではL2習熟度レベルやL1背

景（オランダ語と英語）に関わらず paradigmatic association の得点が最も高く、syntagmatic association は analytic association よりも得点が低いという結果も得られた (Greidanus et al., 2005)。

語彙知識の広さと深さは別個に存在するものではないが、L2 学習者の語彙知識に関する研究においてこの 2 つの側面の区別は有用なものとなっている (Henriksen, 2008; Schmitt & Meara, 1997; Vermeer, 2001; Zareva, 2007; Zareva et al., 2005)。例えば、Zareva et al. (2005) では、外国語としての英語 (EFL) 学習者 (ブルガリア人) の語彙知識は、L2 習熟度が上がるにつれて、広さ (既知語の数と頻度レベル) および、連想の母語話者らしさを除く深さ (連想の数とその同種性) において、母語話者の語彙知識に似通ってくることを示されている。しかし、Zareva (2007) では、2005 年と同じ方法で調査したところ、第二言語としての英語 (ESL) 学習者 (多様な母語背景) と母語話者との間には、量 (連想の数とその同種性) および質 (paradigmatic, syntagmatic, phonological 連想の比率) において違いがみられなかった。この 2 つの研究結果の矛盾は、言語状況 (第二言語か外国語か) と対象者集団 (同質か多様か) などに関係するものであると考えられる。これらの可能性を明らかにするためには更なる研究が必要である。

### 3. 語彙知識と L2 読解に関する先行研究

語彙サイズまたは語彙知識の広さは、L2 読解を予測するのに有効な要因である (Huckin, Hayes, & Coady, 1993; Laufer, 1992; Nation, 2001; Staehr, 2008)。実際、L2 読解研究では、限られた L2 能力の読み手の L2 読解は、(L1 を基盤とする) 一般的な読解スキルに比べて L2 言語知識 (語彙と文法) の方が説明要因として有効であり、中でも語彙は文法よりも重要な要因であることが繰り返し示されている (Bernhardt & Kamil, 1995; Lee & Schallert, 1997)。また、L2 読解における語認識のような低次レベルの処理 (知識とは対照的に) の効率性を強調する研究者もいる (Koda, 1996; Segalowitz, Poulson, & Komeda, 1991)。

しかし、語彙知識の異なる側面がどのように L2 読解に影響を与えるかについてはごく限られた研究しか行われていない。Qian (1999, 2002) では、語彙知識のどちらの側面 (広さか深さ) が EFL 学習者の学術テキスト読解により重要であるかを調べた結果、広さ (VLT で測定) と深さ (WAT で測定) のどちら

も読解を説明する重要な要因であるが、広さよりも深さの方が説明力が大きいという結果を得ている。語彙知識の深さ（WATで測定）がL2テキストにおける未知語の意味推測に関係する可能性も示唆されている（Nassaji, 2006）。しかし、Qianの研究で報告された語彙知識とテキスト理解との関係が他のL2学習者グループや他の種類のテキストに一般化できるかについては検証されていない。この課題についても更なる研究が必要である。

本研究では、語彙知識とそのテキスト理解との関係を、中国語と韓国語という2つの異なるL1背景を持つL2日本語学習者を対象に調査した。語彙知識は広さテストと深さテストの2種類のテストによって測定した。テキスト理解は2つの説明文を用いて測定した。L2学習者の結果は、L1背景による比較分析を行い、さらに母語話者との比較も行った。母語話者の応答パターンはコミュニケーションにおける適切な言語運用の土台となる語彙知識の特徴を反映していると考え、母語話者データをベースラインとして用いた。

本研究で扱う言語は非印欧祖語である日本語、中国語、韓国語で、この3つの言語には対照的な共通点と相違点がある。簡単に述べると、日本語では表意文字である漢字と表音文字である2種類の仮名を組み合わせ用いる。漢字は形態素や単語を表すのに対し、仮名は音節を表す。形態統語的には、日本語は格マーカースystemを持つSOV形式の膠着語である。中国語は日本語の漢字と同様の表意文字を用いるが、日本語とはかなり異なるSVO形式の独立語である。これに対し、韓国語は日本語や中国語と異なるアルファベット式の正書法を用いるが、日本語に似たSOV形式の膠着語である。先行研究の結果から、すでに確立し自動化された言語（L1）知識を有している成人のL2学習者においては、そのL1に基づいた知識や手順が新しい語の処理やそれに伴う新しい言語（L2）知識の発達に影響を与えることが示唆されている（Koda, 2005; Randall, 2007）。中国人学習者は、日本語のテキストを読む際に、漢字で書かれた内容語の意味をより容易に導き出すことができる可能性がある（Matsunaga, 1999）。一方で、韓国語学習者は、日本語の文を読む際に統語的情報（語順や格マーキングなど）の処理がより容易にできる可能性がある（Koda, 1993）。これらの事象はそれぞれのL1背景を持つ上級学習者からもしばしば非公式に報告されている。

本研究のために設定した質問は以下の2つである。

- 1) L2学習者の語彙知識はL1背景によって異なるか。また、それは母語話者の場合と異なるか。
- 2) L2学習者の語彙知識とL2テキスト理解の関係はL1背景によって異なるか。また、それは母語話者の場合と異なるか。

## 4. 本研究

### 4.1. 研究方法

#### 4.1.1. 対象者

中国語母語話者50名と韓国語母語話者20名からなるL2日本語学習者（男性24名・女性46名、平均年齢23.0歳）、および、日本語母語者40名（男性3名・女性37名、平均年齢20.3歳）が調査に参加した。L2学習者は全員、日本国内の語学学校あるいは大学の日本語プログラムにある中上級日本語コースを履修していた。履修コースは各プログラム内のプレースメントテストによって振り分けられていた。日本語母語話者は全員、大学の学部生で日本語学のコースを履修中であった。

