

# 英語母語話者と日本人英語学習者の 心内辞書における語彙項目間類似度の比較

折田 充・小林 景・村里泰昭・Richard S. Lavin・  
吉井 誠・相澤一美・神本忠光

## 1. はじめに

本研究は、英語母語話者と日本人英語学習者の心内辞書について探るものである。心内辞書 (mental lexicons) は、言語の形態、意味、機能、また付随する体験的知識・記憶など語彙に関する多様な情報が格納されている脳内機構であり、人が言語の理解や産出のために使う「辞書」である。折田・小林・村里・神本・吉井・Lavin (2013) は、英語母語話者群と語彙サイズの異なる2群の日本人英語学習者における心内辞書内の英単語群の意味的クラスタリング (semantic clustering: 心内辞書内において意味的に関連する単語群がグループ化されたもの、またその状態) 構造には計量的及び質的差異が存在することを確認した。つまり、心内辞書内の様々な構造化形成の中でも、語彙項目 (lexical items) 間の意味的関連性に基づき形成されている英単語クラスターの構造は、母語話者と日本人英語学習者の間で異なることを明らかにした。加えて、日本人英語学習者に関しては、語彙サイズの違いによってクラスター構造に違いがあることが判明した。本研究は、この折田他において確認された群間の相違が、心内辞書内語彙項目の最も基本的な結び付きの指標である語彙項目間類似度 (degrees of similarity between lexical items) に関しても同様に存在するのかを解明することを目指す。なお、本稿では、語彙項目間類似度として「心内辞書内の語彙項目ペア間の距離。特に、心理言語学実験の一つである単語仕分け課題 (word sorting task) 結果に基づくデンドログラムから計算される各単語ペア間の距離」を用いる。

## 2. 先行研究

心内辞書内の意味的クラスタリング構造は、母語話者 (NS) であるか第二言語話者 (NNS) であるかによって差異が存在するのみならず、NNS の中では目標言語の習熟度の違い、さらには辞書内の個々の語彙項目の特性によって異なる。また、辞書内の語彙項目は、形態や機能よりも意味を基盤に結びついているという側面が強い。Fitzpatrick (2006) は、単語連想テスト (WAT: word association test) を使った心理言語学実験結果から、NS であれ NNS であれ、実験に使った刺激語と反応語の間の形態上の類似性に起因する連想よりも意味的な関係性に基づく連想が圧倒的に多かったと報告している。Schmitt (2000, 2010) は、第二言語語彙の発達において全ての語彙が同じ過程をたどるわけではないと指摘している。初級レベルから上級レベルの4群 (中学2年生、工業高専2年生、非英語専攻の大学2年生、英語専攻の大学3・4年生) の日本人英語学習者を被験者とし、WAT における反応語と英語習熟度の関係を分析した Orita (2002) はこれを支持し、被験者群の習熟度との関係からみて、刺激語によって表出する反応語のタイプの変化は異なると報告している。Wolter (2001) の

WAT 結果でもまた、心内辞書内の個々の単語の発達 (developmental shifts) は、他の単語とは独立して進展するという特徴があることを示し、Schmitt の主張を裏付けている。

この語彙項目間の結び付きにおける意味の優勢や心内辞書における語彙項目発達の個別性という傾向は、心内辞書内で意味の場 (semantic fields) を形成する語彙項目クラスターについても存在する。つまり、意味的クラスタリング構造の違いは、母語であるか第二言語であるか、また後者については第二言語の習熟度の違いによって存在すると同時に、クラスター内の単語群の特性によって被験者群間で類似するものとそうでないものが併存している。単語仕分け課題を使った実験結果から折田・小林 (2011a) は、NS群の意味的クラスタリング構造を基準とすると、心内辞書内の意味的クラスタリング構造のクラスタータイプは4つに分類できると報告している。タイプ I は英語上級話者群では NS群との類似性が高く中級話者群ではむしろ独自の (未発達の) 要素を含む単語群であり、タイプ II は上級話者群では NS群との類似性が高く中級話者群では両群と同様のクラスターを形成しつつも異なる意味特性を持つ語群と結び付く構造であった。さらに、タイプ III は中級話者群と上級話者群の同質性が高く NS群の構造に近似しにくい単語群であり、またタイプ IV では3群間の高い類似性はあるが部分的に構造上の違いを含んでいた (関連して、Orita & Kobayashi (2012), 折田・小林 (2012) を参照のこと)。

そして、NSとNNSの心内辞書内の意味的クラスタリング構造には、計量的違いが存在する。Wilks and Meara (2002) は、WAT の分析結果から NSの方がNNSよりも単語間の結び付きを見出した割合が高かったと報告している。さらに、Wilks and Meara (2007) は、第二言語の習熟度の高いNNSの方が習熟度の低いNNSよりも単語間の結び付きを多く見出したと述べている。折田・小林は、NS群及び英語習熟度の異なる2群のNNS群を被験者にした一連の単語仕分け課題を用いた研究 (折田・小林, 2010, 2011a, 2011b; Orita & Kobayashi, 2010) から、3群の心内辞書間には計量的に有意な差異があることを明らかにした。加えて、NSとNNSの心内辞書内の意味的クラスタリング構造には質的な違いもあることを確認した。Orita and Kobayashi (2009) 及び折田・小林 (2011c) においても、NS群とNNS群間に名詞、動詞、形容詞のいずれの語類についても両群間に質的な違いがあったと報告している。さらに、折田・小林 (2010, 2011b, 2011c) 及び Orita and Kobayashi (2011) は、NS群、日本人上級英語話者群及び中級英語話者群の3群の心内辞書内の意味的クラスタリング構造の間に質的違いがあることを明らかにしている。このように、語彙項目の特性、母語話者と第二言語話者、また後者においては第二言語の習熟度によって心内辞書内の構造には計量的のみならず質的な違いが存在すると指摘できる。

