

漢字使用の面から見た 学習者の表現能力の向上

金澤裕之

◆要旨

日本語学習者における書きことばの能力を知る目安の一つに、漢字の使用状況がある。一般に、漢字の使用割合や難漢字の増加が、能力の向上に結び付いていると考えられるからである。本稿では、『YNU書き言葉コーパス』を調査対象として、韓国語母語話者〔K〕・中国語母語話者〔C〕それぞれのレベル別データ（下位群〈L〉、中位群〈M〉、上位群〈H〉）を分析したところ、〔K〕の場合には、L→M→Hの順で漢字使用率が増加していたのに対して、〔C〕の場合にはこれと反対にH→M→Lの順でそれが増加していた。特に〔C-L〕の場合には、熟語の形になりやすい「音読み」漢字の使用例が多く、その中には日本語では一般的でない語彙も含まれ、これらが全体の使用率を高める要因になったことと、レベルが上がるとそれらの使用は減少し、日本人に近い洗練された状況となることが分かった。

◆キーワード

YNU書き言葉コーパス、学習者作文、漢字使用率、訓読み、音読み

◆ABSTRACT

Kanji use is one of the indices of writing proficiency of Japanese language learners. This is because a larger proportion of kanji, especially the difficult characters, is considered to be related to a higher proficiency. In this article, the written data of the Korean learners of Japanese〔K〕 and the Chinese learners of Japanese〔C〕, acquired from the YNU written corpus, were analyzed according to their proficiency levels (low [L], middle [M], and high [H]). There were 10 learners in each language group. In the Korean group, the proportion of kanji increased according to the proficiency levels (L→M→H), while it went in the opposite direction in the Chinese group (H→M→L). Detailed analyses of the kanji used in the learners' writing revealed that the Chinese low learners used many compound kanji words in a Chinese manner, many of which are not used in Japanese, accounting for the increased proportion of kanji used in their writing.

◆KEY WORDS

YNU Written Corpus, written data of learners, proportion of kanji, kun-yomi (Sino-Japanese reading of kanji), on-yomi (Native reading of kanji)

Improvement of L2 Learners'
Writing Proficiency
A Perspective Based on Kanji Use
HIROYUKI KANAZAWA

1 はじめに

書きことばの面から見た日本語の大きな特徴の一つは、文字種類の豊富さということであろう。主要なものとしての、ひらがな・漢字・カタカナに加え、アルファベットや絵文字など、多様な種類の文字記号が一つの文章の中に存在している。それらの中で、一般に日本語学習者にとっての習得問題と関わりが深いと考えられるものが漢字である。特に、非漢字圏からの学習者の場合には、文の中における漢字使用の割合や難しい漢字の増加が、基本的には日本語能力の向上に結び付いていると言える。

それでは逆に、例えば中国語を母語とする多くの学習者のように、いわゆる漢字圏から来た学習者の場合には、同様のことが言えるであろうか。ここでとりあえず簡体字・繁体字などという字体の問題はさておくとすると、単純に考えれば、元来は漢字のみで文章を綴っているそれらの学習者にとって日本語の文を記述するという事は、文中における漢字使用の割合が、母語の場合に比して極端に減少するという事だけは確かである。このような、日本語学習者における、表現能力の向上と関わるかもしれない漢字使用の状況（その使用割合や使用漢字の変化）について、いわゆる初級や中級レベルにおける観察や研究はこれまでもある程度見られるが、いわば上級以上の学習者の場合に、どのような傾向や特色が見られるかということについては、実証的なアプローチはこれまでほとんど行われてこなかったように思われる。

本稿では、以上に述べたような観点から、一つの書きことばのコーパスに着目し、そこに無意識的に表れている学習者の漢字使用の実態を具体的に確認しながら、これと学習者のレベルや表現能力の向上との関わりを検討してみたいと思う。

2 調査資料について

2.1 概要

1.で述べたような関心に基づき、今回、調査資料として選んだ対象は「YNU 書き言葉コーパス」と言われる資料である。この資料は、金澤裕之編『日本語教育のためのタスク別書き言葉コーパス』（ひつじ書房、2014）に添付されたCD-Rに収録されているデータで、非漢字圏と漢字圏出身の日本語学習者による同一の課題に対する書きことばデータが含まれたコーパスである。その概要をまとめると、次のようになる。