#### 4.1.2. 語彙知識の測定

対象とした語は、旧日本語能力試験基準（Japan Foundation, 2002）による1、2級レベルの内容語48個である<sup>(注1)</sup>。対象語を付表Aに示す。語彙知識の測定には、語の基本的意味についての知識を測定する「広さ」テストと、語連想についての知識を測定する「深さ」テストの2種類の語彙テストを用いた。

広さテストは、対象語とその簡単な定義を結びつけるという形式である（c.f., Nation's VLT, 1990, 2001）。テスト問題はそれぞれ6個の語（3個の対象語と3個の錯乱語）と3つの簡単な定義で構成され、6語はみな同じ品詞、同じ頻度レベルに属すが、意味領域は異なる。6語は日本語の辞書などで一般に用いられるアイウエオ順で提示した。定義は簡単な語句を用いて書かれていた。

深さテストは、対象語について3つの連想語、すなわち、paradigmatic, syntagmatic, analyticの関係を持つ語をそれぞれ1つずつ7つの選択肢から選ぶという形式である（c.f., Read's WAT, 1993）。錯乱語（語）は、対象語と同レベルまたはより高頻度の語で、正答語と同じ種類の品詞に属する語が1つ以上含まれていた<sup>(注2)</sup>。対象語48個を頻度、品詞、語種、および意味領域を考慮し

て2つに分け、2バージョンのテストを作成した。調査協力者は各グループ内でランダムに振り分けられてどちらか一方のバージョンのテストを受けた。この手順を踏んだのは、複数の測定を可能にするための時間的制約と人数の確保を考慮したためである。語彙テストの問題例を付表Bに示す。

テスト材料は、経験のある日本語教師で日本語学、応用言語学、あるいはその両方で博士号・修士号を持つ日本語母語話者のチームによって開発された。錯乱肢（語）は、旧日本語能力試験のための語彙リストから、複数の辞書（例 Shibata et al., 1995, 2002; Yamaguchi, 2003）や日本語能力試験に関する学習教材を参考にして選定した。テストで使用した語や文は慣例の漢字仮名混じりで表記し、漢字には振り仮名をつけた。対象語と答えで同じ漢字が使われる場合は、対象語を仮名で表記した。このような手法を用いたのは、テストで測定するのは語彙知識であって漢字知識ではないからである。テストでは、高頻度の語を提示してから低頻度の語を提示し、各頻度レベル内ではランダムに提示した。テスト材料については、その開発の過程でパイロット調査を複数回行い、将来L2日本語の教師・研究者になることを目指している大学院生（母語話者および非母語話者からなる）を対象に行った予備調査の結果をもとに最終決定した。テストの所用時間は広さテスト10分、深さテスト15分である。

#### 4.1.3. 読解の測定

調査協力者の興味関心および教育的価値を考慮して2つの短い説明文、「旅行」テキストと「文化」テキスト（文の数23・24、文字数747・962）を選出した。両テキストを付表Cに示す。各グループの調査協力者はランダムに配分され、半数が一方のテキストを、残りの半数がもう一方のテキストを読んだ。テキストで使用される漢字はすべて振り仮名をつけて表記した。

読解は再生と要約完成の2つの手法で測定した。再生タスクでは、読んだテキストの内容について覚えていることを全て母語で書くようにと指示した<sup>(注3)</sup>。要約完成タスクでは、5つの文の空所を埋めて完成するようにと指示した（5つの空所それぞれにつき4つの選択肢）<sup>(注4)</sup>。調査協力者はテキストを読む前に、内容理解のためにテキストを読むように指示され、後でテキストについての質問があることも伝えられた。読解テストの材料についても複数回のパイロット調査を行った。所要時間はテキストの読み（10分）と3つのタスク（すなわち



錯乱タスク、再生タスク、要約完成タスク）を合わせて30分とした。

#### 4.1.4. 手順

L2学習者は通常の授業時間内に調査に参加した。母語話者は参加者人数確保のため設けた複数回の調査のうち1回に参加した。調査では、調査助手（試験監督）が調査全体の目的と手順を説明したあと、調査協力者は全てのテスト材料の入った封筒を受け取った。調査協力者は試験監督の指示に従うようにと指示された。調査協力者は初めに背景アンケートを記入した後、3つのテストを広さテスト、読解テスト、深さテストの順で受けた。データ収集は、調査協力者の所属する機関の教室で試験監督2名の在席により実施した。L2学習者からのデータ収集については通常クラスの担任教師も在席した。

#### 4.1.5. 分析

語彙テストと要約完成テストの応答データは全て2名の評定者が個別に採点した。L2学習者の再生データは各言語の母語話者で高度な日本語力を持つ者2名が和訳した。再生データは全てオリジナルのテキストに含まれるイベントをリストしたもの（Trabasso, Secco, & van den Broek, 1984）を基準にして2名の評定者が個別に採点した。旅行テキスト、文化テキストに含まれるイベントの数はそれぞれ65、66である。語彙テストと要約完成テストの評定者間信頼性は1.00であった。再生テストの評定者間信頼性は.93で不一致の部分については評定者1名の再採点により解決した。

### 4.2. 結果

ここで報告する結果は、特に言及がない限り、全て.05基準の信頼性を満たしている。

#### 4.2.1. 語彙知識の広さと深さ

広さテストと深さテストの結果を表1に示す。テスト信頼性は広さテストで  $KR = .87$  (Native:  $KR = .36$ , L2 group:  $KR = .98$ ), 深さテストで  $KR = .96$  (native:  $KR = .87$ , L2 group:  $KR = .99$ ) であった。