折田他 (2013) は、単語仕分け課題結果から得られた群デンドログラムの計量的解析から、日本人英語学習者の心内辞書内構造における高頻度英語動詞の意味的クラスタリング構造は母語話者とは異なること、日本人英語学習者の間でも語彙サイズの大きさにより違いがあること、また3群の意味的クラスタリング構造は質的にも異なることを明らかにした。さらに、確認された心内辞書内クラスターを英語母語話者のクラスタリング構造を基準として群間で比較するとき、3つに分類された。タイプ A クラスターは語彙サイズの小さいNNS群と語彙サイズの大きいNNS群の同質性が高く、NS群の構造との違いが大きい単語群であり、タイプ B クラスターは高い3群間の類似性があり、NNS群には一部 NS群に近似していない要素を含む単語群から構成されていた。そして、タイプ C クラスターは3群間の違いが著しい単語群であった。これらの結果は、NS群とNNS群の心内辞書内の

様々な側面における計量的及び質的相違の存在を明らかにした Wilks and Meara (2002, 2007) や折田・小林 (2010, 2011a, 2011b, 2011c, 2012), また Orita and Kobayashi (2010) を支持している。では、このような心内辞書の被験者群間の相違が、辞書内の最も基本的な語彙間の結び付きの単位である語彙項目間類似度についても存在するのか。また、群間に相違があるとすれば、それにはどのような特徴があるか。本研究では、心内辞書内の意味的クラスタリング構造に関するこれらの問題の解明に取り組みたい。

### 3. 研究課題

本研究は、前述した先行研究の概観を踏まえて、被験者の語彙サイズの違いを指標に、英語母語話者と日本人英語学習者の心内辞書に関して次の2つの研究課題について明らかにすることを旨とする。

- (a) 心内辞書内の語彙項目間類似度に計量的な差異があるか。
- (b) 心内辞書内の語彙項目間類似度分布に特徴的な傾向が存在するか。

## 4. 方法

### 4.1 被験者

本研究の被験者は3群より構成された。1つ目の群は、英語母語話者群 (NS 群) 30名で熊本市内外に在住の英語教師、あるいは熊本大学に在籍する英国や米国などからの留学生や同大学客員研究員である。NS 群には語彙サイズテストを課さなかったが、教養ある英語母語話者が持つとされる20,000語の語彙サイズ (Nation, 2001) を保持していると推定された<sup>1)</sup>。日本人英語学習者 (熊本県内3大学の大学1~3年生) が2つ目及び3つ目の群 (NNS 群) を構成し、The Vocabulary Size Test (Nation & Beglar, 2007) 選択肢日本語版の結果に基づき、大きな語彙サイズを持つ群 (LARGE 群: 平均6273.3語,  $SD = 258.6$ ,  $Max = 6800$ ,  $Min = 6000$ ) 30名と小さな語彙サイズを持つ群 (SMALL 群: 平均4683.3語,  $SD = 155.5$ ,  $Max = 4900$ ,  $Min = 4200$ ) 30名から成った。なお、LARGE 群と SMALL 群の語彙サイズは有意に異なった ( $t = 28.86$ ,  $df = 58$ ,  $p < .01$ ,  $r = .97$ .)。

### 4.2 単語仕分け課題

主データ収集には単語仕分け課題を採用した。まず、Stevenson (1883) の *Treasure Island* 中の “Chapter 32. The Treasure-hunt—The Voice Among the Trees” から、『大学英語教育学会基本語リスト JACET List of 8000 Basic Words』(JACET8000) (大学英語教育学会基本語改訂委員会編, 2003年) のレベル1 (順位1-1000位) に該当する高頻度の動詞を全て抽出した上で、それらの中から無作為に50語を選んだ<sup>2)</sup>。実験に用いた動詞は次の通りである:

*add, ask, beat, begin, believe, break, bring, come, cry, describe, die, face, fall, feel, fight, find, get, go, have, hear, help, hold, increase, keep, leave, listen, look, make, mind, remain, reply, rest, return, rise, run, say, see, show, sing, sit, speak, stare, start, stop, strike, struggle, take, tell, think, walk.*

NS 群には、次の手順で実験を行った。各単語を2.4 cm × 4.4 cm のカードに印刷し、順番をランダ

ム化してゴム輪でくくり封筒に入れた。被験者は封筒の中からカードを取り出し、単語仕分け課題に取り組んだ。被験者の持つ心内辞書内の意味的クラスタリング構造を解明するという本研究の目的から、課題は、与えられた50語を被験者が考える動詞の意味で関連する語群に思う通りにグループ分けする自由仕分け課題とした。実験は被験者一人ずつに実施した。2つのNNS群は、オンラインストレージソフトDropbox (Dropbox, Inc., 2011) にアップロードされた単語仕分け課題ファイルを各自パソコンにダウンロードして英語の授業の中で取り組んだ(課題の考え方は、NS群の場合と同じ。50枚の「単語カード」をドラッグ&ドロップの要領で意味の上で関連しているグループに分けるタスク) (APPENDIX 参照)<sup>3)</sup>。

