

漢字四字成語の受容とその延命

砂岡和子（早稲田大学政治経済学術院）[†]

羅鳳珠（台湾元智大学中国語文学系）^{††}

王雷（北京大学計算語言学研究所, 北京大学英語系）^{†††}

姜柄圭（韓国西江大学中文系）^{††††}

松崎実夏（千駄ヶ谷日本語教育研究所）^{†††††}

Adoption of the Chinese Four-Character Idioms and their Surviving in the Contemporary Written Japanese

Sunaoka Kazuko (School of Political Science and Economics, Waseda University)

Lo Feng Ju (Dept. of Chinese Linguistics & Literature, Yuan Ze University)

Wang Lei (Dept. of Institute of Computational Linguistics & English, Peking University)

Kang Byeong Kwu (Dept. of Chinese Culture, Sogang University)

Matuzaki Mika (Sendagaya Japanese Institute)

1. 概要

筆者らは国際共同プロジェクト事業の一環として、中台韓日英マルチリンガル成語辞典を編纂している¹。中台韓日それぞれの地域を代表する現代語コーパスに基づき、最大公約数の常用成語を選定後、言語文化情報を付与し、英語訳を加え、各言語間のマルチ検索と各地域のコーパスとの閲覧リンク可能な Web 版成語辞典の公開を目指している。本文は主に日本研究チームが、日本国立国語研究所「現代日本語書き言葉均衡コーパス (BCCWJ)」の検索ツール「中納言」を利用し日中同形成語の抽出作業を行った作業経過に基づき、現代日本語に於ける漢字四字成語の受容とその延命状況について報告する。

2. 中日韓英マルチリンガル成語辞典の構築

四字成語は東アジア漢字圏の智慧と文化を凝縮する表現形式のひとつであり、簡潔で口調の良い言語形式と相まって古典文学から新聞記事まで旺盛な生命力を維持している。現代中国語で常用される成語は 2000-3000 語に上るとされる[劉長征他 (2007)]。長らく漢字圏にあった韓国、日本では四字成語を含め漢字語句の使用頻度が低下しているとはいえ、故事成語やことわざの愛好に見られるように、根強い延命力がある[砂岡和子他 (2011b)]。

[†]ksunaoka@gmail.com ^{††}gef.julo@saturn.yzu.edu.tw ^{†††}wangleics@pku.edu.cn

^{††††}kg4335@gmail.com ^{†††††}songqi.shixia@gmail.com

¹ 台湾蔣經國國際學術交流基金會「歴代語言知識庫建置計畫」研究代表：台湾元智大學羅鳳珠〔網路展書讀〕<http://cls.hs.yzu.edu.tw> に概要公開

中台韓日英マルチリンガル成語辞典は、中国古典籍資源を母体とする東アジア歴代語言データベースを基に研究や教育に共同利用する目的で構築中の多国籍プロジェクトである。

成語辞典の編纂作業は以下4段階で進行する[砂岡和子, 羅鳳珠 (2011a) 参照]。

[第一段階] 成語辞典の母資源となる中台成語資源の構築。大陸中国の成語母集団は北京大学計算言語学研究所開発の「成語知識庫」で、同所の「現代漢語語法信息詞典」所収の成語約5000条を核に、「人民日報」ならびに『成語大全』『学生成語詞典』など中国出版の代表的成語辞書5冊から3万余項目の現代中国語成語を収集し、これに頻度情報、語法属性、品詞、異形、類語、反義語、典故、釈義など16種の言語属性を記したコーパスである[A 北大三萬詞簡體成語庫] [2010 王雷] [図1]。このうちもっともポピュラーな中国成語1000語は書籍出版された[A' ' 北大1000] [2012 王雷]。これを台湾で使われる繁体字に変換校正し[A' 北大三萬詞繁體成語庫]、台湾中央研究院「平衡語料庫[バランスコーパス]」[B 台湾中央研究院平衡語料成語庫]、ならびに台湾教育部公布の成語データから台湾香港地域の常用成語約1万語をリストアップした[C 台湾教育部成語詞條]。最後に[A] [B] [C]資源間で重複率の高い成語約3000語を選定し、これ中国語母集団の中核的成語資源データとした[D 三千核心成語]²。

[第二段階]は漢字受容圏の成語資源構築と、母集団資源との共通成語の選出作業である。日韓成語の選定に当たっては、それぞれの母語の表現形式や受容条件を反映するよう、各地域を代表する現代語コーパスから漢字成語を抽出し、上掲[D]資源との照合作業を進めた。日本語の常用成語は名古屋大学佐藤理史研究室作成「基本慣用句五種対照表」所収の成語選定基準を参照し、日本国立国語研究所「現代日本語書き言葉均衡コーパス(BCCWJ)」[E]用のオンラインコンコーダンスである「少納言」「中納言」を利用し、上掲[D 三千核心成語]とのマッチング作業に当たった。韓国は国立国語院「世宗コーパス」から韓国語の成語を抽出する[F]。