表1

グループごとの頻度による語彙知識の広さ・深さテスト得点の平均と標準偏差

	N		広さ		深さ	
			M	SD	M	SD
NS	40	高頻度語	99.2	1.8	88.9	4.6
		低頻度語	98.9	1.7	92.2	4.5
		全体	99.0	1.2	90.7	3.4
NNS						
中国語話者	50	高頻度語	80.9	14.5	67.8	11.1
		低頻度語	71.7	13.0	66.0	9.9
		全体	75.7	12.3	66.8	9.4
韓国語話者	20	高頻度語	66.9	22.8	56.4	20.2
		低頻度語	40.7	24.8	47.1	22.2
		全体	52.2	22.7	51.1	20.4

注 NS = 母語話者; NNS = 非母語話者

広さテストについて、グループを被験者間要因、頻度を被験者内要因とする多変量分散分析 (MANOVA) を行った結果、グループ (M-parameter = 1.67,  $F[2,107] = 89.39$ ) および頻度 (M-parameter = 1.13,  $F[1,107] = 120.89$ ) による有意な効果があった。また、グループと頻度の間の交互作用の効果も有意であった (M-parameter = .76,  $F[2,107] = 40.61$ )。グループ間比較を学生t検定 (Student t-test) で行ったところ、頻度にかかわらず、母語話者が中国語話者 (mean difference = 18.31,  $\pm 95\%CI = 12.51\sim 24.11$ ) および韓国語話者 (mean difference = 32.25,  $\pm 95\%CI = 24.75\sim 39.74$ ) に比べて正答率が高く、中国語話者が韓国語話者を上回っていた (mean difference = 13.94,  $\pm 95\%CI = 6.71 \sim 21.18$ )。頻度の効果は、中国語話者 (Paired-t[49] = 5.41,  $p = .0001$ ) および韓国語話者 (Paired-t[19] = 7.56,  $p = .0001$ ) に見られたが、母語話者には見られなかった (Paired-t[39] < 1)。

深さテストについて、グループとテストバージョンを被験者間要因とする2要因分散分析 (ANOVA) を行った結果、両テストバージョンは同等であることが確認された (テスト:  $F[1,104] < 1$ , n.s.; グループ・テスト交互作用:  $F[2,104] = 1.44$ , n.s.)。広さテストと同様、MANOVAを行った結果、グループ (M-parameter = 1.82,  $F[2,107] = 97.30$ )、頻度 (M-parameter = .08,  $F[1,107] = 8.32$ )、グループと頻度の交互作用 (M-parameter = .25,  $F[2,107]$ )

= 13.54) のそれぞれに有意な効果が検出された。Student t-testを行ったところ、中国語話者が韓国語話者を上回り (mean difference = 15.67,  $\pm 95\%CI = 9.96 \sim 21.38$ )、母語話者が中国語話者 (mean difference = 23.92,  $\pm 95\%CI = 19.34 \sim 28.49$ ) と韓国語話者 (mean difference = 39.58,  $\pm 95\%CI = 33.67 \sim 45.49$ ) を上回っていた。頻度の効果は、韓国語話者については有意であった (Paired-t[19] = 3.41,  $p = .003$ ) が、中国語話者 (Paired-t (49) = 1.42,  $p = .16$ ) には有意でなかった。母語話者については逆方向の頻度の効果が見られた (Paired-t[39]=3.31,  $p = .002$ )。

#### 4.2.2. 語連想の種類

語連想テストにおける連想の種類 (すなわち、paradigmatic, syntagmatic, and analytic association) についての結果を表2に示す。記述的には、中国語話者はparadigmaticの正答率が最も高く、次にanalyticが高く、syntagmaticが最も低かった。韓国語話者はanalyticの正答率が最も高いが、paradigmaticとsyntagmaticも僅差であった。一方、日本語母語話者はparadigmaticの正答率が最も高く、僅差でsyntagmaticが続き、analyticが最も低かった。

表2

グループごとの連想の種類による語彙知識の深さテスト得点の平均と標準偏差

	N		Paradigmatic		Syntagmatic		Analytic	
			M	SD	M	SD	M	SD
NS	40	高頻度語	92.0	7.8	93.7	7.8	81.1	12.7
		低頻度語	95.8	5.2	92.0	9.3	88.7	9.8
		全体	94.3	4.5	92.7	6.7	85.5	9.5
NNS								
中国語話者	50	高頻度語	74.1	16.1	64.9	17.3	64.5	16.2
		低頻度語	74.3	15.8	56.3	16.2	67.4	13.4
		全体	74.2	12.8	59.8	13.1	66.3	11.6
韓国語話者	20	高頻度語	56.5	20.1	61.3	24.9	51.4	22.6
		低頻度語	47.7	25.3	40.0	23.4	53.6	25.3
		全体	51.7	21.0	49.0	20.8	52.7	23.1