### 4.3 データ解析

#### 4.3.1 ハミング距離, デンドログラム距離及び群平均法

単語仕分け課題で収集したデータ(頻度データ)の集計(マトリックス表)をもとにした距離行列(ハミング距離)から、デンドログラム距離行列を計算した。ハミング距離の定義, またデンドログラム距離及び群平均法の考え方を以下に示す。

#### ハミング距離

$$d(i, j) = 1 - n/N$$

ただし,  $N$ 人の被験者のうち $n$ 人が単語  $W_i$  と  $W_j$  を同じグループに分類したとする。

#### デンドログラム距離

全単語ペアについてのハミング距離  $d(i, j)$  をもとに, 群平均法を用いて定義される距離  $d_T(i, j)$

#### 群平均法 (単語数を $M$ とする)

$$\text{初期値: } d(\{i\}, \{j\}) = d(i, j) \text{ for } i, j = 1, \dots, M$$

$$G = \{\{1\}, \dots, \{M\}\}$$

ただし,  $\{\cdot\}$  は集合を表し,  $\{i\}$  は単語  $i$  のみの集合,  $G$  は単語を一つずつ各集合(グループ)に分けるグループ分けに対応する。

STEP 1:  $d(I, J)$  が最も小さいペア  $I, J \in G$  のうちの一つを選び,  $G$  から  $I$  と  $J$  を除き, 代りに  $I \cup J$  を入れて,  $G$  を更新する。

STEP 2:  $I, J$  以外の  $K \in G$  について,

$$d(I \cup J, K) = \frac{n_I}{n_I + n_J} d(I, K) + \frac{n_J}{n_I + n_J} d(J, K)$$

を計算する。ただし,  $n_I, n_J$  はそれぞれ  $I, J$  のサイズである。

STEP 3: 各  $i \in I \cup J$  と  $j \in K$  について,  $d_T(i, j) = d(I \cup J, K)$  と定義する。

STEP 4:  $G$  が全体集合  $\{1, \dots, M\}$  のみになるまで STEP 1 ~ 3 を繰り返す。

以上が完了すると, 全ての  $i, j$  の組について群平均法によるデンドログラム距離  $d_T(i, j)$  が定義できている。

### 4.3.2 語彙項目間類似度の集計

各被験者群について、得られたデンドログラム距離を集計したマトリックス表の全データ（語彙項目間類似度：0.0～1.0の値。値が小さいほど当該の語彙項目間の類似度が高い。各群に1,225個の語彙項目間類似度あり）を0.1間隔ごとに、それぞれの間隔に入る頻度を集計した。さらに、得られたデータの全体構造の特徴を明確にするために、4つの語彙項目間類似度(T)区分に基づき頻度を再集計した。4つの区分は、「極めて高い ( $T < 0.2$ )」、「高い ( $0.2 \leq T < 0.5$ )」、「中程度～低い ( $0.5 \leq T < 0.9$ )」、「極めて低い ( $0.9 \leq T$ )」とした<sup>4)</sup>。

この頻度データについて、2つのデンドログラム距離行列の限定された語彙項目間類似度（上記の語彙項目間類似度の範囲0.0～1.0を4領域に分割したもの）を持つセルの頻度の差を用いて、それぞれの群間の類似度に差はないという帰無仮説のもとに並べ替え検定を行った。並べ替え検定は、実験で得た3群のデータの組み合わせペア数である6回（6通りのペアワイズ比較）実施した。実施した並べ替え検定の繰り返し回数とその信頼性の確認については、「5. 結果」で述べる。なお、デンドログラムの構成法については齋藤・宿久(2006)、並べ替え検定については竹村(1991)、またデンドログラムを距離行列として扱う統計的手法については小林・折田(2009, 2012, 2013)を参照のこと。また、並べ替え検定には、MATLAB (Matrix Laboratory) Version 7.5.0.342 (R2007b) (The MathWorks, Inc., 2007)を用いた。

## 5. 結果

### 5.1 心内辞書内の語彙項目間類似度の計量的差異

研究課題 (a) 「心内辞書内の語彙項目間類似度に計量的な差異があるか」に関して、デンドログラム距離行列における語彙項目間類似度の頻度及びパーセントを表1に示す。

表1. デンドログラム距離行列における語彙項目間類似度の頻度及び% ( $n = 1,225$ )

類似度 (T)	$T < 0.2$	$0.2 \leq T < 0.5$	$0.5 \leq T < 0.9$	$0.9 \leq T$
NS	12 1.0%	52 4.2%	392 32.0%	769 62.8%
LARGE	9 0.7%	58 4.7%	160 13.1%	998 81.5%
SMALL	7 0.6%	21 1.7%	153 12.5%	1044 85.2%

(注) 類似度 (T) = 語彙項目間類似度（値が小さいほど類似度は高く、大きいほど類似度は低い）。NS = 英語母語話者群；LARGE = 大きな語彙サイズを持つ日本人英語学習者群；SMALL = 小さな語彙サイズを持つ日本人英語学習者群。%は総語彙項目間類似度頻度1,225に対する当該頻度の百分率。

表1にまとめた頻度分布に対して並べ替え検定を行った。なお、それぞれのペアワイズ比較において行った並べ替え検定の繰り返し回数は5,000回であった。並べ替え検定の  $p$  値の信頼区間を計算し、検定の繰り返し回数が十分か否かを確認したところ、いずれのペアワイズ比較においても十分であることがわかった。表2に解析結果をまとめた。