- ・日本人大学生30名と留学生60名（韓国語母語話者30名、中国語母語話者30名）に対し、12種類のタスクを課すことによって得られた合計1080編の作文を、コーパスの形にまとめたものである。
- ・採用された12のタスクは、書きことばに関わるさまざまな言語活動・状況・難易度などを考えた上で、大学生や留学生の場合に実際に起こり得そうな具体的な場面が設定されている。
- ・学習者による作文については、主に内容的な面（タスクの達成度、タスクの詳細さ・正確さ、読み手への配慮、体裁・文体）から総合的な評価が明示的に行われ、その結果から、韓国・中国の両言語母語話者のそれぞれについて、下位群・中位群・上位群の各10名ずつのグループ分けが行われている。

なお、資料の収集にあたっては、全ての作文が手書きによって行われており、また、辞書などの使用も認められていない（前掲書、pp.14-15を参照）とのことから、漢字の正確さは別として、調査協力者における文字選択は個人個人の自由に任されていることが分かる^[註1]。

2.2 タスク

課された12のタスクの内容を要約すると、次のようなものである（なお、内

容の後に付した括弧内は、使用媒体、及び、タスクの分類（型・長さ・相手）である。

- タスク1：面識のない先生に図書を借りる
(メール、自発型・長さA・特定〈疎／目上〉)
- タスク2：友人に図書を借りる (メール、自発型・長さA・特定〈親／友人〉)
- タスク3：デジカメの販売台数に関するグラフを説明する
(レポート、自発型・長さA・不特定)
- タスク4：学長に奨学金増額の必要性を訴える
(メール、自発型・長さB・特定〈疎／目上〉)
- タスク5：入院中の後輩に励ましの手紙を書く
(手紙、自発型・長さB・特定〈親／友人〉)
- タスク6：市民病院の閉鎖について新聞に投書する
(投書、自発型・長さB・不特定)
- タスク7：ゼミの先生に故郷の観光スポット・名物を紹介する
(メール、頼まれ型、長さA・特定〈疎／目上〉)
- タスク8：先輩に起こった出来事を友人に伝える
(メール、頼まれ型、長さA・特定〈親／友人〉)
- タスク9：広報紙で故郷や国の料理を紹介する
(原稿、頼まれ型、長さA・不特定)
- タスク10：先生に早期英語教育についての意見を述べる
(メール、頼まれ型、長さB・特定〈疎／目上〉)
- タスク11：友人に早期英語教育についての意見を述べる
(メール、頼まれ型、長さB・特定〈親／友人〉)
- タスク12：小学生新聞で「七夕」の物語を紹介する
(原稿、頼まれ型、長さB・不特定)

3 調査の方法と結果

3.1 量的調査

CD-Rに収録された「オリジナルデータ」^[註3]の作文を対象として、後に述べるような手順を踏んで、調査を行った。なお、作文のタスクの内容は2.2で述べた通りであるが、本研究の目的を勘案した上で、以下に述べるような理由により、次の四つのタスクは対象から除外した。

タスク3⇒簡単な折れ線グラフについての客観的な説明を課題としているので、表現に関する幅が少なくなっているため。

タスク4⇒状況設定にやや無理があり、あまり現実的な状況とは考えられないため。

タスク5⇒相手である後輩との人間関係をどのように設定するかにより、表現の差が非常に大きくなっているため。

タスク8⇒ケータイメールのやり取り(4コマ漫画で表示)を他人(=友人)に伝えるという設定のため、やり取りの内容把握において大きな差異が生じやすいため。

なお、上記の要約からも分かる通り、タスク1と2、及び、タスク10と11は、同一内容の課題について、対象となる相手のみを換えたものである。

3.1.1 具体的方法

本調査における具体的な方法と手順は、次の通りである。

- ①(留学生の場合が、各グループ10名ずつであることを考慮し、)まず、日本人学生の全30名を10名ずつ三つのグループ(I～III)に分けて調査し、その結果として、全体の文字数でも、漢字の使用割合でも、三者の中で中間的な値を示した^[註3]Iのグループ(以下、「J(I)」と略称)を選択して、日本人学生

の代表とした。

②上記の「J (I)」に加えて、韓国語母語話者の下位群 (「K (L)」)、中位群 (「K (M)」)、上位群 (「K (H)」)、及び、中国語母語話者の下位群 (「C (L)」)、中位群 (「C (M)」)、上位群 (「C (H)」) の10名ずつの作文について、次のような字種分類を行った。

③文字の種類を次のa～fの6種に分類して、それぞれの作文における字数を数えた (なお、句読点・中黒・括弧類については、対象から除いた)。

- a. ひらがな
- b. 漢字
- c. カタカナ
- d. アルファベット (単語に加えて、「cm」や「g」などの単位記号も、便宜的にここに加えた)
- e. 記号 (「!」「?」やアラビア数字など)
- f. その他 (絵文字など)