注 NS = 母語話者; NNS = 非母語話者

連想の種類についての全体正答率を、グループを被験者間要因、連想の種類を被験者内要因とする多変量分散分析によって分析した。その結果、グループ (M-parameter = 1.91,  $F[2,107] = 101.99$ )、連想の種類 (M-parameter = .27,  $F[2,106] = 14.49$ )、および、グループと連想の種類の交互作用 ( $F[4,212] = 9.64$ ) の有意な効果が検出された。中国語話者では、paradigmaticがanalytic (Paired- $t[49] = 4.18$ ,  $p = .0001$ ) とsyntagmatic (Paired- $t[49] = 6.90$ ,  $p = .0001$ ) に比べて高く、analyticはsyntagmaticよりも高かった (Paired- $t[49] = 3.12$ ,  $p = .003$ )。また、韓国語話者では、連想の種類による有意な効果は見られなかった。analyticはparadigmatic (Paired- $t[19] < 1$ ) およびsyntagmatic (Paired- $t[19] = 1.38$ ,  $p = .18$ ) と比べて高くはなく、paradigmaticとsyntagmaticの間にも有意差は見られなかった (Paired- $t[19] < 1$ )。日本語母語話者については、paradigmaticとsyntagmaticがanalyticと比べて高く (Paired- $t[39] = 4.79$ ,  $p = .0001$ ; Paired- $t[39] = 3.55$ ,  $p = .001$ )、paradigmaticとsyntagmaticの間にも有意差は見られなかった (Paired- $t[39] = 1.32$ , n.s.)。

#### 4.2.3. 語彙知識の要素間の関係

L2学習者の語彙知識の要素間の相関関係を表3に示す。中国語話者、韓国語話者ともに、広さテストと深さテストの間に有意な相関があり、深さテストにおける連想の種類の間にも有意な相関が見られた。さらに、相関係数は一貫して韓国語話者が中国語話者に比べて高かった。母語話者については、広さテストと深さテストの間、および、連想の種類の間にも有意な相関は見られなかった。

表3

L2学習者の語彙知識の要素間の相関関係

中国語話者 (N = 50)				
	広さ-全体	深さ-全体	深さ-Par	深さ-Syn
広さ-全体	1			
深さ-全体	.51***	1		
深さ-Par	.34*	.79**	1	
深さ-Syn	.60****	.76****	.36**	1
深さ-Ana	.19	.73****	.41**	.29*
韓国語話者 (N = 20)				
	広さ-全体	深さ-全体	深さ-Par	深さ-Syn
広さ-全体	1			
深さ-全体	.90****	1		
深さ-Par	.91****	.93****	1	
深さ-Syn	.77****	.93****	.79****	1
深さ-Ana	.86****	.96****	.85****	.85****

\* p &lt; .05, \*\* p &lt; .001, \*\*\* p &lt; .001, \*\*\*\* p &lt; .0001

注 Par = Paradigmatic, Syn = Syntagmatic, Ana = Analytic

#### 4.2.4. テクスト理解

再生テストの正答率と要約完成テストの得点の平均と標準偏差を表4に示す。再生については、重回帰分析 (full model:  $F [5,104] = 32.65$ ,  $MSE = 97.57$ ,  $p = .0001$ ) を行った結果、グループ ( $F [2,109] = 70.92$ ,  $d = 11.22$ )、グループとテキストの交互作用 ( $F [2,109] = 7.82$ ,  $d = 3.73$ ) に有意な効果がみられたが、テキストの主効果は有意ではなかった ( $F [1,109] = 2.91$ ,  $d = 1.61$ )。中国語話者と韓国語話者の正答率は同程度で (mean difference = .49,  $\pm 95\%CI = -5.69 \sim 4.72$ )、両L2グループともに母語話者に比べて低かった (中国語話者: mean difference = 23.56,  $\pm 95\%CI = 19.39 \sim 27.73$ ; 韓国語話者: mean difference = 23.07,  $\pm 95\%CI = 17.68 \sim 28.47$ )。中国語話者にとっては旅テキストが文化テキストより易しかった (mean difference = 10.53,  $\pm 95\%CI = 4.98 \sim 16.07$ ) が、韓国語話者 (mean difference = 5.87,  $\pm 95\%CI = -2.94 \sim 14.68$ ) と母語話者 (mean difference = 5.96,  $\pm 95\%CI = -.27 \sim 12.18$ ) にとってはそうでなかった。

要約完成テストの結果は、再生テストの結果と同様のパターンであった。中国語話者と韓国語話者の得点は同程度で (mean difference = 4.69,  $\pm 95\%CI = -7.27 \sim 16.65$ )、両グループはともに母語話者に比べて低かった (中国語話者: mean difference = 19.99,  $\pm 95\%CI = 10.41 \sim 29.57$ ; 韓国語話者: mean difference = 15.30,  $\pm 95\%CI = 2.91 \sim 27.70$ )。テキストによる効果は、中国語話者に見られた (mean difference = 18.65,  $\pm 95\%CI = 5.91 \sim 31.40$ ) が、韓国語話者 (mean difference = 9.70,  $\pm 95\%CI = -10.54 \sim 29.94$ ) と母語話者 (mean difference = 6.36,  $\pm 95\%CI = -7.95 \sim 20.68$ ) には見られなかった。

各グループについて再生と要約完成の間に信頼できる中程度からやや高い正の相関が見られた (中国語話者:  $r = .62$ ; 韓国語話者:  $r = .69$ ; 母語話者:  $r = .30$ )。以下に述べる語彙知識とテキスト理解の関係についての分析では総合得点をテキスト理解の指標として用いる。

表4

グループごとの再生テストと要約完成テストの得点の平均と標準偏差

		N	再生 (%)		要約完成 (Max = 5)	
			M	SD	M	SD
NS	旅テキスト	18	34.8	10.9	4.3	.8
	文化テキスト	22	40.8	10.2	4.5	.8
	全体	40	38.1	10.8	4.3	.8
NNS						
中国語話者	旅テキスト	26	19.5	11.6	3.8	1.5
	文化テキスト	24	9.0	6.4	2.9	1.1
	全体	50	14.4	10.8	3.4	1.4
韓国語話者	旅テキスト	11	17.6	9.1	3.8	1.3
	文化テキスト	9	11.8	10.2	3.3	1.2
	全体	20	15.0	9.8	3.6	1.2