[第三段階]は[D]の成語に典故、例文などの百科知識、およびその使用頻度、分布密度、難度レベル、語法属性、品詞機能などデジタル言語情報を付与する。台湾元智大学[網路展書讀]にはすでに「三国演義」「水滸伝」「紅樓夢」など歴代文学資料が構築されており、これら白話小説中の成語を抽出してマルチ辞書に加える。図2に成語知識庫構築図を示す。

[第四段階]で英語、韓国語、日本語の翻訳をつけ、各言語間のマルチ検索と各地域コーパスとリンクして閲覧可能なWeb版成語辞典として公開する。

3. 現代日本語常用漢字成語の選定

既述のように日本チームは日本語常用成語の選定あたり、日本語基本慣用句データベースとして定評のある名古屋大学「基本慣用句五種対照表(以後「基本慣用句表」)」所収の成語を被検出語とした。BCCWJコーパスには成語を抽出するための言語標識が無い。BCCWJのDVD媒体でマッチングしようと試行錯誤したが、異形の多さに頓挫した。

「基本慣用句表」は1982—2006年にかけ日本で出版された辞書から慣用句を抽出し、「慣

² [劉長征他(2007)]に拠ると現代中国語の常用成語数は2000-3000とされる。

序号	汉语	类别	全拼	释义	反义	直译	释义	出处	例句	频率	来源
13305	艰苦卓绝	1	jiǎnkǔ zhuōjué	艰苦: 艰难困苦; 绝: 断绝。形容处境极其困难, 斗争极其艰苦。	安逸	to work arduously	艰苦: 艰难困苦; 绝: 断绝。形容处境极其困难, 斗争极其艰苦。	《中国革命和战争的总路线》: “我们要有‘艰苦卓绝’的精神。”	33080	yu1	
37816	自力更生	1	zìlì gēngshēng	自力: 自己的力量; 更生: 重新获得生命。	依赖	regeneration through self	自力: 自己的力量; 更生: 重新获得生命。	《中共中央关于中国农村的社会主义》: “我们要有‘自力更生’的精神。”	30881	yu1	
28002	实事求是	1	shíshì qiúshì	实: 实际; 是: 客观; 求: 研究; 是: 符合实际。	主观	be practical and real	实: 实际; 是: 客观; 求: 研究; 是: 符合实际。	《汉书》: “实事求是。”	26373	yu1	
21005	千方百计	1	qiān fāng bǎi fǎ	千: 形容多; 方: 方法; 百: 形容多; 计: 计谋。	简单	by every possible way	千: 形容多; 方: 方法; 百: 形容多; 计: 计谋。	《宋史》: “千方百计。”	22341	yu1	
22276	一心一意	1	yìxīn yìyì	一: 一个; 心: 心思; 一: 一个; 意: 意思。	三心二意	with all one's heart	一: 一个; 心: 心思; 一: 一个; 意: 意思。	《晋书》: “一心一意。”	20439	yu1	
13288	坚定不移	1	jiǎndìng bù yí	坚: 坚定; 定: 稳定; 移: 移动。	动摇	to insist on firmly	坚: 坚定; 定: 稳定; 移: 移动。	《宋史》: “坚定不移。”	18005	yu1	
33841	因地制宜	1	yīn dì zhì yì	因: 根据; 地: 地方; 制: 制定; 宜: 适宜。	一成不变	to take actions	因: 根据; 地: 地方; 制: 制定; 宜: 适宜。	《汉书》: “因地制宜。”	14410	yu1	
6826	独立自主	1	dúzì lìzhù	独: 单独; 立: 建立; 自: 自己; 主: 自主。	依赖	be one's own	独: 单独; 立: 建立; 自: 自己; 主: 自主。	《毛泽东选集》: “独立自主。”	11271	yu1	
3812	成千上万	1	chéngqiān shàngwàn	千: 形容多; 上: 到; 万: 形容多。	寥寥无几	thousands upon thousands	千: 形容多; 上: 到; 万: 形容多。	《汉书》: “成千上万。”	8888	yu1	
13284	坚持不懈	1	jiǎn chí bù xiè	坚: 坚定; 持: 坚持; 懈: 懈怠; 不: 不; 懈: 懈怠。	半途而废	to persist to the end	坚: 坚定; 持: 坚持; 懈: 懈怠; 不: 不; 懈: 懈怠。	《汉书》: “坚持不懈。”	8887	yu1	
38280	战无不胜	1	zhàn wú bù shèng	战: 战斗; 无: 没有; 不: 不; 胜: 胜利。	一触即溃	invincible	战: 战斗; 无: 没有; 不: 不; 胜: 胜利。	《汉书》: “战无不胜。”	8820	yu1	
11285	轰轰烈烈	1	hōnghōng lièliè	轰: 轰响; 烈: 强烈。	冷冷清清	be dynamic and vivid	轰: 轰响; 烈: 强烈。	《汉书》: “轰轰烈烈。”	6154	yu1	
