

# 母語話者と第二言語話者の心内辞書

## — 語類の混在する単語群における意味的クラスタリング構造 —

折田 充・小林 景

### はじめに

私たちが日頃日本語で読み書きするときであれ、英語を学び始めた中学生が何とか英語で意志疎通を図ろうとするときであれ、あるいは日本語と英語に同程度の熟達度を持つバイリンガルが英語で話すときであれ、ひとは脳内にある「辞書」を使う。この脳内機構は心内辞書 (mental lexicon) と呼ばれ、言語の形態、意味、機能、また付随する体験的知識や記憶など語彙に関する多様な情報が格納されている。心内辞書内では各単語はアルファベット順には並んでおらず、発音や綴りのみに基づく構造でもなく、並列分散処理 (parallel distributed processing) 的な索引機能を持つネットワーク構造体と言える。加えて、心内辞書には常に新しい情報が書き加えられ、書き換えられ、あるいは消去され、その内容や規模は変化する。例えば、私たちが英語を学び続ける中で、語彙サイズは大きくなり既知の単語に新たな意味を発見しそれを心内辞書に加えることがそれにあたる。心内辞書はこのように可塑性に富み、膨大な量の情報が関係するネットワーク構造を持つ脳内機構である。本稿では、心内辞書内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) に注目する。名詞、あるいは動詞などと語類を特定しない、つまり語類の混在する単語群における意味的クラスタリング構造について、母語話者と第二言語話者間の相違を明らかにしたい。

### 1. 先行研究

これまで母語話者と第二言語話者間の心内辞書内構造の違いについて、単語連想テスト (word association test: 被験者に刺激語 (stimulus) を見せたり聞かせたりして、それに対して表出する反応語 (response) を分析することで、心内辞書内構造を探ろうとする心理言語学実験) を用いた Fitzpatrick (2006) や Meara (1978, 1983)、また Nissen & Henriksen (2006) など多くの先行研究が、母語話者よりも第二言語話者の方が実験結果にばらつきが大きいと報告してきた (同様の報告として Kikuchi, Yamamoto, Yoshimura, Yabuuchi, & Tanimura, 2001; Meara & Schur, 2002; Orita, 2002; Postman & Keppel, 1970; Riegel & Zivian, 1972; Schmitt, 1998; Szalay & Deese, 1978; Söderman, 1989, 1993; Yokokawa Yabuuchi, Kadota, Nakanishi, & Noro, 2002; Zareva, 2007を参照のこと)。心内辞書内の構造は母語話者間では一定の安定した傾向がある反面、第二言語話者の間では不安定で個人差が大きいことが反映している。心内辞書内のクラスター構造の解明に焦点を当て、単語仕分け課題 (word sorting task) を用いた折田・小林 (2011a, 2011b, 2011c) では、名詞については前述の先行研究を支持する結果となったが、動詞と形容詞については両群間に有意な差異は検出されなかった。第二言語

話者においても語類により心内辞書内のクラスタリング構造のばらつき度に違いがあることが判明した。

心内辞書内の語彙項目が形成しているクラスターについて Meara & Schur (2002) は、母語話者の生成する平均クラスター数は第二言語話者よりも多く、平均クラスターサイズ（個々のクラスターを構成する単語数）は母語話者の方が小さかったと報告している。この結果は、母語話者の方が第二言語話者よりも心内辞書内の単語間の意味的関連性を明確に意識化できていることに拠ると Meara & Schur は説明している。折田・小林 (2011a) では、動詞や形容詞においては両群間に有意な差異は検出されず、名詞についてのみ有意に母語話者の方が第二言語話者よりも平均クラスター数が多く、平均クラスターサイズが小さいという結果となった。ばらつき度と同様に、心内辞書内の平均クラスター数や平均クラスターサイズは母語話者と第二言語話者間の違いのある語類とそうでないものが併存していると言える。

一方、折田・小林による一連の研究（折田・小林, 2011a, 2011b, 2011c）から、母語話者と第二言語話者の心内辞書内の意味的クラスタリング構造に関して、名詞、動詞、形容詞のいずれにおいても質的な差異があることが明らかになった。まず名詞について、上級レベルの日本人英語話者でも母語で形成された概念知識（conceptual knowledge）や“偽りの友（false friends）”の関係を持つ英語からの借用語の影響が強固に残り母語からの転移として作用し、日本語の中で借用語として定着しているカタカナ語に対応する英単語は、英語の熟達度が上がっても英語心内辞書内において母語話者に近似する構造を形成することが難しいことが判明した（折田・小林, 2011a）。動詞については、母語話者の持つ構造をベースラインとするとき4つのタイプに分かれることが判明した。上級話者では母語話者との類似性が高いにも関わらず中級話者ではむしろ独自の未発達な要素を含む単語群、上級話者では母語話者との類似性が高く中級話者では両群と同様のクラスターを形成しつつも異なる語群と結び付く単語群、中級話者と上級話者の同質性は高い半面母語話者の構造に近似しにくい単語群、そして高い3群間の類似性が確認できるが部分的には構造上の違いがある単語群の4つである（折田・小林, 2011c）。形容詞についても、第二言語話者には母語話者の持つクラスタリング構造の獲得が困難な領域、また第二言語話者の間でも上級話者と中級話者の間で異なる構造を持つものなどがあり様ではない（折田・小林, 2011b）。例えば、母語話者では *hot-light-hard* の3語からなる「程度」クラスターが形成されている一方、上級話者における「程度」は「時間」「空間」などの意味領域について26語からなる巨大なクラスターの一部を構成していた。また、*hot-hard* は *strong* と類似度が高いが、*light* は *little-small-large-big* と類似度が高く、母語話者と構造が異なる。中級話者では、*light* は「空間」「形状」「大きさ」「状態」クラスターに入り、*big-small-large-wide-little-long-deep-high-low* サブクラスターと類似度が高かった。心内辞書内の意味的クラスタリング構造は、熟達度が上がるにつれて発達・変容していく領域と熟達度に関わらず母語話者と第二言語話者の間で異なる領域が併存している。