注 NS = 母語話者; NNS = 非母語話者

#### 4.2.5. 語彙知識とテキスト理解の関係

語彙知識の要素とテキスト理解の相関関係を表5に示す。中国語話者と韓国語話者については共に、テキスト理解は語彙知識の広さ、深さの両側面と信頼できる中程度の正の相関があった。しかし、語連想の種類についての相関は、

グループ間で異なるパターンが見られた。中国語話者の場合は、テキスト理解とsyntagmaticとの間に信頼できる中程度の相関が見られたが、他の連想の種類には見られなかった。韓国語話者の場合は、テキスト理解とparadigmaticおよびanalyticとの間に中程度の相関が見られたが、syntagmaticには見られなかった。母語話者については、テキスト理解と語彙知識との間に信頼できる相関は見られなかった。

表5

グループごとの語彙知識の要素とテキスト理解の相関関係

	N	語彙知識の要素				
		広さ-全体	深さ-全体	深さ-Par	深さ-Syn	深さ-Ana
NS	40	.16	.21	.13	.12	.21
NNS						
中国語話者	50	.45***	.33*	.12	.49***	.13
韓国語話者	20	.60**	.49*	.50*	.43	.46*

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

注 NS = 母語話者; NNS = 非母語話者

Par = Paradigmatic, Syn = Syntagmatic, Ana = Analytic

語彙知識のいろいろな要素がテキスト理解にどの程度貢献しているかを調べるために、各グループについて語彙知識の要素を予測変数として系統的に選択した多重およびステップワイズ回帰分析を行った。第1モデルでは、2つの語彙知識得点、すなわち、広さと深さを予測変数として用いた。第2モデルでは、4つの語彙知識得点、すなわち、広さ、paradigmatic、syntagmatic、analyticを予測変数として用いた。第1モデルを使った結果、中国語話者については、広さがテキスト理解の変動の20%を説明する要因として選択されたが、深さは選択されなかった。韓国語話者については、広さがテキスト理解の変動の36%を説明する要因として選択されたが、深さは選択されなかった。母語話者については、広さではなく、深さがテキスト理解の変動の4%を説明する要因として選択された。

第2モデルを使った結果では、中国語話者については、syntagmaticが初めに選択されテキスト理解の変動の24%を説明し、続いて広さが選択された(4%の増大)。韓国語話者については、広さのみが選択されテキスト理解の変動の

36%を説明していた。母語話者については、analyticのみが選択され4%を説明していた。

## 5. 考察

### 5.1. L2語彙知識の特徴

語彙知識の分析結果から、L2日本語の語彙知識の広さと深さについて、中国語話者と韓国語話者の間にはいくつかの異なる特徴があることが示唆された。まず、中国語話者は、両語彙テストにおいて、韓国語話者に比べて正答率が高かったことから、より高度な語彙知識を有していると考えられる。本研究で使用した語彙テストは、受容テストで、語とその定義を結びつけたり（広さテスト）3つの意味的に関連する語を選択したり（深さテスト）するために、表記された語についての知識が要求されるテストである。漢字には全て振り仮名をつけて読みを表記したが、中国語話者は漢字で表記された語彙の再認について有利であったかもしれない。同様のテストを音声媒介で行った場合はテストでのパフォーマンスもかなり異なるのではないだろうか。

第二に、深さテストの結果から、語連想の種類に関する知識の要素の強度が中国語話者と韓国語話者とでは異なっていることが示唆された。中国語話者は、paradigmaticの知識がもっとも強く、syntagmaticの知識が特に弱い。それに対して、韓国語話者はsyntagmaticの知識は他の種類の知識と同程度の強さがある。深さテストにおける彼らのパフォーマンスは現在の語彙知識を反映していると考えられるが、現在の語彙知識の特徴には母語背景の影響による痕跡が現れている可能性も考えられる。より具体的には、中国語は孤立語でSVO型の語順を持つ言語であり、膠着語でSOV型の語順を持つ日本語とは大きく異なる。従って、中国語話者は日本語の文を処理する際に、母語に基づく統語知識が有効ではなく、語と語の間のsyntagmaticの関係についての知識の習得が遅れるのではないだろうか。それに対して、韓国語は膠着語でSOV型の語順を持つ、統語的には日本語に近い言語である。従って、韓国語話者は日本語の文を処理する際に、母語に基づく統語知識が有効で、文の語と語の間のsyntagmaticの関係を学ぶ機会が増えるのではないだろうか。L2語彙知識の習得における母語の影響についてのこのような可能性は、今後の研究で検証される必要がある。

L2学習者と母語話者の比較結果から、L2学習者の語彙知識は母語話者の語



彙知識と比べて量的にも質的にも異なる特徴があることが示唆された。中国語話者と韓国語話者は共に、両テストにおいて母語話者に比べて有意に低い正答率であった。加えて、語の頻度の一般的な効果は両L2学習者グループで見られたのに対して、母語話者については頻度の効果は広さテストで見られず、深さテストでは逆方向で見られた。さらに、L2学習者については、語彙知識の広さと深さの間、および、異なる種類の連想の間にそれぞれ信頼できる相関が見られたが、母語話者については語彙知識の要素間で信頼できる相関は見られなかった。このような結果は概ね他言語を対象とした先行研究の結果を支持するものである（Greidanus et al., 2005; Greidanus & Niehuis, 2001; Schmitt & Meara, 1997）。よって、語彙テストで用いた対象語は、その頻度レベルがL2学習者の語彙知識の熟達度を識別するのに適当なものであり、母語話者にとっては既知語であることが確認された。