21273	前所未有	1	qiǎn wèi yǒu	前: 以前; 未: 没有; 有: 有。	史无前例	unprecedented	前: 以前; 未: 没有; 有: 有。	《汉书》: “前所未有。”	6878	yu1	
11887	记忆犹新	1	jiùyì yǒu xīn	忆: 回忆; 犹: 好像; 新: 新鲜。	恍如隔世	be great overjoyed	忆: 回忆; 犹: 好像; 新: 新鲜。	《汉书》: “记忆犹新。”	6884	yu1	
28806	滔天罪行	1	tāotiān zuìxíng	滔: 滔天; 天: 天; 罪: 罪恶; 行: 行为。	微不足道	One's crimes cry aloud	滔: 滔天; 天: 天; 罪: 罪恶; 行: 行为。	《汉书》: “滔天罪行。”	6487	yu1	
8232	丰富多彩	1	fēngfēng cǎicǎi	丰: 丰富; 多: 多; 彩: 色彩; 采: 色彩。	单调乏味	be rich and varied	丰: 丰富; 多: 多; 彩: 色彩; 采: 色彩。	《汉书》: “丰富多彩。”	5300	yu1	
3274	朝气蓬勃	1	zhāoqì péngbō	朝: 早晨; 气: 气息; 蓬: 蓬; 勃: 勃。	暮气沉沉	be full of vigor	朝: 早晨; 气: 气息; 蓬: 蓬; 勃: 勃。	《汉书》: “朝气蓬勃。”	5811	yu1	
18820	目光远大	1	mùguāng yuǎndà	目: 眼睛; 光: 光线; 远: 远; 大: 大。	鼠目寸光	to open one's eyes	目: 眼睛; 光: 光线; 远: 远; 大: 大。	《汉书》: “目光远大。”	5641	yu1	
33879	阴谋诡计	1	yīnyōu guǐjì	阴: 阴险; 谋: 计谋; 诡: 诡计; 计: 计谋。	光明正大	intrigues and plots	阴: 阴险; 谋: 计谋; 诡: 诡计; 计: 计谋。	《汉书》: “阴谋诡计。”	5385	yu1	
33486	以身作则	1	yǐshēn zuòfàn	以: 用; 身: 身体; 作: 做; 范: 模范。	言行不一	to set an example	以: 用; 身: 身体; 作: 做; 范: 模范。	《汉书》: “以身作则。”	5239	yu1	
30794	一如既往	1	yīwú yǐzhàng	一: 一个; 如: 像; 往: 以前; 往: 以前。	一反常态	to continue as always	一: 一个; 如: 像; 往: 以前; 往: 以前。	《汉书》: “一如既往。”	5108	yu1	
38998	针锋相对	1	zhēnfēng xiāngduì	针: 针; 锋: 锋芒; 相: 相对; 对: 对。	冰炭不相容	in action, step by step	针: 针; 锋: 锋芒; 相: 相对; 对: 对。	《汉书》: “针锋相对。”	4988	yu1	
24408	声势浩大	1	shēngshì hàodà	声: 声音; 势: 气势; 浩: 浩; 大: 大。	悄无声息	powerful and dinant	声: 声音; 势: 气势; 浩: 浩; 大: 大。	《汉书》: “声势浩大。”	4937	yu1	
37448	卓有成效	1	zhuó yǒu chéngxiào	卓: 卓越; 有: 有; 效: 效果; 成: 成就。	华而不实	with outstanding results	卓: 卓越; 有: 有; 效: 效果; 成: 成就。	《汉书》: “卓有成效。”	4805	yu1	
2688	不屈不挠	1	bùqū bùniǎo	不: 不; 屈: 屈服; 不: 不; 挠: 挠。	畏葸不前	indomitable	不: 不; 屈: 屈服; 不: 不; 挠: 挠。	《汉书》: “不屈不挠。”	4685	yu1	
726	百家争鸣	1	bǎijiā zhēngmíng	百: 百; 家: 家; 争: 争; 鸣: 鸣。	一统天下	contention of a hundred schools	百: 百; 家: 家; 争: 争; 鸣: 鸣。	《汉书》: “百家争鸣。”	4484	yu1	
33547	再接再厉	1	zài jī zài lì	再: 再; 励: 励; 再: 再; 厉: 厉。	一蹶不振	to make persistent	再: 再; 励: 励; 再: 再; 厉: 厉。	《汉书》: “再接再厉。”	4438	yu1	
30077	欣欣向荣	1	xīnxīn xiàngróng	欣: 高兴; 欣: 高兴; 向: 向; 荣: 繁荣。	门可罗雀	flowers in spring	欣: 高兴; 欣: 高兴; 向: 向; 荣: 繁荣。	《汉书》: “欣欣向荣。”	4356	yu1	
5820	当之无愧	1	dāng wú yì zhī	当: 应当; 无: 没有; 愧: 惭愧; 之: 之。	愧不敢当	deserve	当: 应当; 无: 没有; 愧: 惭愧; 之: 之。	《汉书》: “当之无愧。”	4196	yu1	