表2. 並べ替え検定結果：各ペアワイズ比較の帰無仮説に対する  $p$  値

帰無仮説	$T < 0.2$	$0.2 \leq T < 0.5$	$0.5 \leq T < 0.9$	$0.9 \leq T$
$T\_SMALL \leq T\_LARGE$	0.286	0.476	0.969	0.043*
$T\_SMALL \leq T\_NS$	0.381	0.970	0.997	0.000*
$T\_LARGE \leq T\_NS$	0.691	0.993	0.960	0.031*
$T\_SMALL \geq T\_LARGE$	0.714	0.524	0.031*	0.957
$T\_SMALL \geq T\_NS$	0.619	0.030*	0.003*	1.000
$T\_LARGE \geq T\_NS$	0.309	0.007*	0.040*	0.969

(注)  $T$  = 語彙項目間類似度。  $T\_SMALL$  = SMALL群のデンドログラム距離行列における語彙項目間類似度の頻度；  $T\_LARGE$  = LARGE群のデンドログラム距離行列における語彙項目間類似度の頻度；  $T\_NS$  = NS群のデンドログラム距離行列における語彙項目間類似度の頻度。

\* $p < 0.05$ .

表2から、6つの並べ替え検定のいずれについても5%有意水準で有意差のあるセルの存在が確認され、すべてのペアワイズ比較における帰無仮説が棄却された。つまり、3群間のデンドログラム距離行列における語彙項目間類似度の頻度分布には有意な偏りがあることが確認され、心内辞書内の語彙項目間類似度に計量的な差異があることが明らかとなった。

## 5.2 心内辞書内の語彙項目間類似度の分布傾向

研究課題 (b) 「心内辞書内の語彙項目間類似度分布に特徴的な傾向が存在するか」に関して、表1及び表2から次の5点が指摘できる。

- (1) 語彙項目間類似度が極めて高い ( $T < 0.2$ ) 語彙項目ペア (*start-begin*, *walk-run*, *listen-hear* など) の頻度に関して、3群間に統計的に有意な差はない。また、いずれの群でも語彙項目間類似度が極めて高いこの項目ペアの産出は極めて少ない (7~12個)。
- (2) 語彙項目間類似度が高い ( $0.2 \leq T < 0.5$ ) 語彙項目ペア (*add-increase*, *tell-ask*, *hold-have* など) の頻度には、群間に有意な違いがあるものかないものが併存している。具体的には、SMALL群とNS群、またNS群とLARGE群の間にはこの語彙項目間類似度が高い語彙項目ペアの頻度に有意な差異があった (いずれにおいても、各組み合わせの后者の群の頻度が前者より有意に高い) が、他の4つのペアワイズ比較においては頻度に有意な差異は認められなかった。加えて、この語彙項目間類似度区分の語彙項目ペアは21~58個で、語彙項目間類似度が極めて高い ( $T < 0.2$ ) 語彙項目ペア数よりもすべての群で多かった。
- (3) 語彙項目間類似度が中程度~低い ( $0.5 \leq T < 0.9$ ) 語彙項目ペア (*listen-ask*, *speak-sing*, *make-bring* など) の頻度には、群間に有意な違いがあるものかないものが併存している。具体的には、SMALL群とLARGE群、SMALL群とNS群、またLARGE群とNS群の間にはこの語彙項目間類似度が中程度~低い語彙項目ペアの頻度に有意な差異があった (すべての組み合わせについて后者の群の頻度が前者より有意に高い) が、他の3つのペアワイズ比較においては頻度に有意な差異はなかった。加えて、この語彙項目間類似度区分の語彙項目ペアは153~392個で、語彙項目間類似度が高い ( $0.2 \leq T < 0.5$ ) 語彙項目ペア数よりもいずれの群でも多かった。

- (4) 語彙項目間類似度が極めて低い ( $0.9 \leq T$ ) 語彙項目ペア (*die-hear, look-make, rise-show* など) の頻度には、群間に有意な違いがあるものとないものが併存している。具体的には、LARGE 群と SMALL 群、NS 群と SMALL 群、また NS 群と LARGE 群の間にはこの語彙項目間類似度が極めて低い語彙項目ペアの頻度に有意な差異があった (いずれにおいても、各組み合わせの後者の群の頻度が前者より有意に高い) が、他の3つのペアワイズ比較においては頻度に有意な差異は認められなかった。(2) 及び (3) と異なり、母語話者より第二言語話者が、また後者では語彙サイズの小さい群が大きい群よりも、この語彙項目間類似度が極めて低い語彙項目ペアの頻度が有意に高かった。加えて、この語彙項目間類似度区分の語彙項目ペアは769～1044個で、語彙項目間類似度が中程度～低い ( $0.5 \leq T < 0.9$ ) 語彙項目ペア数よりもいずれの群でも多かった。
- (5) すべての群において、語彙項目間類似度値が高い領域ほど産出した頻度が高い。つまり、類似度の低い語彙項目ペア数が高いペア数よりも多いという傾向があった。

このように、心内辞書内の語彙項目間類似度分布には、4つの語彙項目間類似度区分のうち3つにおいて群間の構造の違いを示す特徴があることが明らかとなった。折田他 (2013) において、高頻度英語動詞の意味的クラスタリング構造における母語話者群と第二言語話者群の間で、また後者では語彙サイズの大きさの異なる被験者群の間で群間の相違が確認されたことと同様に、語彙項目間類似度分布においても群間に差異があることが判明した。