このように、心内辞書内の質的構造について、母語と第二言語、また後者については習熟度の違いによって互いに異なる側面が少なくなく、その違いは母語話者と日本人英語話者間に存在するもの、中級レベルの日本人英語話者と他の2群間にあるもの、また群間の同質性は高いが部分的に異なる要素を含むものなどと多様である。第二言語話者の意味的クラスタリング構造の質的な発達は、母語話者の構造を基準に置くと、母語話者のものへの発達・近似が比較的容易なものと困難なものが併

存する可能性が高いと推定できる。しかし、折田・小林 (2011a, 2011b, 2011c) は、名詞、動詞、形容詞という主要な単一の語類を用いた研究である。特定の語類内でのみ単語間のネットワークが形成されているわけではない心内辞書本来の特性からみて、語類の混在する単語群におけるクラスタリング構造にはどのような特徴があるのか解明することは、心内辞書の実相解明のためには不可欠である。本研究では、特定の語類に注目して得られた折田・小林による一連の研究結果が、語類の混在する単語群における心内辞書内の意味的クラスタリング構造についても当てはまるのか明らかにしたい。

## 2. 研究課題

前述した先行研究の概観を踏まえて、本研究では次の3つの研究課題について解明することを目指す。

- (a) 母語話者は第二言語話者よりも単語間の結び付きを明確に意識し、単語仕分け課題結果において母語話者の作る平均クラスター数は第二言語話者よりも多く、また平均クラスターサイズは小さいか。
- (b) 心内辞書内構造の個人差は母語話者よりも第二言語話者の方が大きく、仕分け課題結果において母語話者よりも第二言語話者の方が被験者間のばらつき度が大きいのか。
- (c) 単語仕分け課題結果において、母語話者と第二言語話者の心内辞書内のクラスター構造には質的な違いがあるか。

## 3. 方法

### 3.1 被験者

本研究の被験者は2つの群から構成された。一つは英語母語話者群28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。もう一つの群 (28名) は、上級レベルの英語力を持つ日本人英語話者で、熊本大学の常勤あるいは非常勤の英語教師、または TOEIC で730点以上、あるいは他の公的な英語熟達度試験結果から同等の英語力を持つと判断された大学生や社会人である。被験者として協力いただいた方には薄謝を差し上げた。

### 3.2 単語仕分け課題

本研究では、折田・小林 (2011a, 2011b, 2011c) と同様に、心内辞書内の意味的クラスタリング構造解明において妥当性が高い単語仕分け課題を採用した。課題で用いる単語として、『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) の順位 1 - 500位の高頻度英単語の中から無作為に50語を選んだ。用いた英単語は次の通り: *ago, air, all, already, also, area, arm, arrive, believe, boy, business, century, clear, close, country, cry, doctor, dream, early, figure, find, form, help, history, keep, law, lot, matter, nature, next, nothing, open, our, person, place, police, power, president, reason, shop, social, stand, step, street, then, understand, very, walk, war, while*。各語類の内訳は、名詞30語 (60%)、動詞5語 (10%)、形容詞6語 (12%)、副詞6語 (12%)、代名詞1語 (2%)、指示代名詞1語 (2%)、接続詞1語 (2%) であった。各単語を2.4cm × 4.4cm 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、手渡された指示文 (Appendix

を参照のこと)に従い課題に取り組んだ。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、作成するクラスターの数やサイズに制限のない、つまり与えられた50語を被験者が考える意味のまとまりで思う通りにグループ分け(クラスター化)する自由仕分け課題(free sorting task)とした(制限時間は20分。なお、指示文は英語を母語とする被験者には英語版で示した)。高頻度の英単語は複数の品詞を担う場合が多く、本研究で用いた単語にもそのような単語(close, lot, open, reasonなど)が含まれていたが、心内辞書内の語類の混在する単語群における意味的クラスタリング構造を明らかにするという本研究の目的から、被験者が実験で用いた単語についていずれの品詞に基づき課題に取り組むかについて統制は加えなかった。実験は折田が被験者一人ずつに実施した。

### 3.3 データ解析

研究課題(a)について、クラスター数及びクラスターサイズに関して両群間の差異の有無を検定した( $t$ 検定)。なお、仕分け課題における各個人の結果に関して、被験者がいずれのクラスターにも入らないと判断した単語(作成されたどのクラスターにも入らなかった単語)は計算から除外した。次に、研究課題(b)について、まず研究課題(a)で得たクラスター数及びクラスターサイズのばらつき度を $F$ 検定結果に注目し確認した。併せて、各被験者の仕分け課題の計算結果の表象として得られる個人デンドログラムが、非類似度(dissimilarity)ゼロ(仕分けされて同じクラスター内に入った単語間の類似度)、あるいは1(仕分けされて異なるクラスターに入った単語間の類似度)の“高さ”から構成される距離行列であることに着目した。得られた両群データのばらつき度(分散)について群間で差異があるか否かを解析するために、各被験者の個人デンドログラムの群デンドログラム(距離行列)へのあてはまり度(各個人のデータの距離行列から各群のデンドログラムの距離行列へのフロベニウス距離の2乗平均。値が小さいほどあてはまり度が高い)を、並べ替え検定(permutation test)により、群内データのばらつき度は同一であるという帰無仮説のもとに解析した。その上で、得られた両群の結果についてウェルチの $t$ 検定により群間差異の有無を解析した。なお、デンドログラム間の距離の計算においては、ミンコフスキー距離が最も自然であると判断され、これを採用した(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)を参照のこと)。そして、研究課題(c)については、まず、クラスター分析(群平均法)によって得られた各実験結果における群デンドログラムのクラスター数を比較した。クラスター数(群デンドログラムをいくつに切り分けるか)の決定においては様々な方法が提案されているが、「結合距離が極端に変化する段階をクラスタリング結果として採用する。」(齋藤・宿久, 2006, p. 152)を踏まえ、またクラスター数は2~10個、望ましくは2~7個に決定するという原則(Hair, Black, Babin, Anderson, & Tatham, 2006)のもとに、群デンドログラムにおける各単語の類似度を降順に並べ、そのプロットにギャップの生じる箇所で切り分ける決定法(長谷川四郎, personal communication, 14 May 2007)により各群デンドログラムの最終クラスター数を決定した。さらに研究課題(c)への取り組みにおいては、両群の群デンドログラムの距離行列間に有意差はないという帰無仮説のもとに並べ替え検定を行い、その有意差の有無を検討した。研究課題(b)及び(c)における並べ替え検定には、MATLAB(Matrix Laboratory) Version 7.5.0.342 (R2007b) (The MathWorks, Inc., 2007)を用いた。なお、実施