图 1 「北大三萬詞簡體成語庫」部分

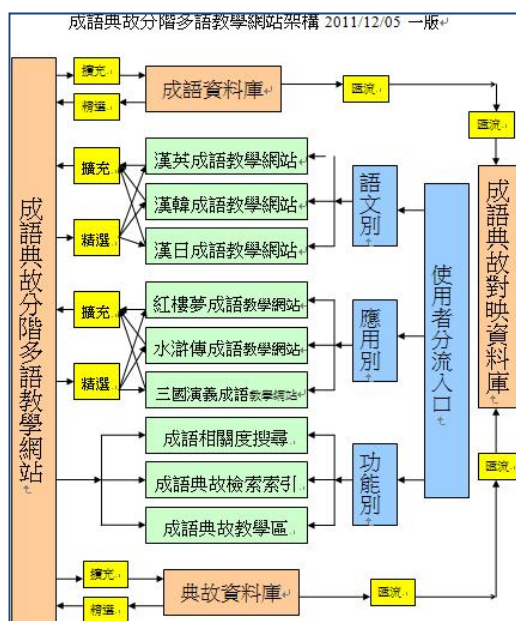


图 2 成語コーパス構築図

用句 1918 語」「ことわざ 344 語」「四字熟語 265 語」「故事成語 63 語」「その他不明 1038 語」に分類、計 3628 の慣用的表現を対照表としたデータベースである。大半は“我に返る”“わき目も振らず”“目を丸くする”“弱音を吐く”など和語の「慣用句」で、漢字成語よりこちらのほうが多い[図 3 日本名大和語辞参照]。

我々はまず「基本慣用句」中、四字漢語を多く含む「四字熟語」と「故事成語」の計 328 語を被検索対象語とし、「少納言」「中納言」で上掲[D 三千核心成語]とのマッチングを開始した。[D]には中台母集団成語データベース[A] [B] [C]間で重複率の高い成語約 3000 語に対し、重複回数に応じ 4 レベルに分けてある。最初に[D]中の重複回数 4-3 回の上位中国成語 2005 条について BCCWJ でマッチングした結果、ヒットした日本語は僅か 211 語、全体の約 10%に留まった。

ヒット率低迷の原因は、日本語の漢字四字成語が少ないこと、日中成語の使用文字や語形の異なるためである。成語に限らず日本語の四字漢語は二字+二字の緩い連合形式から成るものが全体の 90 パーセント以上を占め、一語としての凝縮力が弱い³。中国の成語は“自相矛盾”“完璧歸趙”“画蛇添足”“杞人憂天”“不可思議”のように四字が典型であるが、日本語ではそれぞれ“矛盾”“完璧”“蛇足”“杞憂”“不思議”のように基幹二字もしくは三字となる。四字でも“三顧の礼(中国語原文は“茅廬三顧)”“背水の陣(同“背水一戰)”)のように平仮名混じりや、“枚挙にいとまがない(中国語は“不勝枚舉(不可枚

³ [野村雅昭(1974)]の調査は約 19 万新聞語彙の四字漢語全体についてであるが、日本語四字成語にも当てはまる特徴と言える。

舉、不遑枚舉とも)”など日本語の統語法に則して形を変えた成語例は、それこそ枚挙にいとまがない。このため日本語の四字成語は形式上、異形や類語が極めて多く[1982 宮地裕)、マッチングの障害となる。韓国語の成語も同様の傾向を示す[2011 姜柄圭]。

日中同形で抽出できる成語数には限界があるため、次に「基本慣用句」の日本語慣用句を中国語に翻訳し、中台成語データベースの [A] [B] [C] をダイレクトにマッチングした。成語の中訳に当たって各種日中成語辞書や、北京大学が開発する類義語検索サイト[関連 URL 参照]で類似表現から適訳を検出した。すでに BCCWJ とのマッチングで抽出済みの 211 語と併せ、最終的に「基本慣用句」所収全 3628 句のうち 1108 語の中日成語ペアを選定した。

4. 「少納言」「中納言」による日本語常用漢字成語の抽出

われわれは BCCWJ 現代日本語コーパスを他言語資源とのマッチングデータとしても利用し、「少納言」「中納言」を多言語検索ツールに利用した。以下、「少納言」による日中漢字成語検索と、BCCWJ によるマッチングの過程をチャートにまとめる⁴。実際に多用したのは「少納言」より高度な検索機能を備える「中納言」のほうで、「少納言」の 500 項までの検索結果表示制限も無く、複雑な結果を得ることができる。ただ「中納言」は日本語を理解しない外国人には利用申請手続きが難しいため、アクセスが簡単な「少納言」での利用で代表する。

4-1 前処理

マッチング作業用に整理加工したデータを準備する。中台成語資源[D 三千核心成語]は、データを中国大陆の簡体字と台湾香港地区の繁体字で併記し、どちらの字体でもできる。BCCWJ データ、「少納言」「中納言」とも繁体字であれば検索成功率が高いため、簡体字成語の場合は[D]で対応する繁体字同形語を取得する。このほかデータには各資源の[重複回数]や原データのコード番号、頻度情報などが記され、随時データの復元処理を担保する。図 3 は中台成語 DB[データベース]から簡体字形成語“新陈代谢”とその繁体字形“新陳代謝”を特定し、字体変換する場面。