母語話者について逆方向の頻度の効果が現れたのは予想外であった。この結果は、母語話者の長期記憶の中のある語の表象が持つ（他の語の表象との）結びつきが当該言語における語彙の中でのその語の「立ち位置」を反映している可能性を示唆している。より詳しく述べると、母語話者の洗練された語彙知識システムの中で、頻度の高い語は多くの他の語と比較的弱い連想で繋がっているが、頻度の低い語は限られた数の他の語とよりはっきりとした連想で繋がっている。受容テストで応答する際に、頻度の低い語の連想語は、高頻度語の場合と比べてより容易に特定化され個人間で一致しやすいのに対し、高頻度語は異なる多数の連想語の間で競争が起きやすい。対象語の使用頻度が低く母語話者もほとんど知らない語であれば、いくつかの先行研究（Greidanus et al., 2005）で示されているように、母語話者もL2学習者のようなテストパフォーマンスを見せたのではないだろうか。

深さテスト結果について観察されたもう一つの興味深いことは、異なる種類の連想の知識の強度について、L2学習者と母語話者との間で異なるパターンが見られたことである。L2学習者についてはanalyticの知識はparadigmaticやsyntagmaticの連想の知識と比べて強度が高いか同程度であったのに対し、母語話者についてはanalyticの知識は他の種類の連想の知識と比べて明らかに強度が低い。このような違いは、異なる種類の語連想の知識がどのように習得され使用されているかということと関係していると考えられる。Paradigmatic

の連想は同義語やカテゴリーと事例の関係についての知識に関わるもので、syntagmaticの連想は文の中で語と語の共起についての知識に関わるものである。このような種類の語連想の知識は、人がコミュニケーションする際、談話やテキストを処理する最中に自然に使用されるものである。それに対して、analyticの連想は、辞書の定義の記述のように、要素に分析された語の意味についての知識に関わるものである。このような知識は明示的で分析された語知識に関わるものである。教室内学習を基盤とする成人のL2学習者は、教師やコース資料（教科書や辞書など）によって提供される説明や定義を通して、語の明示的な分析に晒される傾向があり、analyticの連想の知識が比較的強くなるのではないだろうか。しかし、母語話者（言語学者以外）の場合は、通常の状況下で辞書の定義の記述のように明示的に語を分析することはなく、analyticの連想の知識は、他の種類の連想と比べて発達しにくいアクセスしにくいのではないだろうか。このような考えについては今後の研究で検証する必要がある。

## 5.2. 語彙知識とテキスト理解の関係

再生と要約完成のテスト結果から、中国語話者と韓国語話者はテキスト理解については同程度であることが示唆された。この結果は、2つの語彙テストで韓国語話者が中国語話者と比べて正答率が有意に低いことを考えると、意外であった。

語彙知識とテキスト理解の関係については、分析結果から、中国語話者と韓国語話者の間に共通点と相違点があることが示唆された。どちらのL2学習者グループにおいても語彙知識の広さと深さは共にテキスト理解と有意な相関関係があることが分かったが、これはEFL学習者を対象とした先行研究（Qian, 1999, 2002）の結果を支持するものである。しかし、語彙知識の深さがテキスト理解とどのように関係しているかについては、中国語話者と韓国語話者との間で異なるようである。中国語話者の場合は、syntagmaticの連想はテキスト理解との間に有意な中程度の相関があったが、paradigmaticおよびanalyticの連想はテキスト理解との間に有意な相関はなかった。それに対して、韓国語話者の場合は、3種類の連想すべてがテキスト理解との間に中程度の相関があったが、syntagmaticの連想は有意ではなかった。また、回帰分析の結

果からも、中国語話者のテキスト理解についてはsyntagmaticの連想が特に重要であるが、韓国語話者の場合はそうではないということが示唆された。

このような中国語話者と韓国語話者の間の相違は、少なくとも部分的には、以下のように説明できるのではないだろうか。syntagmaticの連想の知識とこの知識を使用する能力は、テキストに描写記述されている出来事や命題の理解の際に必要な連続的处理に関係するものである。テキスト処理においてsyntagmaticの連想は、記憶の中にテキストの（命題）テキストベースや状況モデルを構築する際に仲介的な役割を担うものであると考えられる。おそらく、中国語話者の場合は、日本語のテキストを読む際に、漢字で表記された内容語から意味を抽出するために書記形態素処理に依存し、統語処理にあまり積極的に従事しない傾向がある。テキスト理解におけるこのような書記形態素処理と形態意味処理への過度な依存が、その結果として、語彙の深さテストの結果が示唆するように、syntagmaticの連想の知識の発達度が低い語彙知識に繋がっている。よって、より強度なsyntagmaticの連想の知識を持っている中国語話者は、弱いsyntagmaticの連想の知識を持つ者と比べて、テキスト理解に成功する可能性が高い。加えて、中国語話者は、内容語の再認を重要視し、語と語の関係に関する情報を充分に取り込まずに個々の語についての情報をテキスト表象に取り入れる可能性がある。

それに対して、韓国語話者の場合は、L1（韓国語）が日本語と統語的に似通っており、日本語で書かれたテキストの中の文を読む際に語順や文構造の情報を積極的に使用する。このような積極的な統語処理が、L2日本語のsyntagmaticの連想の知識の習得に影響し、後にその知識がテキスト処理において使用されると考えられる。さらに、韓国語話者は、より整合性のあるテキスト表象を構築するために、L1やL2を基盤とする統語に関する知識と能力など他の種類の知識を使用して、限られた語彙知識を補っているのではないだろうか。その結果、韓国語話者は、より高度な語彙知識を持つ中国語話者と同じくらい、テキスト理解に成功した。別の言い方をすれば、L1を基盤とする要因がL2テキスト理解に及ぼす影響は2つの経路で現れる可能性である。1つはテキストの処理と理解に対する直接的な影響で、もう1つは語彙知識の習得と使用を通して起こる間接的な影響である。これらの可能性は、このようなL1背景（中国語と韓国語）を持つL2日本語学習者においてよく観察されることと一致してお