した並べ替え検定の回数とその信頼性の確認については、「4. 結果」で述べる。

## 4. 結果

### 4.1 平均クラスター数、平均クラスターサイズ及びばらつき度

研究課題 (a) に関して、表 1 に英語母語話者及び日本人上級英語話者に課した単語仕分け課題の結果について、平均クラスター数と *SD* (標準偏差) 及び *t* 検定結果 (*t* 値) をまとめた。平均クラスター数は母語話者が 9.11、日本人が 8.54 と母語話者が若干多いという結果となったが、5%水準で両群間に有意差は認められなかった ( $t(54) = 0.66$ , n.s.)。また、標準偏差は、母語話者が 2.95、日本人が 3.50 で、日本人被験者間のばらつき度が大きいという結果となったが、両群間の有意差は検出されず ( $F(27) = 1.41$ , n.s.)、クラスター数のばらつき度に関しても両群間に有意な差異はないことが明らかとなった。つまり、名詞以外の形容詞や動詞において、英語母語話者と日本人英語話者が生成した平均クラスター数には差異がないという一連の折田・小林の研究結果 (折田・小林, 2011a, 2011b, 2011c) を支持する結果となった。これらのことから、クラスター数は母語話者と日本人英語学習者を区別する指標にはならない可能性が高いことが判明した。

表 1. 平均クラスター数

母語話者 ( $n = 28$ )		日本人 ( $n = 28$ )		<i>t</i> 値
平均	<i>SD</i>	平均	<i>SD</i>	
9.11	2.95	8.54	3.50	0.66

表 2 は、英語母語話者及び日本人上級英語話者に課した単語仕分け課題の結果について、クラスターサイズの平均と *SD* (標準偏差) 及び *t* 検定結果 (*t* 値) をまとめたものである。平均クラスターサイズは母語話者が 5.78、日本人が 6.41 と日本人の方がやや大きいという結果となったが、5%水準で両群間に有意差は認められなかった ( $t(54) = 0.91$ , n.s.)。標準偏差は、母語話者で 2.16、日本人で 2.95 と日本人が大きく、クラスター数と同様に日本人被験者間のばらつき度が大きいという傾向を示しているが、両群間の有意差は検出されなかった ( $F(27) = 1.87$ , n.s.)。クラスターサイズ及びそのばらつき度についても両群間に有意な差異はないことが判明した。

表 2. 平均クラスターサイズ

母語話者 ( $n = 28$ )		日本人 ( $n = 28$ )		<i>t</i> 値
平均	<i>SD</i>	平均	<i>SD</i>	
5.78	2.16	6.41	2.95	0.91

### 4.2 個人デンドログラムのばらつき度

研究課題 (b) に関して、英語母語話者及び日本人上級英語話者に課した単語仕分け課題の結果について並べ替え検定を行った。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、十分であることがわかった。表 3 に、個人デンドログラムのばらつき度の解析結果をまとめた。

表3. 個人デンドログラムのばらつき度

T_NS	T_JP	p 値
108.20	112.37	0.399

(注) T\_NS = 英語母語話者各個人のデータの距離行列から英語母語話者群の群デンドログラムの距離行列へのフロベニウス距離の2乗平均; T\_JP = 日本人英語話者各個人のデータの距離行列から日本人英語話者群の群デンドログラムの距離行列へのフロベニウス距離の2乗平均。T\_NSとT\_JP間の有意差の検定はウェルチのt検定による。

表3から、T\_NS（英語母語話者各個人のデータの距離行列から英語母語話者群の群デンドログラムの距離行列へのフロベニウス距離の2乗平均）のばらつき度（108.20）とT\_JP（日本人英語話者各個人のデータの距離行列から日本人英語話者群の群デンドログラムの距離行列へのフロベニウス距離の2乗平均）のばらつき度（112.37）について、母語話者群の方が日本人よりもばらつき度が小さいという傾向はあったが、両群間に有意差は認められなかった（ $p = 0.399$ ）。上述した平均クラスター数や平均クラスターサイズにおけるばらつき度の場合と同様に、個人デンドログラムのばらつき度についても両群間に有意な差異はなく、ばらつき度は英語母語話者と日本人英語話者の心内辞書内のクラスター構造の相違を示す指標とはならないことが明らかとなった。

#### 4.3 意味的クラスタリング構造の質的差異

研究課題(c)について、単語仕分け課題結果における両群の群デンドログラムの距離行列間に有意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。並べ替え検定の繰り返し回数は、個人デンドログラムのばらつき度に関する解析と同様に、5,000回とした。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十分か否かを確認したところ、いずれの場合も十分であることがわかった。得られた並べ替え検定のp値0.012（5%有意水準、両側検定）から、英語母語話者群と日本人英語話者群には有意な差異があることが明らかとなった。本研究による語類の混在する単語群を用いた仕分け課題結果の群デンドログラム距離行列の距離に対する並べ替え検定においても、動詞、名詞、形容詞を用いた単語仕分け課題結果の群間比較の場合と同様に、英語母語話者群と日本人英語話者群の心内辞書内の意味的クラスタリング構造は有意に異なることが判明した。これらの結果は、全ての語類に関して両群の構造が有意に異なることを報告した折田・小林（2011a）を支持し、母語話者と日本人英語話者の心内辞書内の意味的クラスタリング構造は明確に異なると結論付けられる。