図 3

4-2 検索作業

繁体字の成語を「少納言」「中納言」で検索。それぞれ検索条件を指定できる[図 4]。

⁴ チャート原作者は沈睿(早稲田大学人間科学学術院助手)、原文は中国語。

あることが分かる。中国の成語は古い典故をもつ語が多いが、現代成語には近世のものも多く、[D 三千核心成語] の出典は時代別に以下の順位となる。

宋 (16%)、春秋戦國 (15%)、清 (14%)、明 (11%)、漢 (11%)、唐 (11%)、元 (7%)、南北朝 (5%)、當代 (4%)、晉 (3%)、以下、三國、五代十國、金、隋、十六國、周が各 1%未滿。対して日本語には“大同小異”“四面楚歌”“吳越同舟”“朝令暮改”“異口同音”など、春秋戦國、漢代に起源をもつ故事成語が多く残り、歴史的文化遺産を継承している。

北大3000[D]	頻度	台灣教育[E]	頻度	台灣平衡語料[C]	頻度	韓國世宗[F]	頻度	日本名大漢字[G]	頻度	日本名大和語	頻度
艰苦奋斗	33066	一毛不拔	27	不知不覺(D)	86	견견금금(戰戰兢兢)	98	老若男女	98	輪を掛ける	6639
自力更生	30981	掩耳盜鈴	27	不可思議(VH)	79	명실상부(名實相符)	93	臨機応変	93	我を忘れる	6634
实事求是	26373	一丘之貉	27	理所當然(VH)	74	유언비어(流言蜚語)	81	流言飛語	81	横車を押す	6416
千方百计	22341	囫圇吞棗	27	莫名其妙(VH)	53	자유자재(自由自在)	80	立身出世	80	我に返る	6288
全心全意	20159	班門弄斧	27	前所未有(VH)	52	자초지종(自初至終)	80	油断大敵	80	訳は無い	6252
坚定不移	16005	朝三暮四	27	截然不同(VH)	50	자유분방(自由奔放)	79	優柔不断	79	わき目も振らず	6248
因地制宜	14410	竭澤而漁	27	自然而然(D)	49	이목구비(耳目口鼻)	74	無味乾燥	74	路頭に迷う	6225
独立自主	11271	天衣無縫	27	脫穎而出(VH)	43	동서고금(東西古今)	71	無我夢中	71	目を見張る	6191
成千上万	8886	杞人憂天	27	迫不及待(D)	43	시시각각(時時刻刻)	70	三日坊主	70	目を丸くする	6182
坚持不懈	8837	水落石出	26	不約而同(D)	43	동서남북(東西南北)	69	満場一致	69	目を細くする	6175
战无不胜	8820	名落孫山	26	層出不窮(VH)	41	방방곡곡(坊坊曲曲)	69	本末転倒	69	目を通す	6147
轰轰烈烈	8154	胸有成竹	26	無可奈何(VI)	41	비일비재(非一非再)	68	傍若無人	68	世を去る	6135
前所未有	6876	指鹿為馬	26	依依不捨(VH)	39	이기양양(意氣揚揚)	68	暴飲暴食	68	目を回す	6134
欣欣鼓舞	6684	舉一反三	26	自由自在(VH)	39	신진대사(新陳代謝)	68	不眠不休	68	弱音を吐く	6128
滔天罪行	6467	江郎才盡	26	總而言之(DR)	38	좌지우지(左之右之)	67	百發百中	67	目を白黒させる	6123
丰富多彩	5900	青出於藍	26	成千上萬(Nega)	36	우유부단(優柔不斷)	66	和洋折衷	66	寄ると触ると	6121
朝气蓬勃	5811	口若懸河	25	不遺餘力(VI)	35	각양각색(各樣各色)	66	理路整然と	66	目を凝らす	6110
明目張胆	5641	驚弓之鳥	25	與眾不同(VH)	34	일사불란(一絲不亂)	63	理路整然	63	目を配る	6103
阴谋诡计	5585	含沙射影	25	大街小巷(Ne)	33	유명무실(有名無實)	63	竜頭蛇尾	63	目を掛ける	6096
以身作則	5239	一敗塗地	25	取而代之(VB)	33	요지부동(搖之不動)	63	八方美人	63	目を疑う	6078

表 1

検索語列	学生不知順位 (全328語中)	BCCWJ(中納言) 出現回数
多士濟濟	2	20
特筆大書	3	9
螻蛄の釜	4	10
大所高所	5	53
豪放磊落	7	21
衆議一決	8	2
古希	10	50
立錐の余地も無い	10	19
四角四面	16	26
四分五裂	16	10
一木一草	20	13
出処進退	20	10
切齒扼腕	23	3
一言半句	23	11
青雲の志	26	9
前門の虎、後門の狼	26	6
故事來歷	26	5
得手勝手	26	6
人事不省	26	8
青息吐息	31	22