## 6. 考察

「5. 結果」から、高頻度英語動詞群から構成される NS 群、LARGE 群、そして SMALL 群の心内辞書内の語彙項目間類似度には計量的な差異があること、そして語彙項目間類似度分布には群間の違いを特徴付ける差異が存在することがわかった。語彙項目間類似度が高い ( $0.2 \leq T < 0.5$ ) 語彙項目ペア、語彙項目間類似度が中程度～低い ( $0.5 \leq T < 0.9$ ) 語彙項目ペア、さらには語彙項目間類似度が極めて低い ( $0.9 \leq T$ ) 語彙項目ペアの3つの頻度分布においては群間に統計的に有意な差異があるものとないものが併存したが、語彙項目間類似度が極めて高い ( $T < 0.2$ ) 語彙項目ペアの頻度には群間に有意な違いはなかった。このように、極めて高い語彙項目間類似度を持つ語彙項目ペアの頻度を除いて、産出した頻度に群間の違いがあった。加えて、全ての群で類似度が低い語彙項目ペアほどその産出数が多くなるという傾向が確認できた。以下、この結果を踏まえて、3点 (①類似度が極めて高い、及び高い語彙項目ペア頻度データの示唆すること、②語彙サイズと語彙項目間類似度分布の関係、③仕分け課題結果による英語習熟度予測の可能性) 考察する。

まず、語彙項目間類似度が極めて高い ( $T < 0.2$ ) 語彙項目ペア、また高い ( $0.2 \leq T < 0.5$ ) 語彙項目ペアの頻度 (%) はいずれの群でも少ないことがある。NS 群で前者が12個 (1.0%)、後者が52個 (4.2%)、LARGE 群で前者が9個 (0.7%)、後者が58個 (4.7%)、また SMALL 群で前者が7個 (0.6%)、後者が21個 (1.7%) にすぎなかった。加えて、この2つの語彙項目間類似度区分に関しては、3群間に明確な頻度の違いがなかった (ただし、語彙項目間類似度が高い ( $0.2 \leq T < 0.5$ ) 語彙項目ペアの頻度については SMALL 群と NS 群の間、また LARGE 群と NS 群の間の2つに有意差あり)。この結果は、本研究で用いた語彙項目には母語と第二言語の違いや被験者の語彙サイズに拘わらず、心内辞書内において意味的距離が近い語彙項目ペアが少なかったことを反映している。換言すると、3

つの群間で頻度に有意な差がなかったこれらの区分に入る *start-begin*, *walk-run*, *listen-hear* といった少数の語彙項目ペアは、母語であるかそうでないか、あるいはまた語彙サイズが大きいか小さいかに拘わらず、被験者の心内辞書内で同じように構造化されていると考えられる。第二言語語彙の発達において、全ての語彙が同じ過程をたどるわけではないと Schmitt (2000, 2010) は述べているが、心内辞書内で意味的距離が近い語彙項目ペア間においては群間に差異はないものが少数は存在することが指摘できる。そして、少なくとも本研究で用いたこれらの語彙項目ペアでは、母語と第二言語の間で形態上の違いはあっても心内辞書内に格納されている概念知識に基づく意味的な結び付きにおいては実質的な違いはない可能性が高いと言える<sup>5)</sup>。

2点目として、被験者群の語彙サイズと語彙項目類似度分布に一定の関係性が見られたのは、語彙項目間類似度が中程度～低い ( $0.5 \leq T < 0.9$ ) 語彙項目ペア及び語彙項目間類似度が極めて低い ( $0.9 \leq T$ ) 語彙項目ペアの2つの区分のみであったことである。類似度が中程度～低い語彙項目ペアにおいては、被験者群の語彙サイズが大きくなるほど、この区分の頻度が高くなる (SMALL群 153 < LARGE群 160 < NS群 392)。また、類似度が極めて高い語彙項目ペア、そして高い語彙項目ペアでは語彙サイズと類似度分布に一定の関係性はなく、前述の通り、中程度～低い類似度水準になって初めて語彙サイズの違いが語彙項目間類似度頻度の違いに反映している。これとは逆の傾向、つまり、語彙サイズが大きくなればなるほど産出が少ないのが語彙項目間類似度の極めて低い ( $0.9 \leq T$ ) 語彙項目ペアの頻度である (SMALL群 1044 < LARGE群 998 < NS群 769)。これは、本研究で用いた50個の英語動詞について、NS群がNNS群よりも、またLARGE群がSMALL群よりも多くの語彙項目間類似性を見出し得ることを示している。翻って、3群の中ではSMALL群において語彙項目間類似度が中程度～低い ( $0.5 \leq T < 0.9$ ) 語彙項目ペアの頻度が一番少なく、極めて低い ( $0.9 \leq T$ ) 語彙項目ペア頻度が一番高い。語彙サイズが小さい被験者群では、そもそも、どのような語彙項目の間においてもペア間の意味的關係性を見出すことが困難であることが示唆される。語彙サイズが小さい被験者の場合、NS群もしくはLARGE群が関係を見出すペアについて、関係性を見出すことができない (つまり、心内辞書内のネットワークの密度が低い) 可能性が高い。LARGE群、そしてNS群と、語彙サイズが大きくなるにつれて、この傾向は消失し語彙項目間の関係性をよりの確に認識していると言えよう。被験者が持つ語彙サイズが実験に用いた50語の各組み合わせ間の意味的距離、とりわけ「近さ」の認識に強く関係していると推定される<sup>6)</sup>。