クラスター分析結果（群デンドログラム）を図1（英語母語話者）及び図2（日本人英語話者）に示した。両群の群デンドログラムについて、「3.3 データ解析」に示した方法に基づき、クラスター数は2～10個、望ましくは2～7個に決定するという原則のもとに、群デンドログラムにおける各単語の類似度を降順に並べ、そのプロットにギャップの生じる箇所を切り分け、各群デンドログラムの最終クラスター数を決定することとした。その結果、英語母語話者群では6つ、日本人英語話者群では4つに切り分けるのが妥当であることが確認できた。これは、最終クラスター数について3つの語類（動詞、名詞、形容詞）いずれを用いた実験結果においても、母語話者群が日本人英語話者群よりも多かった折田・小林（2011a）の結果を支持する。Meara & Schur（2002）が指摘するように、母語話者は第二言語話者よりも単語間の結び付きを明確に意識し心内辞書内で構造化している可能性が高いと言える。この結果に基づきクラスターの意味的クラスタリングを構成する英単語を特定し、併せて

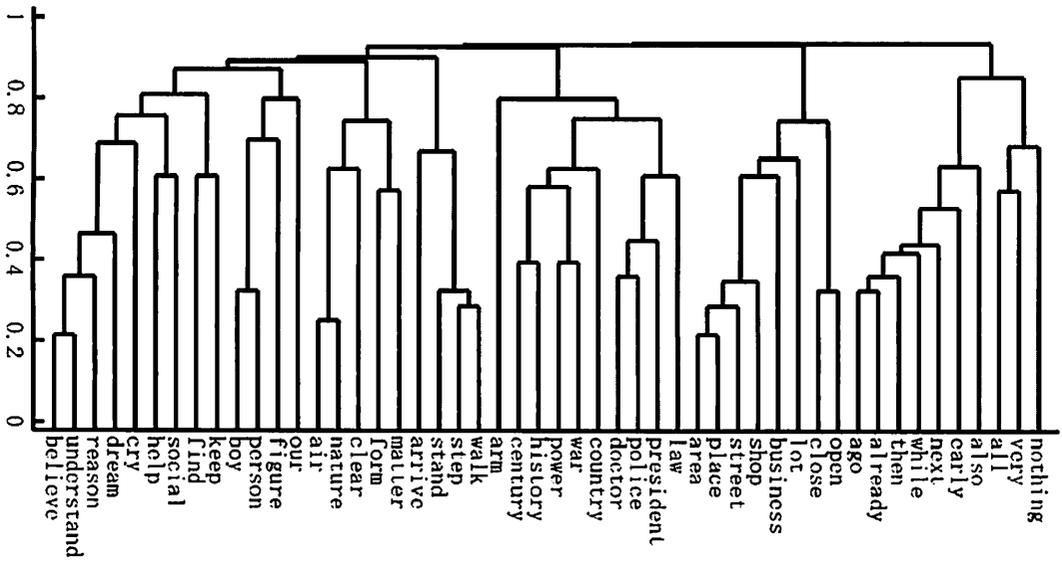


図1. 英語母語話者の仕分け課題結果の群デンドログラム：群平均法によるクラスター分析

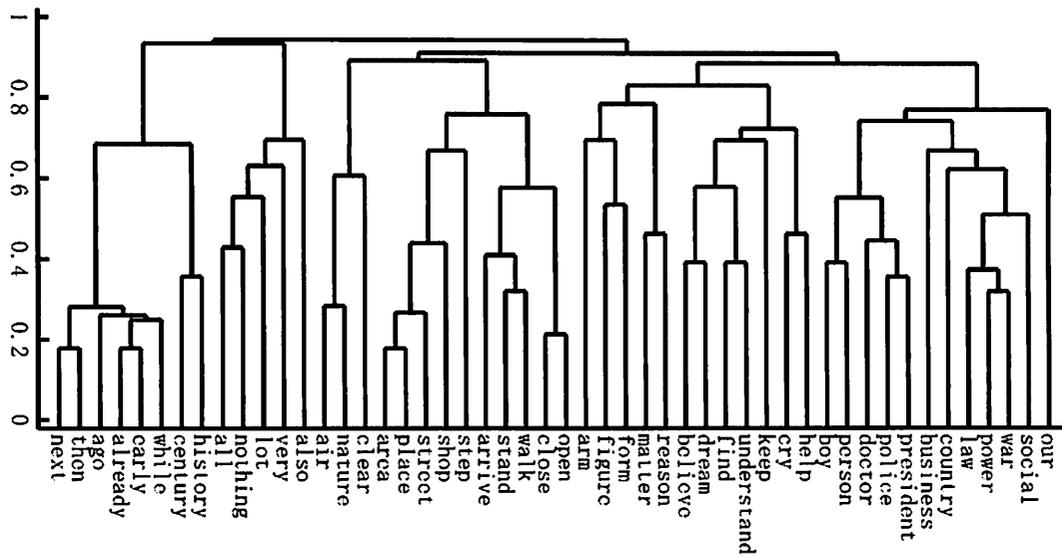


図2. 日本人英語話者の仕分け課題結果の群デンドログラム：群平均法によるクラスター分析

それぞれのクラスター名を決定した。表4a（英語母語話者）及び表4b（日本人英語話者）にクラスター名、各クラスターを構成する単語の頻度及びその単語を示す。

表 4 a. 英語母語話者の意味的クラスタリング ( $k = 50$ )

No.	クラスター名	頻度	単語
1	「程度」「時間」	10	<i>nothing, very, all, also, early, next, while, then, already, ago</i>
2	「状態」「場所」	8	<i>open, close, lot, business, shop, street, place, area</i>
3	「力」「歴史」	10	<i>law, president, police, doctor, country, war, power, history, century, arm</i>
4	「動き」	4	<i>walk, step, stand, arrive</i>
5	「物質」「自然」	5	<i>matter, form, clear, nature, air</i>
6	「人間」「認知」	13	<i>our, figure, person, boy, keep, find, social, help, cry, dream, reason, understand, believe</i>