表 2

ジャンル	成語数	%	ヒット数	%
書籍文学	93	47	1346	65%
書籍社会科学	35	18	359	17%
書籍歴史	24	12	140	7%
Yahoo!ブログ	11	6	28	1%
書籍言語	11	6	35	2%
書籍芸術	7	4	13	1%
書籍哲学	7	4	48	2%
Yahoo!知恵袋	6	3	14	1%
国会会議録	4	2	31	1%
書籍自然科学	3	2	7	0%
雑誌総合	3	2	9	0%
教科書	2	1	3	0%
書籍工学	2	1	30	1%
韻文	1	1	1	0%
新聞	1	1	1	0%
書籍総記	1	1	10	0%
計	211	100	2075	100
検索不能	1		1836(*500)	

表 3

中日同形成語のマッチング作業に先立ち、2010年と2011年にかけて日本人大学生計53名を対象に名大「基本慣用句」の認知度調査を行なった⁶。漢字成語を含め慣用句が現代日本

⁶ 第二外国語として中国語を履修中の政治経済学専攻の学部生対象、名大「(日本語)基本慣用表」所収計3628語から選んだ漢字成語由来の328語を一覧表に作成、瞬時に大体の意味が理解できるか振り分け方

の若年層にどの程度受容されているのか知るためである。結果は、“多士濟濟”“螻蛄の斧”“豪放磊落”“大所高所”“同病相哀れむ”“衆議一決”“古希”“苦心慘澹”“立錐の余地も無い”“功成り名遂げる”“粗製乱造”“四分五裂”などの日本語の成語、慣用句を大学生の80%以上が理解できないことが判明した。“灯火親しむべき候”“老いては子に従え”“蛩雪”“晴耕雨読”“粗製乱造”“薄利多売”“一利一害”は半数が知らない。現代青年層の成語認知度は親世代の50-60歳代に比べ半減している[砂岡和子他(2012)]。

若年世代の漢字成語離れはこのまま加速し続けるのであろうか。成語認知度に差があるのは、記憶に留めやすい成語と忘れやすい成語の別があることを示唆している。両者の原因を明らかにできれば現代日本語中の漢字成語の延命を支援できよう。以下、日中成語の統語法や語用法の違いとコロケーションの特色についてBCCWJ検索を通し見てゆく。

孤立語系統に属する中国語は現在に至るまで一漢字の生産性が高く、意味結合による統辞の自由度が高い。中国語の四字成語は意合結合の縮図であり、日本語や韓国語など膠着語系言語に見られない柔軟な句組織が可能である。[表4]は上掲[北大1000]所収の成語をその内部構造から7分類し、それぞれの出現比率を示したものである⁷。連合関係に加え、主述、動詞補語、動詞目的語、動詞目的補語関係など多用な結合をもつことが分かる。

調査した328語は“呉越同舟”のような中国語伝来の成語(62%)と、“一所懸命”“油断大敵”“絶体絶命”など日本製の慣用句(38%)に分類できる。大多数が知っていると回答した成語は“一喜一憂”“十人十色”“春夏秋冬”“自由自在”“自暴自棄”など後者の慣用句に多い。後者は慣用句の内部統語構造が比較的単純で、並列関係(“春夏秋冬”“自由自在”“自暴自棄”)、もしくは連合関係(“一喜一憂”“十人十色”)など、漢語文法を知らなくとも推理可能な四字句が多い。対して“呉越同舟”{主語(呉越)が述語(舟を同じくする)}や“異口同音”{動詞+目的語(口を異にしながら)動詞+目的語(音[声]を異にする)}は典故や漢語の構造を知らないと意味の理解が難しい。

中国成語内部構造	出現比率	語例
連合関係	44%	不学无术、打草惊蛇
主述関係	23%	草木皆兵、枯木逢春
修飾関係	15%	孜孜不倦、一衣带水
動詞補語関係	4%	重于泰山、无动于衷
非四字成語	8%	三十六计、走为上计
動詞目的語関係	5%	蹉跎岁月、饱经风霜
動詞目的補語関係	2%	问道于盲、入木三分

表4

主要語法機能	出現比率
動詞性成語	72.7%
名詞性成語	11.6%
形容詞性成語	5.3%
副詞性成語	4.7%
分類詞性成語	2.6%
修飾性成語	1.7%
補語性成語	1.4%

表5

四字成語の語用(品詞)機能にも違いがある。[表5]は「現代漢語語法信息詞典」所収の

式で選択させた。回答所要時間はパソコン使用で10-15分、紙媒体で20-30分と、どちらも短時間で回収した。詳しくは砂岡和子、羅鳳珠(2011b)参照。

⁷ 王雷(2010)原表に砂岡が一部修正を加え作成。

成語の語法機能を頻度準に並べたもので、中国語の成語が動詞として振舞うケースが圧倒的に多いことが分かる [姜柄圭(2012)]。

対して日本語の四字漢語は単語としての統辞機能に極めて制約があり、二字漢語や三字漢語と比べサ変動詞や形容動詞の語幹になることが少ないとされる。[野村雅昭(1974)]。例えば以下の四字成語は中国語では動詞として機能するが、BCCWJの日本語用例はともに名詞性用法である。