3点目に、仕分け課題結果による英語習熟度予測の可能性について述べる。望月 (2010) は、日本人英語学習者に関して、心内辞書内の語彙構成度 (測定にはLOT (Lexical Organization Test) (望月, 2006) を使用) と英語習熟度や語彙サイズの間には中程度 (.37 ~ .54) の相関があったと報告している。同様に、Wolter (2002) は、日本人英語学習者を被験者として、単語連想テスト結果とC-testの結果の相関度を解析し、単語連想テストが被験者の英語習熟度を予測できるか否かを探ったが、同じく中程度の相関 (.44) しか得られなかった。このように、いずれの研究でも変数間の説明率 (相関係数の二乗) は30%にも満たない。心理言語学実験結果から精度の高い英語習熟度の予測を行うことは容易でないことがうかがえる。本研究においては、「5. 結果」で明らかになったように、語彙項目間類似度という指標について、「極めて高い」「高い」語彙項目ペアの頻度は高くなく、これに依拠して被験者の英語習熟度との関係を探ることは少なくとも本実験で使用した英単語からは可能性が低いと言える。被験者群の語彙サイズと語彙項目間類似度の頻度分布に一定の関係性が確認できたのは、



語彙項目間類似度が中程度～低い語彙項目ペア及び極めて低い語彙項目ペアであった。この分布区分においては仕分け課題結果から英語習熟度予測の可能性はあると言えよう。しかし、語彙サイズであれ、あるいはTOEFL や TOEIC などであれ、日本人英語学習者の英語習熟度を仕分け課題結果から予測するには、あくまでも仕分け課題で得られたデータ全体の構造を解析した上で、予測可能性を探る科学的な視点が欠かせない。そのためには、例えば、仕分け課題に取り組んだ被験者のTOEIC得点を説明変数(X)とし、従属変数(Y)を以下の(Y1)～(Y4)の4通りに対してそれぞれ回帰分析を行い、従属変数が説明変数からどの程度説明可能であるかを定量的に解析する必要がある。

- (Y1) 母語話者の心内辞書からの各被験者の心内辞書の距離（デンドログラム距離行列のプロベニウスノルム）
- (Y2) 各被験者の用いた皿（単語クラスター）の数（未仕分けの単語はそれぞれ別の皿に分けられたとしてカウント）
- (Y3) 各被験者の皿ごとの平均単語数（全単語数を皿の数で割ったもの）
- (Y4) 各被験者の皿ごとの単語数の分散

上記の観点による解析結果から、仕分け課題結果の被験者のばらつき度を確認し、仕分け課題結果から英語習熟度の予測が可能か否か判断する必要がある。このことは今後の研究課題としたい。

## 7. 結論

本研究は、折田他(2013)で明らかとなった、英語心内辞書内の語彙項目の結び付きにおける意味的クラスタリング構造が母語話者と日本人英語学習者、また後者では語彙サイズの違いによって異なることを出発点とした。このことを踏まえ、心内辞書内語彙項目の最も基本的な結び付き単位である語彙項目ペアの間にも、同様に群間の相違が存在するのかが否かを明らかにすることを目的に、仕分け課題結果の語彙項目間類似度に着目し解析を行った。解析結果から、3群の語彙項目間類似度の頻度は類似度の度合いによって分布に有意な偏りがあることが判明した。つまり、語彙項目間類似度という指標からも、母語話者と日本人英語学習者、また後者では語彙サイズの違いによって心内辞書内の意味的クラスタリング構造には違いがあることがわかった。また、語彙項目間類似度の程度によって、群間の違いが明らかな語彙項目ペアとそうでないものが併存することが明らかとなった。類似度が極めて高い、また高い語彙項目ペアについては一定の方向性のある3群間の差異はなく、一方、類似度が中程度～低い、また極めて低い語彙項目ペアにおいては、語彙サイズが大きい被験者群ほど（さらに、日本人英語学習者よりも英語母語話者の方が）前者（類似度が中程度～低い語彙項目ペア）の頻度が高く、後者（極めて低いペア）の頻度が少ないことが明らかとなった。このように、語彙項目間類似度の違いによる語彙項目ペアの分布は一樣ではなく、高い類似度を持つペアにおいては母語話者、日本人英語学習者間にはっきりした相違が見られないのに対して、類似度が中程度～低い、そして極めて低いペアにおいて母語話者と日本人英語学習者、また後者では語彙サイズの違いによる相違が顕著であると結論付けられる。

## 注

- 1) 英語母語話者の語彙サイズに関しては、研究者や語彙数の数え方でかなり幅がある。ここでは、第二語彙習得研究で第一人者の Nation (2001) の研究を参考にした。
- 2) 単語仕分け課題に用いる単語がたとえ無作為に選ばれていても、被験者はそれらを意味的に関連する単語群にグループ分けする (折田・小林, 2012)。本研究では、実験に用いる英語動詞抽出の初期手続きにおいて、語彙的結束性に優れていると判断される *Treasure Island* (1883) に依拠した。課題で被験者が、単語間の意味的関連性を発見しやすい単語群を抽出するための手続きである。なお、実験に用いた高頻度の英語動詞はいずれも多義性を少なからず持ち、また現代英語と異なる意味もある単語が含まれている可能性があり、現代英語で使われていることを確認した。その結果、抽出した50語に変更はなかった。
- 3) NNS群の被験者であった大学生には授業の中で単語仕分け課題に取り組みさせることが可能なことから、各自パソコンを使う方法を採用した。被験者個々に実施せざるを得なかったNS群についてはこの方法を用いることができず、印刷したカードを使う方法を採用した。方法の違いによる何らかの影響があった可能性は否めないが、単語を意味の上で関連するグループに分けるという課題の眼目は共通しており、本質的な問題はないと考える。
- 4) Tの分布は非常に非対称であり、特に語彙項目間類似度が極めて高い語彙項目ペアについては、 $T < 0.1$  ではサンプル数が少なく検定の信頼性を得られない可能性があり、代わりに  $T < 0.2$  を用いることとした。  
また、頻度データについては  $\chi^2$  検定を用いて解析することが少なくないが、本研究で収集したデータに対して  $\chi^2$  検定を行うには統計学的に以下の問題があり採用しなかった。
  - i) 距離行列の各要素は独立ではない。そのため通常の  $\chi^2$  検定での自由度は大きすぎ、帰無仮説を棄却しやすくなる。
  - ii) 群平均法などの構成法によるデンドログラム距離行列については、さらにその傾向が強くなる恐れがある。
- 5) 母語と第二言語の概念知識の近さに起因するとするこれらの語彙項目ペアに対する解釈は、レベル1の中でもより高い頻度の語と低い頻度の語があり、たまたまこれらは高頻度語で被験者が意味を知っていたため、という可能性があるのではないかという指摘が一人の査読者からあった。首肯できる指摘であり、ここの解釈を本研究で用いた語彙項目ペアについて限定する記述に改めた。
- 6) 本稿執筆に際し、語彙項目類似度の低いデータと語彙サイズの小さい被験者群の関係について、坂田直樹先生の指摘を踏まえて解釈を修正した。