なお、先にも少し触れたが、今回の調査は、あくまでも調査協力者における用字意識を調べることが目的であるため、誤字・脱字や漢字ミス、及び、表現ミスや曖昧な表現などについては、一切の顧慮を行わなかった^[註4]。

3.1.2 調査の結果

それらの総合的な結果は、次の「表1」の通りである (なお、上記分類の「d～f」については、併せた数値が圧倒的に少ない (しかも、その90%以上が「e」にあたる) ため、ここではまとめて「その他」として示す)。

表1 タスクごとの字種別の結果

	T1	T2	T6	T7	T9	T10	T11	T12	合計	
J (I)	ひ	1144 (64.4)	598 (66.9)	2183 (61.2)	1406 (62.7)	2237 (67.0)	1419 (63.7)	1017 (61.4)	5351 (74.8)	15355 (67.2)
	漢	566 (31.9)	207 (23.2)	1217 (34.1)	688 (30.7)	910 (27.2)	759 (34.0)	521 (31.4)	1650 (23.1)	6518 (28.5)
	カ	66 (3.7)	52 (5.8)	160 (4.5)	119 (5.3)	139 (4.2)	31 (1.4)	70 (4.2)	55 (0.8)	692 (3.0)
	他	—	37 (4.1)	7 (0.2)	29 (1.3)	55 (1.6)	20 (0.9)	50 (3.0)	95 (1.3)	293 (1.3)
		1776	894	3567	2242	3341	2229	1658	7151	22858

K (L)	ひ	1241 (68.7)	832 (65.1)	1977 (67.7)	1984 (67.7)	2160 (66.5)	1208 (62.6)	1105 (66.5)	3188 (65.3)	13695 (66.3)
	漢	476 (26.4)	327 (25.6)	896 (30.7)	668 (22.8)	719 (22.1)	675 (35.0)	499 (30.0)	1312 (26.9)	5572 (27.0)
	カ	84 (4.7)	76 (5.9)	33 (1.1)	264 (9.0)	345 (10.6)	34 (1.7)	28 (1.7)	339 (6.9)	1203 (5.8)
	他	4 (0.2)	43 (3.4)	15 (0.5)	16 (0.5)	25 (0.8)	13 (0.7)	30 (1.8)	46 (0.9)	192 (0.9)
		1805	1278	2921	2932	3249	1930	1662	4885	20662
K (M)	ひ	1438 (67.8)	1012 (72.3)	2501 (63.3)	2207 (67.3)	3141 (63.1)	1855 (62.7)	1639 (65.5)	4061 (65.6)	17854 (65.2)
	漢	594 (28.0)	287 (20.5)	1364 (34.5)	767 (23.4)	1154 (23.2)	1007 (34.1)	768 (30.7)	1605 (25.9)	7546 (27.6)
	カ	87 (4.1)	67 (4.8)	64 (1.6)	268 (8.2)	621 (12.5)	85 (2.9)	45 (1.8)	452 (7.3)	1689 (6.2)
	他	3 (0.1)	34 (2.4)	25 (0.6)	35 (1.1)	63 (1.2)	9 (0.3)	49 (2.0)	72 (1.2)	290 (1.0)
		2122	1400	3954	3277	4979	2956	2501	6190	27379
K (H)	ひ	1544 (64.3)	1085 (65.6)	2323 (58.8)	2535 (61.9)	3033 (62.0)	2221 (62.4)	1697 (64.3)	4201 (65.2)	18639 (62.9)
	漢	746 (31.0)	424 (25.6)	1550 (39.2)	1043 (25.4)	1319 (27.0)	1249 (35.1)	821 (31.1)	1863 (28.9)	9015 (30.4)
	カ	103 (4.3)	92 (5.6)	61 (1.6)	474 (11.6)	423 (8.7)	59 (1.7)	62 (2.4)	327 (5.1)	1601 (5.4)
	他	9 (0.4)	53 (3.2)	17 (0.4)	46 (1.1)	113 (2.3)	30 (0.8)	58 (2.2)	50 (.8)	376 (1.3)
		2402	1654	3951	4098	4888	3559	2638	6441	29631
C (L)	ひ	1131 (66.0)	851 (71.9)	2015 (60.1)	1843 (62.0)	2640 (60.4)	1486 (59.9)	1198 (60.9)	3457 (64.0)	14621 (62.4)
	漢	542 (31.6)	306 (25.9)	1249 (37.3)	975 (32.8)	1504 (34.4)	899 (36.2)	713 (36.3)	1803 (33.4)	7991 (34.1)
	カ	37 (2.2)	13 (1.1)	76 (2.3)	97 (3.3)	132 (3.0)	84 (3.4)	34 (1.7)	95 (1.7)	568 (2.4)
	他	4 (0.2)	13 (1.1)	10 (0.3)	58 (1.9)	93 (2.2)	12 (0.5)	22 (1.1)	47 (0.9)	259 (1.1)
		1714	1183	3350	2973	4369	2481	1967	5402	23439
C (M)	ひ	1309 (67.8)	955 (70.8)	2246 (61.0)	2504 (61.4)	3639 (61.5)	2134 (62.8)	1607 (63.3)	4019 (65.2)	18413 (63.4)
	漢	588 (30.4)	354 (26.3)	1291 (35.1)	1343 (33.0)	1850 (31.3)	1178 (34.7)	780 (30.8)	2007 (32.5)	9391 (32.3)
	カ	33 (1.7)	13 (1.0)	136 (3.7)	205 (5.0)	367 (6.2)	80 (2.3)	125 (4.9)	97 (1.6)	1056 (3.6)
	他	1 (0.1)	26 (1.9)	9 (0.2)	23 (0.6)	60 (1.0)	6 (0.2)	25 (1.0)	43 (0.7)	193 (0.7)
		1931	1348	3682	4075	5916	3398	2537	6166	29053
C (H)	ひ	1319 (67.5)	1019 (68.2)	2440 (59.8)	2217 (62.9)	3614 (62.2)	2049 (64.3)	1605 (64.6)	5044 (67.3)	19307 (64.3)
	漢	592 (30.3)	410 (27.5)	1440 (35.3)	1136 (32.3)	1850 (31.8)	1049 (32.9)	802 (32.3)	2337 (31.2)	9616 (32.0)
	カ	42 (2.1)	27 (1.8)	171 (4.2)	147 (4.2)	290 (5.0)	54 (1.7)	45 (1.8)	59 (0.8)	835 (2.8)
	他	1 (0.1)	37 (2.5)	30 (0.7)	22 (0.6)	60 (1.0)	33 (1.1)	33 (1.3)	57 (0.7)	273 (0.9)
		1954	1493	4081	3522	5814	3185	2485	7497	30031