り、今後の研究において直接的に研究されなければならない。

最後に、L2学習者にとっての語彙知識とテキスト理解の関係は、母語話者の場合と比べて、量的にも質的にも異なる部分があるようである。本研究では、L2学習者のテキスト理解は、(語彙知識とテキスト理解の間の高い相関に示されているように、)母語話者の場合と比べて、語彙知識への依存度が高いことが明らかにされた。さらに、語彙知識の広さと深さ、および、語連想の種類についての相対的な重要度のパターンがL2学習者と母語話者の間で異なることが明らかにされた。このような解釈は、語彙知識とテキスト理解についての複数の測定から得た結果をもとにおこなったものであるが、直接的な観察によるものではない。語彙知識とテキスト理解の関係についての直接的な研究は、言語知識の異なる要素がどのようにL2テキスト理解に貢献するかを解明する一助となるであろう (Bernhardt, 2005; Koda, 2005)。そのような研究は、一般にテキスト理解の複雑さの問題、特にL2テキスト理解に関わる補償と呼ばれる問題を解明する手助けとなるであろう。

## 6. 結論

本研究で得られた主な結果を以下にまとめる。第一に、語彙知識の要素の相対的な強度は、中国語を母語とするL2学習者と韓国語を母語とするL2学習者とで異なり、L2学習者は日本語母語話者と異なる。第二に、韓国語話者は、中国語話者と比べて語彙知識の発達が遅れているが、中国語話者と同程度のテキスト理解ができるようである。L2テキスト理解には語彙知識の広さ(すなわち、基本的な語の意味に関する知識)と語彙知識の深さ(すなわち、語連想に関する知識)の両方が重要であるが、syntagmaticの連想に関する知識は中国語話者にとって特に重要だが、韓国語話者にとってはそうではない。これらの結果は、L2語彙知識とそのテキスト理解との関係に、L1に関する要因が複雑な形で影響していることを示唆する。

もちろん、本研究にも限界があり、本研究の結果から新たな質問も派生する。第一に、研究で使用する特定の測定方法は語彙知識やテキスト理解として何が観察できるかを制限する。本研究では、語彙知識を測定するために、特定の言語テスト(すなわち、日本語能力試験)のために用意された語彙リストから選んだ語を対象語とする受容テストを使用した。その語彙リストの特徴から、本

研究での対象語の頻度レベルは、対象語を辞書やその他の資料といったより大きなデータベースから選んだ場合と比べて、L2学習者の馴染み度とより強い相関があった可能性がある。また、テストで用いた特定の手法によって、自然なコミュニケーションの中で観察されるものとは異なる行動を引き出した可能性もある。第二に、本研究は相関関係を調べたものであって、要因間の因果関係を確立するものではない。L2学習者の目標言語経験が増すにつれて語彙知識がどのように発達し変化するかを理解するためには、縦断的な研究が必要である。さらに、語彙知識の異なる要素がテキスト理解にどのように貢献するかを調べるためには直接的な研究が必要である。例えば、テキスト処理における意味カテゴリーと精緻化についての研究、作動記憶とsyntagmaticの連想の関係についての研究は有益であろう。

最後に、本研究の結果をもとに教育的示唆を導き出すことができる。まず、L2学習者には意味のある文脈の中で語彙を処理する機会を数多く与え、テキストや談話の中で語が他の語とどのように関係し合うかについての知識を発達させ強化できるようにしなければならない。ひとつの有効なアクティビティは、繋がった談話の中で使われている語彙を、学習者が談話の内容を理解した後で、明示的に分析するというアクティビティである。語彙がどのように順序づけられて、あるいは、配列されて、命題やより大きな意味の単位を形成しているかを分析することによって、L2学習者は、テキストの中の語と語の繋がりや自己のテキスト心表象に対してより敏感になることができる。第二に、複数のL2学習者が読解において（少なくとも結果から判断すると）同様のパフォーマンスをしているように見えても、実際は言語知識や処理能力における異なる長所と短所を持っており、そこにL1背景による影響が部分的に関わっているかもしれない。このようなL2学習者間の相違点はカリキュラムの設計と実行、および、評価において注意深く考慮されなければならない。

## 注

本稿はHoriba（2012）. Word knowledge and its relation to text comprehension: A comparative study of Chinese- and Korean-speaking L2 learners and L1 speakers of Japanese, *Modern Language Journal*, 96, 108-121を和訳したものである。翻訳に協力して下さった田所直子さんに感謝申し上げます。