- ・新陳代謝/「この辺、新陳代謝がうまくいってないんだろなあ」「司令部、何やってるんだろ」1996 筒井康隆(著)最後の伝令, 新潮社
- ・厚顔無恥/二人ともぼくの厚顔無恥には頭をふった。
2004 フリードリヒ・グラウザー(著) 種村季弘(訳)『外人部隊』国書刊行会
- ・粉骨砕身/それがし劉芳亮、商洛山中の危機を救うために粉骨砕身をいとみません
1992 姚雪垠(著)陳舜臣, 陳謙臣(訳)『叛旗』徳間書店

ただし BCCWJ の用例を調べると必ずしもそうとは言い切れない。日本語成語のレキシコン分析のため、NINJAL-LWP for BCCWJ (NLB) ver. 1.00 を利用し⁸、現代日本語で比較的出現頻度の高い漢字四字 24 語について、そのコロケーションをサンプル分析した結果、体言用法に劣らず用言用法が健在で、かつその統語パターンは体言用法のそれより少数精鋭で安定している[表 7]。[表 6] は (NLB) の「総合頻度」と「最多統語比率」の値が特に高い成語につき、それぞれ体言用法 3 語、用言用法 6 語のコーパス情報を引用したものである。限られた資料ではあるが、安定した成語の類型には次の 2 タイプの特徴を認めることができよう。

- 1、統語パターンが一定の成語：“起死回生”“一網打尽”“四面楚歌”“虎視眈々”など
- 2、新しい統語パターンを派生する成語；“喜怒哀楽”“新陳代謝”“厚顔無恥”など

語形	総合頻度	パターンの種類	最多パターン	最多パターン頻度比率	最多パターン中の最多用法	同左頻度比率	最大MI用例	MI値	最多出現ジャンル	同左頻度比率
起死回生	32	7	起死回生+助詞	90.6	起死回生+の+名詞	53.1	の靈薬	18.03	書籍地の文	0.63
喜怒哀楽	68	14	喜怒哀楽+助詞	89.7	喜怒哀楽+の+名詞	26.5	の情	14.10	書籍地の文	1.33
新陳代謝	118	17	新陳代謝+助詞	95.8	新陳代謝を	51.7	をする	14.58	yahoo!知恵袋	4.87
一網打尽	26	3	一網打尽+助詞	88.5	一網打尽に…	92.3	にする	16.38	書籍会話文	0.9
一朝一夕	68	5	一朝一夕+助詞	98.5	一朝一夕に	85.3	～にできる	7.63	書籍会話文	1.25
四面楚歌	26	6	四面楚歌+助詞	80.8	四面楚歌に…	26.9	に陥る	13.40	国会会議録	0.6
虎視眈々	29	3	虎視眈々+助詞	93.5	虎視眈々と	90.3	と狙う	13.58	書籍地の文	0.5
正々堂々	55	5	正々堂々+助詞	83.6	正々堂々と	74.5	と勝ち上がる	15.68	書籍会話文	2.69
意気揚々	74	7	意気揚々+助詞	93.2	意気揚々と…	56.8	と拡張する	15.01	書籍地の文	1.6

表 6

		総合頻度	頻度計	パターンの種類		例	ジャンル
体言的用法	12例	68~3(平均頻度17)	209	14~2(平均パターン7種)	90.6	起死回生+の+名詞	書籍地の文
用言的用法	12例	74~5(平均頻度29)	347	11~6(平均パターン7種)	93.5	虎視眈々と	書籍地の文

表 7

⁸ <http://ninjal-lwp-bccwj.ninjal.ac.jp/>

- 1、 は“起死回生（の靈薬）”“一網打尽（にする）”“虎視眈々（と狙う）”のように、内部構造が比較的複雑な漢字四字成語がより長い統語表現と結合した形式が多い。漢字四字は日本語表現と不適合を起こしやすいが、より長いコロケーションに包摂されることで安定した日本語表現に収まる。定型慣用表現単位の獲得は漢字四字成語の保持と延命に有効であり、四字漢字語であっても動詞や形容動詞の語幹になることが保証される。
- 2、 は“喜怒哀楽”“新陳代謝”“厚顔無恥”のように並列構造など比較的単純な四漢字句で、新しい文脈の中で表現形式の拡張が容易なため、延命が可能と言えよう。

表2は前掲「基本慣用句」328語の認知度テストで、学生が知らない慣用句（成語）トップ20語⁹を[中納言]で検索した結果である。“大所高所”53件、“古希”50件、“四角四面”26件、“青息吐息”22件、“豪放磊落”21件、“多士濟濟”20件を始め、現代日本語中にかんがりの用例数があり、出版年も比較的新しい。コーパスに多数出現するにも拘らず、これらの成語の認知度が低い理由に、現代社会の家族関係や習慣の変化の影響に加え、意味を想起しにくい漢字（古希＝古稀が正字、士、濟、磊、落、所）を含むことが考えられる。そこで学生自身に自分が知らない成語や慣用句を[少納言]で検索体験してもらったところ、身近な現代日本語中に確かにこれらの成語が出現することを、実テキスト付きで確かめることができ、印象に深く刻まれた¹⁰。漢字四字成語の延命支援にもBCCWJは役立つ。