## 謝辞

本稿は、大学英語教育学会第52回国際大会 (2013年8月31日, 京都大学吉田キャンパス) において筆者らが行った口頭発表「母語話者と第二言語話者の英語心内辞書における語彙項目間類似度の比較」を加筆修正したものである。発表に対して、坂田直樹先生 (近畿大学) から貴重なご指摘を賜った。また、査読者のお二人から本稿全般について有益なご助言を頂いた。記して深謝申し上げたい。本研究は、科学研究費補助金基盤研究(C)「語彙知識のネットワーク化を促進するオンライン自学教材の開発」(課題番号 25370634: 研究代表者 折田充)、科学研究費補助金若手研究(B)「代数的性質を用いた新しい統計解析手法の開発」(課題番号 24700288: 研究代表者 小林景) 及び統計数理研究所共同利用所要経費一般研究2「英語心内辞書データの統計的解析」(25-共研-2067: 研究代表者 小林景) の助成を受けている。

## 参考文献

- 折田充・小林景 (2010, 12月). 「心内辞書内の意味的クラスタリング—母語話者と第二言語話者の相違」. 第39回九州英語教育学会, 鹿児島大学.
- 折田充・小林景 (2011a). 「心内辞書内の意味的クラスタリング構造—L1とL2の違いの指標となり得る語彙類の特定」『熊本大学社会文化研究』, 9, 19-37.
- 折田充・小林景 (2011b). 「心内辞書内の意味的クラスタリング—高頻度英語形容詞における母語話者と第二言語話者の相違」, *KASELE Bulletin*, 39, 1-11.
- 折田充・小林景 (2011c). 「心内辞書内の意味的クラスタリング構造 (3) —高頻度英語動詞における英語母語話者と日本人英語話者の相違」『第37回全国英語教育学会山形研究大会発表予稿集』, 340-341.
- 折田充・小林景 (2012). 「母語の心内辞書と第二言語の心内辞書—日本人英語学習者における日英語間で訳語関係にある語彙項目群の構造」, *KASELE Bulletin*, 40, 1-10.
- 折田充・小林景・村里泰昭・神本忠光・吉井誠・Lavin, R. (2013). 「語彙サイズと心内辞書内の意味的クラスタリング構造の関係」, *KASELE Bulletin*, 41, 1-10.
- 小林景・折田充 (2009, 9月). 「日本人と英語母語話者との心内辞書構造の相違の統計的解析」. 2009年度統計関連学会連合大会, 同志社大学.
- 小林景・折田充 (2012). 「英語心内辞書の木構造データ解析新手法」『行動計量学会第40回大会抄録集』, 101-104.
- 小林景・折田充 (2013). 「木構造およびクラスター構造をもつデータの測地的解析手法」『2013年度統計関連学会連合大会予稿集』, 313.
- 齋藤堯幸・宿久洋 (2006). 『関連性データの解析法—多次元尺度構成法とクラスター分析法』. 東京: 共立出版.
- 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 JACET List of 8000 Basic Words』. 東京: 大学英語教育学会.
- 竹村彰通 (1991). 『現代数理統計学』. 東京: 創文社.
- 望月正道 (2006, 8月). 「新しい語彙構成テスト開発の試み」. 第32回全国英語教育学会高知研究大会, 高知大学.
- 望月正道 (2010). 『技能別及び総合的英語能力を推定する語彙テストの開発: サイズ, 構成, 認知速度の融合』. 平成19-21年度科学研究費補助金基盤研究 (B) 研究成果報告書 研究課題番号19320084.
- Fitzpatrick, T. (2006). Habits and rabbits: Word associations and the L2 lexicon. *EUROSLA Yearbook*, 6, 121-145.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(1), 9-17.
- Orita, M. (2002). Word associations of Japanese EFL learners and native speakers: Shifts in response type distribution and the associative development of individual words. *Annual Review of English Language Education in Japan*, 13, 111-120.
- Orita, M., & Kobayashi, K. (2009, December). *Predictors of L1 and L2 differences in lexical organisation*. Paper presented at The 6th JACET Vocabulary Research Group Annual Conference, Reitaku University Tokyo Research Center, Tokyo.
- Orita, M., & Kobayashi, K. (2010, March). *Effects of intra-lexical features on the completion time of sorting tasks*. Paper presented at The 20th Vocabulary Acquisition Research Group Network