- 1) 旧日本語能力試験には4つのレベルがあり、出題基準 (Japan Foundation, 2002) によると1級で約2,000の漢字と約10,000の語、2級で約1,000の漢字と約6,000の語を知っているとされている。語のレベル分けは、複数のコーパスにおける頻度と専門家の判断に基づいて行われている。本研究のL2協力者はほとんどが2級に合格しており、これから1級を受けようとしていた。本研究で使用した高頻度語は2級レベルから、低頻度語は1級レベルから選択されたものである。本研究の対象語48個は別の調査研究の結果 (Horiba, Kobayashi, Matsumoto, & Suzuki, 2004) を基に選定された。その研究では、中国語母語89名、英語母語49名のL2学習者を対象に、1級と2級の語彙リストよりランダムに取り出し、領域を考慮して選出した200語の内容語 (名詞124個、動詞35個、形容詞20個、副詞9個、その他15個) について、伝統的な語彙テスト (多肢選択式による文中空所補充) を行った。その結果から正答率60%以下の語を同定し、本研究の対象語として選定した。選出した語と全体の200語との間には、信頼できる中程度の相関 ( $r = .61, p < .002$ ) があった。
- 2) 品詞による推測を減らすために元々、各対象語について8つの選択肢 (正答3、錯乱肢5) が用意されていた。パイロット調査の結果、いくつかの錯乱肢が母語話者に選択されることが判明したため、これらの錯乱肢は削除することとした。よって、本調査で用いたテストでは各対象語について4つの錯乱肢が含まれていた。
- 3) 先行研究によると、L2学習者はL2よりもL1でより多くのテキスト内容を再生する傾向があることが示されている (Lee, 1986)。
- 4) 要約のテンプレートは、日本語を母語とする経験豊富な日本語教師5名のチームによって作成された。手順としては、まず各自が個別に要約を作り、それから全ての要約に含まれる内容を確認した。(テストで) 要約を正確に完成するためには読み手が別々の文にあるアイデアを統合しなければならないように、注意を払った。

## 参考文献

- Aitchison, J. (1994). *Words in the mind: An introduction to the mental lexicon*. (2nd ed.). Oxford, England: Blackwell.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- Bernhardt, E. B. (2005). Progress and procrastination in second-language reading. In M. McGroarty (Ed.), *Annual Review of Applied Linguistics* (pp. 133-150). Cambridge: Cambridge University Press.
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationship between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypothesis.

*Applied Linguistics*, 16, 15–34.

- Greidanus, T., Beks, B., & Wakely, R. (2005). Testing the development of French word knowledge by advanced Dutch- and English-speaking learners and native speakers. *Modern Language Journal*, 89, 221–233.
- Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distracters and types of associations. *Modern Language Journal*, 85, 567– 577.
- Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10, 221–239.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303–317.
- Henriksen, B. (2008). Declarative lexical knowledge. In D. Albrechtsen, K. Haastrup, & B. Henriksen, (Eds.), *Vocabulary and writing in a first and second language: Processes and development* (pp. 22–66). New York: Macmillan.
- Horiba, Y., Kobayashi, H., Matsumoto, J., & Suzuki, H. (2004). Acquisition of vocabulary and grammar in Japanese as a second language: A comparison between Chinese-speaking learners and Englishspeaking learners. In Y. Horiba (Ed.), *Text understanding and learning* (pp. 1–14). Research Report. Grant-in-aid for scientific research (B). Japan Society for the Promotion of Science.
- Huckin, T., Hayes, M., & Coady, J. (1993). *Second language reading and vocabulary learning*. Norwood, NJ: Ablex.
- Japan Foundation (2002). *Nihongo-noryoku-shiken shutsudai kijun* [Standards for questions in the Japanese language proficiency test]. Tokyo: Bonjinsha.
- Koda, K. (1993). Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *Modern Language Journal*, 77, 490–500.
- Koda, K. (1996). L2 word recognition research: A critical review. *Modern Language Journal*, 80, 450–460.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York: Cambridge University Press.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? H. Béjoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 125– 132). London: Macmillan.
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, 365–391.
- Lee, J. F. (1986). On the use of recall tasks to measure L2 reading comprehension. *Studies in*

*Second Language Acquisition*, 8, 201–212.

- Lee, J. W., & Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*, 31, 713–739.
- Matsunaga, S. (1999). The role of kanji knowledge transfer in acquisition of Japanese as a foreign language. *Japanese-Language Education around the Globe*, 9, 87–100.
- Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & A. P. R. Howett (Eds.), *Interlanguage* (pp. 225–235). Edinburgh, Scotland: University of Edinburgh Press.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Competence and performance in language learning* (pp. 35–53). New York: Cambridge University Press.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304–330.
- Nagy, W. E., & Scott, J. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Mahwah, NJ: Erlbaum.
- Namei, S. (2004). Bilingual lexical development: A Persian–Swedish word association study. *International Journal of Applied Linguistics*, 14, 363–388. doi: 10.1111/j.1473-4192.2004.00070.x
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *Modern Language Journal*, 90, 387–401.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nissen, H. B., & Henriksen, B. (2006). Word class influence on word association test results. *International Journal of Applied Linguistics*, 16, 389–408.
- Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282–307.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 53, 513–536.
- Randall, M. (2007). *Memory, psychology and second language learning*. Amsterdam: Benjamins.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355–371.
- Read, J. (2000). *Assessing vocabulary*. New York: Cambridge University Press.



- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 209–227). Amsterdam: Benjamins.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17–36.
- Segalowitz, N., Poulson, C., & Komeda, M. (1991). Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. *AILA Review*, 8, 15–30.
- Shibata, T., et al. (Eds.). (1995). *Shinmeikai kokugo-jiten* [New lucid Japanese dictionary]. (4th ed.). Tokyo: Sanseido.
- Shibata, T., et al. (Eds.). (2002). *Ruigo-daijiten* [Synonym dictionary]. Tokyo: Kodansha.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36, 139–152.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.
- Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Erlbaum.
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217–234.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Journal Review*, 53, 13–40.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition*, 23, 41–69.
- Yamaguchi, T. (Ed.) (2003). *Nihongo daihesaurus: Ruigokensaku daijiten* [Synonym dictionary]. Tokyo: Taishukan.
- Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research*, 23, 123–153.
- Zareva, A., Schwanenflugel, P., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition*, 27, 567–595.

付表A、B、CはHoriba (2012) を参照のこと。