6. まとめ

BCCWJを利用した多言語成語辞典の構築の実際とその活用について紹介を行った。「少納言」「中納言」の利用によってデータの電子的検索と収集分析の利便性がより向上し、他言語とのマッチングデータとしても有用である。言語コーパスは共時的な言語情報だけでなく通時的な流通状況を知ることができる言葉の履歴書である。多地域のコーパスを連携すれば言語間の接触と変容を正確なデータで把握可能となり、対照言語学や比較文化の研究、翻訳をはじめ、語学教育の科学性と効率向上に寄与するに違いない。

謝 辞

本研究は平成22-24年度文部科学省科学研究費補助金〔基盤(C) 課題番号:22520445: 研究代表者砂岡和子〕および2010-2013年度台湾蔣經國國際學術交流基金會(課題番号:RG013-D-09:「歴代語言知識庫建置計畫」:研究代表台湾元智大學羅鳳珠)による補助を得ている。

文 献

野村雅昭(1974)「四字漢語の構造」『国立国語研究所報告54』『秀英出版』pp36～80

⁹ 「学生が知らない」ほど数字が若い、上位31位からBCCWJでヒットしない語を除き、同位があるため20語となった。

¹⁰ 検索作業の感想を授業用チャットに書き込んでいる。

- 宮地裕編(1982) 『慣用句の意味と用法』明治書院』
- 石井正彦(2001) 「文章における臨時一語化の諸形式—新聞の四字漢語の場合—」『現代日本語研究』8, pp. 1-34
- 村木新次郎(2002) 「四字熟語の品詞性を問う」『日本語学と言語学』明治書院, pp. 123-135
- 砂岡和子, 羅鳳珠(2011a) 「『歴代語言知識庫』發展的中日韓漢字成語教學模式」第四屆華語文教學國際研討會發表, 台灣銘傳大學 2011 年 3 月 11~12 日
- 砂岡和子, 羅鳳珠(2011b) 「日本語常用漢字熟語の選好変化と自然言語処理」日本語言語処理学会 (NLP2011) 論文集 CDROM, E3-2
- 砂岡和子, 羅鳳珠, 王雷, 姜柄圭(2012) 「BCCWJ で知る東アジア漢字圏四字成語の受容と変容」言語処理学会第 18 回年次大会 (NLP2012) 論文集 pp. 899-902
- 劉長征&秦鵬(2007) 「基于中国主流报纸动态流通语料库的成语使用情况调查」『語言文字應用』No3
- 王雷, 俞士汶, 朱學鋒, 李芸(2010) 「面向中文資訊處理的成語知識庫的建設與應用」『第五屆文學與資訊科技國際會議』台灣亞洲大學主辦, 2010 年 1 月 22~23 日
- 姜柄圭(2011) 「漢韓成語的語法結構對比分析及教學策略」第四屆華語文教學國際研討會發表, 台灣銘傳大學 2011 年 3 月 11~12 日
- 王雷(2012) 『中国成语 1000』北京大学出版社

関連 URL

- 国立国語研究所『現代日本語書き言葉均衡コーパス (BCCWJ)』
- 「中納言」<https://chunagon.ninjal.ac.jp/>
- 「少納言」<http://shonagon.ninjal.ac.jp/>
- NINJAL-LWP for BCCWJ (NLB) ver. 1.00 <http://ninjal-lwp-bccwj.ninjal.ac.jp/>
- Open MWE for Japanese multiword expressions (MWEs)
<http://openmwe.sourceforge.jp/pukiwiki-j/index.php?Idioms#s0e1b7f3>
- 佐藤理史(2007) 「基本慣用句五種対照表の作成」NLP178, pp. 1~6、『情報処理学会研究報告』 (<http://sslslab.nuee.nagoya-u.ac.jp/index.html> よりダウンロード可能)
- 台灣中央研究院平衡語料成語庫：<http://dblx.sinica.edu.tw/kiwi/mkiwi/>
- 韓国国立国語研究院世宗コーパス：<http://kkma-sc.snu.ac.kr>
- 台灣元智大学[網路展書讀]羅鳳珠：<http://cls.hs.yzu.edu.tw>
- 台灣中央研究院[全球華語文數位教與學資源中心]<http://elearning.ling.sinica.edu.tw/>
- 北京大学信息科学技术学院计算语言学研究所〔语义相关度检索平台〕吴云芳：
<http://klcl.pku.edu.cn:8008/seek/index.php>
- 成語知識庫計畫；砂岡和子 <http://www.f.waseda.jpksunaokacorporuschengyuzhishiku.pdf>