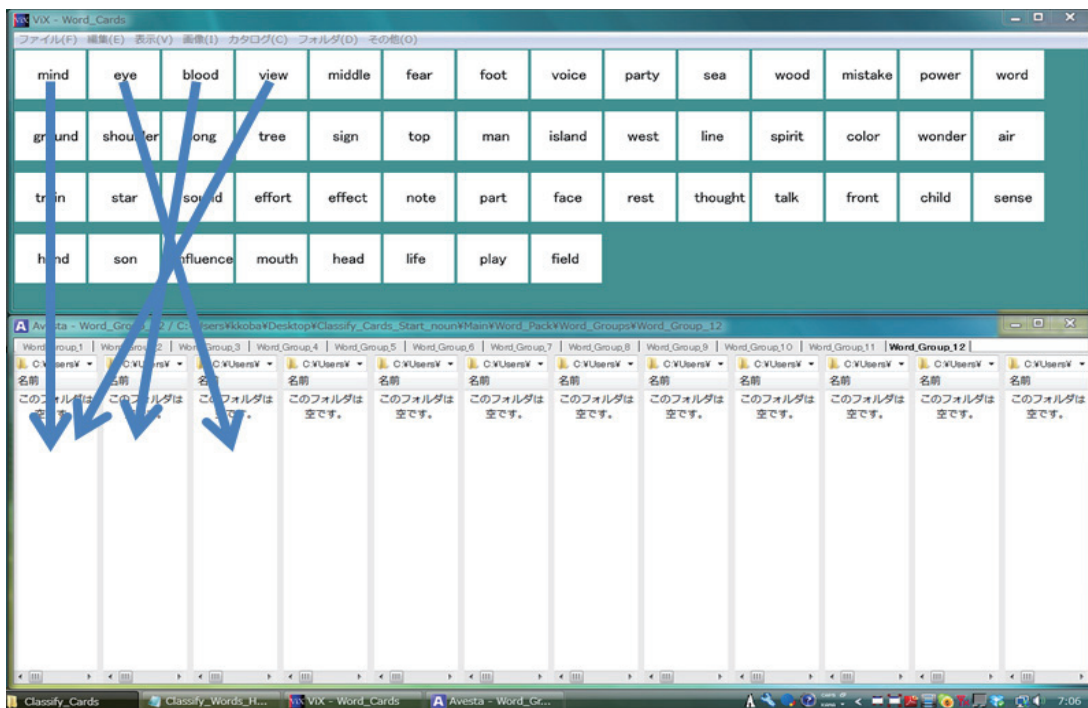
- Conference, Newtown, Wales, UK.
- Orita, M., & Kobayashi, K. (2011, March). *Semantic clustering of high frequency nouns in L1 and L2 mental lexicons*. Paper presented at Learners and Networks Conference 2011, Swansea, Wales, UK.
- Orita, M., & Kobayashi, K. (2012, March). *Semantically equivalent lexical items between L1 and L2 mental lexicons*. Paper presented at The 22nd Vocabulary Acquisition Research Network Conference, Newtown, Wales, UK.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Stevenson, R. (1883). *Treasure Island*. Electronic Text Center, University of Virginia Library. Retrieved 4 December 2005, from <http://etext.lib.virginia.edu/toc/modeng/public/SteTrea.html>
- Wilks, C., & Meara, P. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research*, 18, 303-324.
- Wilks, C., & Meara, P. (2007). Implementing graph theory approaches to the exploration of density and structure in L1 and L2 word association networks. In H. Daller, J. Milton & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 167-181), Cambridge: Cambridge University Press.
- 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.
- Wolter, B. (2002). Assessing proficiency through word associations: Is there still hope? *System*, 30, 315-329.

## APPENDIX

## 単語仕分け課題

## 日本人被験者用指示文及びスクリーンショットイメージ

Word\_Cardsの中の50枚のカードを、自分が思う、意味の上で関連しているグループに分けてWord\_Groupsの「皿」に移動して下さい。



## Mental Lexicons of Native English Speakers and Japanese Learners of English: Degrees of Similarity Between Lexical Items

Mitsuru ORITA, Kei KOBAYASHI, Yasuaki MURASATO, Richard S. LAVIN,  
Makoto YOSHII, Kazumi AIZAWA & Tadamitsu KAMIMOTO

This study addresses the question as to whether the degrees of similarity between lexical items differ between an L1 mental lexicon and L2 mental lexicons of different sizes. T is a measure of the distance between a pair of lexical items, ranging from 0.0 to 1.0; the lower T is, the higher the degree of similarity. A free sorting task using a set of 50 high-frequency English verbs was given to 30 native speakers of English (NS), 30 Japanese speakers of English with large vocabulary size ( $M=6273.3$ ,  $SD=258.6$ ) (LARGE) and 30 Japanese speakers of English with small vocabulary size ( $M=4683.3$ ,  $SD=155.5$ ) (SMALL),  $t=28.86$ ,  $df=58$ ,  $p<.01$ ,  $r=.97$ . The obtained T results of dendrogram distance, based on the Hamming distances of all the lexical item pairs, for each group were put into matrices and converted into frequency data. The data of the three groups were then submitted to permutation tests (six different pair-wise comparisons). The analysis revealed that the Ts of the three groups were statistically different from each other at the 5% significance level. NS produced more lexical item pairs having moderate to low Ts ( $0.5 \leq T < 0.9$ ), followed by LARGE, and SMALL produced the least. In the case of lexical item pairs with very low Ts ( $0.9 \leq T$ ), however, the results were reversed, with SMALL producing them in the highest number, followed by LARGE, and NS the lowest.