表 4 b. 日本人英語話者の意味的クラスタリング ( $k = 50$ )

No.	クラスター名	頻度	単語
1	「力」「人間」 「認知」「物質」	24	<i>our, social, war, power, law, country, business, president, police, doctor, person, boy, help, cry, keep, understand, find, dream, believe, reason, matter, form, figure, arm</i>
2	「状態」「動き」 「場所」「自然」	13	<i>open, close, walk, stand, arrive, step, shop, street, place, area, clear, nature, air</i>
3	「程度」	5	<i>also, very, lot, nothing, all</i>
4	「歴史」「時間」	8	<i>history, century, while, early, already, ago, then, next</i>

表 4 a 及び表 4 b に示したように、「程度」「時間」「状態」「場所」など母語話者と第二言語話者間でクラスターの内部構造が部分的に類似するクラスタリングの傾向が確認できるが、両群間に共通するクラスターはない。母語話者と第二言語話者間で心内辞書内のクラスタリング構造には質的違いがあることが確認され、名詞、動詞、形容詞という特定の語類における両群間の質的違いを報告した折田・小林 (2011a, 2011b, 2011c) を支持することとなった。

## 5. 考察

語類の混在する単語群における心内辞書内の意味的クラスタリング構造について、まず、平均クラスター数、平均クラスターサイズ及びこれらのばらつき度、また個人デンドログラムのばらつき度のいずれに関しても英語母語話者と日本人英語話者の間には有意な差異が検出されず、これらの変数は両群間の違いを示す指標とならないことが判明した。名詞に関しては、「被験者の母語である日本語への借用語として定着しているものが多く、母語獲得におけるその強固な影響が L2 獲得における個々の単語レベルの学び直しの難しさのみならず、意味的クラスタリング構造の再構築の困難さにも深く関係している」(折田・小林, 2011a, p. 30) という特異性がある。しかし、本研究で対象とした語類の混在する語群、また折田・小林 (2011b) で報告した漸次性・段階性 (gradation) や反意語・同意語を意味関係の基盤とする英語形容詞 (Gross, Fischer, & Miller, 1989; Miller & Fellbaum, 1991)、そして折田・小林 (2011c) で報告した複雑な意味構造を持ち言語間の違いが大きく (Källkvist, 1999)、他の語類よりも意味の幅の広い (Gentner, 1981, 1982) 動詞では、名詞に見られた特徴は検出されな

かった。これらのことから、クラスター数、クラスターサイズ、またばらつき度からみた心内辞書の発達・変容に関しては、名詞あるいは名詞的特性を持つ語群以外は母語話者的な構造に発達・変容する可能性があると言える。一方、心内辞書内の意味的クラスタリング構造については、母語話者と日本人英語話者の間に明確な差異が確認され、母語話者では6つ、日本人英語話者では4つのクラスターに構造化されていることが判明した。また、両群が生成したこれらのクラスターは、部分的には類似する構造を持ちながらも同一のものはないこともわかった。以上のまとめを踏まえて、両群の心内辞書内意味的クラスタリング構造の質的違いの特徴を、6つのクラスターからなる母語話者の構造をベースラインとして、表4a・表4b及び図1・図2の各群の群デンドログラムの組成を検討しながら考察する。

まず、母語話者では「程度」「時間」という2つの意味領域から構成されるクラスター（10語）が形成されていることである。母語話者では「程度」に関わる *nothing-very-all* と「時間」に関わる *also-early-next-while-then-already-ago* の語群は、明確な2つの下位領域に分化しながらも密接に結び付いている。日本人英語話者では、「程度」（5語）は独立したクラスターを形成し、「時間」（6語）は「歴史」（2語）と結合して一つのクラスターを形成しており、両群の構造化の様相は異なる。Alsoは母語話者では「時間」クラスターの構成要素であるが、日本人英語話者では「程度」の一部を構成している。また、日本人英語話者では *lot* が「程度」の一部をなしており *nothing-all* と類似度が高い。母語話者では *lot* は、語群 *business-shop-street-place-area* と類似度が高く、この語に関する両群間の心内辞書内の構造化の違いは際立っている。このように、「程度」「時間」クラスターにおける両群間の相違は少なくない。

母語話者の生成した2つ目のクラスターは、2語からなる「状態」(*open-close*)と6語からなる「場所」(*lot-business-shop-street-place-area*)から構成されている。母語話者では「場所」を構成する *business* が、日本人英語話者では「力」「人間」「認知」「物質」クラスターの構成要素となり、語群 *social-war-power-law-country* と類似度が高い。*Business* は母語話者では仕事や事業の行われる場所という意味成分が優勢に活性化され、日本人英語話者では仕事や事業それ自体として活性化される傾向がうかがえる。*Open-close* は、日本人では「状態」「動き」「場所」「自然」クラスター（13語）の一部を形成し、*walk-stand-arrive* という3語の動詞からなる語群と類似度が高い。母語話者では *open-close* は「場所」に関する6語 (*lot-business-shop-street-place-area*) と密接に結び付いていることと対照的である。しかし、両群とも *open* と *close* は本実験で用いた50語における2語間の類似度が最も高く、母語、第二言語の違いに関わらず反意語の関係により心内辞書内に内在化されている。上に述べたように、*lot* は母語話者では「場所」の、また日本人英語話者では「程度」の一部を形成し、活性化される優勢な語彙知識に違いがある。

母語話者が生成した3つ目のクラスターは「力」「歴史」クラスターである。「力」は *law-president-police-doctor*（4語）から、「歴史」は *country-war-power-history-century-arm*（6語）からなる。日本人英語話者では *war-power-law* は *social* と、*president-police-doctor* は *person-boy* と類似度が高い。*War-power* が *history-century* と、また *president-police-doctor* が *law* と強く結び付く母語話者と対照的である。日本人ではまた、*arm* は *form-figure* と最も密接に結び付いている。つまり、本実験で用いた語群において *arm* に関して語彙間ネットワークが活性化される意味的クラスタリングは、母語話者では「武器」という意味組成であり、日本人では「腕」であった。多義性を

持つ語彙項目の優勢な意味組成が母語話者と第二言語話者の間で異なる場合があることが示唆される。*Country* は、母語話者では *war-power-history-century* と、日本人では *social-war-power-law* と結び付き、「国」という意味組成を同様に活性化させてもクラスタリングする対象が両群間で異なっている。そして、母語話者が「力」と「歴史」を最近接する意味の場 (semantic field = 共通の意味属性 (semantic properties) を持つ群化された心内辞書内の語彙項目群) として構造化しているのに対し、日本人英語話者では「歴史」と「時間」が一つのクラスターを形成し、「力」は「人間」「認知」「物質」と結び付いている。このように、「力」「歴史」語群における心内辞書内の内部構造やその活性化の様相、また意味的クラスタリングの基盤となる個々の語彙項目における優勢な意味組成は英語母語話者と日本人英語話者で大きく異なる。

「動き」クラスターは母語話者が生成した4つ目のクラスターで、*walk-step-stand-arrive* の4語の動詞から構成される。一方、日本人英語話者ではこの4語は「状態」「動き」「場所」「自然」クラスター (13語) の一部を形成している。この4語間のクラスタリング構造は両群で異なり、たとえば *walk* について、母語話者では *step* と最も類似度が高く日本人英語話者では *stand* と類似度が高い。また、母語話者では *walk, step, stand, arrive* はすべて動詞としてクラスタリングされているのに対し、日本人英語話者では *walk, stand, arrive* の3語については母語話者と同様であるが、*step* は *shop-street-place-area* とクラスタリングされ名詞の意味組成が活性化されている。ステップ (*step*) は、日本語の中に英語からの借用語 (『広辞苑』(第六版) で確認) として日常的に使われており、足取り、歩調、昇降口の階段といった名詞の意味組成でカタカナ語として定着していることが関係している可能性が高い。名詞や動詞が混在し、当該の単語が動詞を中心としても、また名詞を中心としてもクラスタリングされることがありうる場合、母語で形成されたものが優位に活性化されることがあると示唆される。母語話者と比較して第二言語話者は、「緩やかな概念的反応 (loose conceptual responses)」を生成する比率が高く (Fitzpatrick, 2006, p. 143)、動詞と名詞の別を明確に意識せずに *step* が *shop-street-place-area* とクラスタリングされたことも考えられる。折田・小林 (2011a) が指摘したように、母語における語彙獲得の強固な影響は第二言語の語彙獲得における個々の単語レベルの学び直しの困難さや意味的クラスタリング構造の再構築の難しさに深く関係し、「言語獲得の初期において確立された語彙レベルのつながりは、たとえ学習者が流暢な第二言語話者になったとしても言語使用の状況によっては作動し続ける」(Kroll & Tokowicz, 2001, p. 52) ことを示す例と言える。

母語話者が生成した5つ目のクラスターは「物質」「自然」クラスターである。「物質」に関する2語 (*matter-form*) と「自然」に関する3語 (*clear-natural-air*) から構成された。一方、日本人英語話者の場合は *matter* と *form* は24語からなる「力」「人間」「認知」「物質」クラスターの一部をなした。そこでは *matter* は *reason* と結び付き、*form* は *figure* と結び付いている。このように、母語話者では *matter* は物質という意味組成が活性化され、日本人英語話者では事柄・問題という意味組成が活性化されている。心内辞書にアクセスする語彙間意味処理において、母語話者は単語の具象的な意味を第一義としてクラスタリングする一方、日本人英語話者は抽象的な意味を第一義としてクラスタリングする場合があると言える。さらに、母語話者では「物質」と「自然」は一つのクラスターを形成し、その中では *matter* と *nature* は密接に結び付いている。日本人英語話者においては、*matter* は「力」「人間」「認知」「物質」クラスターに入り、*nature* は「状態」「動き」「場所」「自然」クラスターの中で構造化されている。これは、英語母語話者には (西洋的な価値観のもとでは)

自然は科学の対象となりえても、日本人には一般に自然は科学の対象となりえないことに起因する可能性が高い<sup>1)</sup>。つまり、*matter* と *nature* のクラスタリングにおける母語話者と日本人の違いは語彙知識や母語の影響によるものとするよりも、むしろ被験者が属する社会や文化を通して保持するようになった自然観や世界認識に基づく違いに関係すると指摘できる。

そして、英語母語話者が生成した6つ目のクラスターは「人間」「認知」クラスターである。サブクラスター「人間」は *our-figure-person-boy* の4語から、またサブクラスター「認知」は *keep-find-social-help-cry-dream-reason-understand-believe* の9語から構成されている。一方、日本人英語話者の場合は、これら13語はすべて「力」「人間」「認知」「物質」クラスター(24語)に入ったが、その構造化の様相は母語話者とは著しく異なっている。具体的に4点指摘する。まず *our* は、母語話者では *figure-person-boy* と類似度が高く、明らかに「人間」に関する単語群と結び付いている。しかしながら、日本人英語話者の場合は *our* は *social-war-power-law-country-business-president-police-doctor-person-boy* という11もの語彙項目から構成されるサブクラスターと類似度が高く、*our* を含むこれらの語群が明確に「人間」という意味領域を構成しているとは言えない。日本人が生成した *our* を含む12語からなる語群は、Fitzpatrick (2006) の指摘する「緩やかな概念的反応」による結果としてクラスタリングされており、母語話者ほど明確に「人間」に関する意味の場としてクラスター化されていない。2点目に、母語話者では *figure* は *person-boy* と結び付き、ひと(の姿)という意味組成が活性化され、一方日本人英語話者では *figure* は *form* と最も類似度が高く、形状という意味組成が活性化されており、活性化される心内辞書内の優勢な意味はここでも両群間で異なった。3点目に、上述したように母語話者では *person-boy* は *figure* と類似度が高かったが、日本人英語話者では *president-police-doctor* と類似度が高かった。母語話者が意味組成として純粋に人間に関わるもののみを基盤にクラスタリングするのに対し、日本人英語話者では「力」クラスターに入る職業名とクラスタリングするという違いがある。そして4点目として、両群間の *reason* のとらえ方の違いがある。*Reason* は、母語話者では *understand-believe* と最も類似度が高く動詞としてクラスタリングされ、日本人英語話者では *matter* と最も類似度が高く名詞としてクラスタリングされている。母語であるか第二言語であるかの違いにより、心内辞書へのアクセスの状況によっては活性化される当該語彙項目の統語知識に違いが発生することもあると言える。

## 6. 結論

本研究は、英語母語話者と日本人英語話者の心内辞書内の意味的クラスタリング構造について、名詞や動詞といった特定の語類に限定しない語群における違いを明らかにすることを目指した。そのために、語類の混在する単語群を用いた単語仕分け課題結果について両群の意味的クラスタリング構造の計量的及び質的相違を解析した。解析結果から、母語話者の方が日本人英語話者よりも生成する平均クラスター数が多く、平均クラスターサイズが小さいという傾向、またクラスター数やクラスターサイズのばらつき度、さらには個人デンドログラムのばらつき度は母語話者の方が小さいという傾向は見られたが、いずれも有意な差異ではなかった。このことから、少なくとも計量的な差異という観点からは、母語話者の方が第二言語話者よりも心内辞書内の単語間の意味的関連性を明確に意識化しているとは必ずしも言えないと指摘できる。両群間の違いはむしろ意味的クラスタリングの様相、つまりクラスター構造における質的な違いにある。両群の群デンドログラムの比較から、クラスターを

構成する単語群には「程度」「時間」「状態」「場所」など共通のものが確認できた。しかし、詳細に分析するとクラスターの構造化の様相は両群間で大きく異なっていた。そしてその違いは、実験課題に取り組む際に活性化される心内辞書内の個々の語彙項目の優勢な意味組成の違いに起因する場合が多いことが判明した。母語話者は *business* を「場所」語群の中でクラスタリングするが、日本人英語話者は仕事や事業それ自体の意味を基盤に意味的に関連する語群の中でクラスタリングし、また *arm* を母語話者は「武器」の意味で、日本人英語話者は「腕」の意味でクラスタリングしたことがその例である。加えて、*step* を母語話者は意味的に関連する動詞群とクラスタリングし、日本人英語話者は名詞群とクラスタリングした。このように、心内辞書内で優勢に活性化される語類が両群間で異なる場合もあることが明らかになった。さらに、母語話者は *nature* と *matter* を同一のクラスター内で内在化したのに対し、日本人英語話者では *nature* と *matter* は別個のクラスターに入った。これは語彙知識の違いや母語の影響よりも、むしろ自然観や科学観の違いに関係すると言える。以上のことから、母語話者と日本人英語話者の心内辞書内の意味的クラスタリング構造の質的な違いの要因は多様であるが、最も優勢に内在化され、また活性化されやすい個々の単語の意味や統語知識などが、日本人英語話者の心内辞書において母語話者的なものに全て書き換えられるわけではないことが示唆される。

## 謝辞

本稿は、第54回熊本大学英文学会（2010年11月20日、熊本大学）において折田が行った口頭発表「心内辞書内のネットワーク構造—Sorting tasks を用いた母語話者と第二言語話者の違いの解明」を加筆修正したものである。発表に対して、吉田正憲（熊本大学名誉教授）・堀正広（熊本学園大学）の両氏から貴重なご助言を賜った。また、小林美代子（社会文化科学研究科）・永尾悟（文学部）の両氏から本稿全般について有益なご指摘を頂いた。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）「日本人英語学習者のメンタルレキシコン：意味的クラスタリング構造の解明」（課題番号21520444：研究代表者折田充）、同若手研究（B）「大規模ランダム行列を用いたモデル選択と機械学習理論」（課題番号20700258：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」（23-共研-2056：研究代表者小林景）の助成を受けたものである。

## 注

- 1) 日本文学者の自然観について、たとえば夏目漱石に関して吉田（2010）は、「彼は「日本人は山川崇拝と云ふべき国民性故」、これが我々の関心を惹く問題だと考えたようだが、日本人の自然への感情が西洋人のそれと同じなのかどうか、この点は少し検討してみる必要がある」（p. 85）と指摘している。一般の日本人が抱く自然観について本川（2010）は、1953年以来5年ごとに継続的に実施されている「日本人の国民性調査」（統計数理研究所）のデータに基づき、「自然との共存」か「自然の支配」かという国際アンケートでは、日本は他国と比べて最も「自然との共存」が多く、これが日本人の自然観の特徴となっている。（中略）日本人も高度成長期の後半には、経済の成功につられて欧米型の自然観に近づいた時期もあったがその後従来の姿勢に戻ったということができよう。」としている。また深谷・舛田（2006）は、主要全国紙4紙（データ総件数3,681件）に対するスク립ト分析により、日本人が共有する可能性の高い自然観の抽出を試みている。そのまとめとして、「現

代日本人が自然を自律的存在であるべきだとし、人間も自然も互いに自身の立てた自己規範によって生きていると捉えている点は従来「日本人の自然観」に関していわれてきたことと大まかには一致する。なぜ *nature* の訳語に自然という語があてられたかが明確ではないことは前述の通りだが、当時の日本人が現代日本人と同じように「おのずからしかる」という自律の性質を草木に見出し、*nature* を自然と訳したという推測もできる。」(p. 28) と述べている。これらの研究から、自然は人間が支配するものでなくむしろ共存するものであり、また自律的存在として尊重するものであるとする自然観は、少なくとも一般の日本人には共有されていると言える。

### 参考文献

- 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」『九州英語教育学会紀要』, 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.
- 小林景・折田充 (2009). 「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, 9月9日.
- 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』. 共立出版.
- 新村出 (編) (2008). 『広辞苑』 (第六版). 岩波書店.
- 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』. 大学英語教育学会.
- 竹村彰通 (1991). 『現代数理統計学』. 創文社.
- 深谷昌弘・榎田晶子 (2006). 「人々の意味世界から読み解く日本人の自然観」 (総合政策学ワーキングペーパーシリーズ No. 96). 慶應義塾大学大学院政策・メディア研究科.
- 本川裕 (2010). 「日本人の自然観の長期推移」. Retrieved 28 December 2011, from <http://www2.ttcn.ne.jp/honkawa/4250.html>
- 吉田正憲 (2010). 「若き漱石の英国「自然詩」研究」坂元昌樹・田中雄次・西楨侖・福澤清 (編)『漱石文学の水脈』, 64-98. 思文閣出版.
- Fitzpatrick, T. (2006). Habits and rabbits: Word associations and the L2 lexicon. *EUROSLA Yearbook*, 6, 121-145.
- Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory*, 4, 161-178.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum.
- Gross, D., Fischer, U., & Miller, G. A. (1989). The organization of adjectival meanings. *Journal of Memory and Language*, 28, 92-106.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Kikuchi, M., Yamamoto, M., Yoshimura, M., Yabuuchi, S., & Tanimura, M. (2001). Assessing the hierarchical structure of L2 mental lexicon: An analysis of categorical and non-categorical word association. *JACET*

- Bulletin*, 34, 27-35.
- Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. L. Nicol (Ed.), *One mind, two languages* (pp. 48-71). Malden, MA: Blackwell.
- Källkvist, M. (1999). *Form-class and task-type effects in learner English: A study of advanced Swedish learners*. Lund, Sweden: Lund University Press.
- Meara, P. (1978). Learners' word associations in French. *Interlanguage Studies Bulletin*, 2, 192-211.
- Meara, P. (1983). Word associations in a foreign language. Retrieved 23 June 2009, from <http://www.lognostics.co.uk/vlibrary/meara1983.pdf>
- Meara, P., & Schur, E. (2002). Random association networks: A baseline measure of lexical complexity. *British Studies in Applied Linguistics*, 17, 169-182.
- Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition*, 41, 197-229.
- Nissen, H. B., & Henriksen, B. (2006). Word class influence on word association test results. *International Journal of Applied Linguistics*, 16, 389-408.
- 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.
- Postman, L., & Keppel, G. (1970). *Norms of word associations*. New York: Academic Press.
- Riegel, K. F., & Zivian, I. W. M. (1972). A study of inter- and intralingual associations in English and German. *Language Learning*, 22, 51-63.
- Schmitt, N. (1998). Quantifying word association responses: What is native-like?. *System*, 26, 389-402.
- Szalay, L. B., & Deese, J. (1978). *Subjective meaning and culture: An assessment through word associations*. Hillsdale, NJ: Lawrence Erlbaum.
- Söderman, T. (1989). Word associations of foreign language learners and native speakers: A shift in response type and its relevance for a theory of lexical development. *Scandinavian Working Papers on Bilingualism*, 8, 114-121.
- Söderman, T. (1993). Word associations of foreign language learners and native speakers: The phenomenon of a shift in response type and its relevance for lexical development. In H. Ringbom (Ed.), *Near-native proficiency in English* (pp. 91-182). *English Department Publications 2*, Åbo Akademi University, Åbo, Finland.
- Yokokawa, H., Yabuuchi, S., Kadota, S., Nakanishi, Y., & Noro, T. (2002). Lexical networks in L2 mental lexicon: Evidence from a word-association task for Japanese EFL learners. *Language Education & Technology*, 39, 21-39.
- Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization?. *Second Language Research*, 23, 13-153.

## APPENDIX

## 単語仕分け課題\*

封筒の中に、別々の英単語が印刷された、50枚のカードが入っています。これらのカードを、自分が思う、意味の上で関連しているグループに分けてください。どのようにグループ分けしてもかまいません。(ただし、名詞や動詞といった、品詞では分けないでください。)単語によっては、他の単語とグループを作らないと思うものもあるかもしれません。そういった単語は、そのままにしておいてください。最終的にいくつのグループを作ってもかまいません。制限時間は20分です。

\* 英語母語話者へは同じ内容を英語で示した。

Mental Lexicons of Native and Non-Native English Speakers:  
Semantic Clustering of Lexical Items from Mixed Word Classes

MITSURU ORITA and KEI KOBAYASHI

This study addresses the question of whether or not the semantic clustering of lexical items in L1 and L2 mental lexicons is structurally different either quantitatively or qualitatively. A free sorting task using a set of 50 high-frequency English words taken from mixed word classes (30 nouns, five verbs, six adjectives, six adverbs, one pronoun, one determiner and one conjunction) were given to 28 native speakers of English (NS) and 28 advanced-level Japanese speakers of English (NNS). The experiment found that there were no statistically significant differences between the two groups in mean cluster number, mean cluster size, and variability in mean individual dendrogram distance. However, cluster analysis revealed that L1 lexical organisation differed from its L2 counterpart qualitatively. Both groups similarly produced the clusters of POWER, TIME, PLACE, MATERIAL, NATURE, etc., but the connected lexical items were differently structured. For example, NS linked *arm* to *country*, *war*, *power* and others, but NNS connected *arm* with *form* and *figure*. The word *lot* was linked to *business*, *shop*, *street*, *place* and *area* by NS, whereas *lot* was connected with *nothing* and *all* in the NNS results. L1 and L2 difference was also detected in the way the two groups conceptualised “nature and materials” in that NS put *nature* and *matter* in the same MATERIAL and NATURE cluster, while NNS placed *nature* and *matter* in separate clusters. It is concluded that a significant portion of the semantic properties of L2 lexical knowledge and structure are difficult for non-native speakers to relearn and restructure into native-like ones.