今回、主な対象となる漢字については後に詳しく述べるので、先にひらがなとカタカナについて、日本人学生の場合も参照しながら、韓国語・中国語母語話者の場合の概要について触れておきたい（以下では、三者のそれぞれを、必要に応じて「J」「K」「C」と略称する）。まず先に、カタカナについては、Kでおよそ6%、Cでおよそ3%ということ、両者ともに比較的安定していると言える。Jの場合も3%ということ、Kの場合が多めであると言えるが、この点に関しては、記事の中に出てくる韓国の地名・名物・料理名などといった固有名詞類が、一般にはカタカナで表記されるため、その影響が現れたものと推察される。一方ひらがなについては、Kではレベルが上がるにつれて数値が少しずつ下降しているのに対し、Cでは逆に、レベルが上がるにつれて数値が少しずつ上昇しているのが分かる。この点に関しては、次に述べる漢字の状況の影響を受けたために、両者が数値的に裏返しのような状況になっていると言えるのではないだろうか。

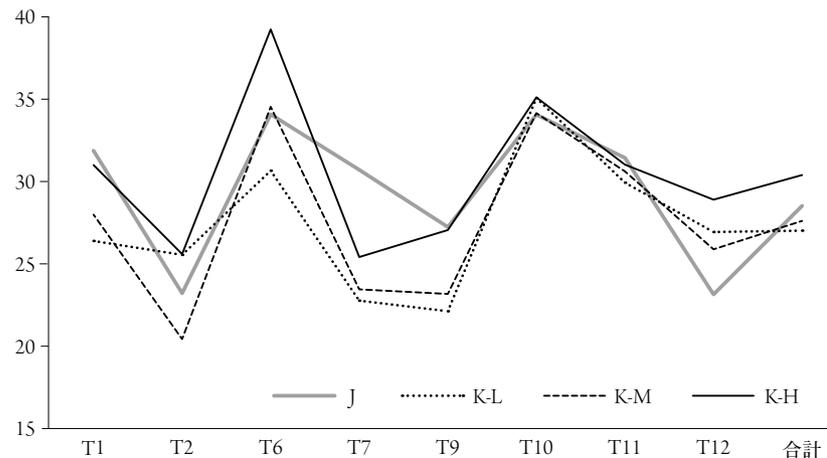


図1

3.1.3 母語別の漢字の使用割合

次に、今回の調査において注目している漢字の使用割合について、八つのタスク別の値と合計の分の結果を、韓国語、及び、中国語の母語話者のグループ別にして「図1」「図2」で示す。なお、それらとの比較対照のために、「J (1)」の結果のグラフも、二つの図の両方に載せておく。

上記の、「表1」ならびに「図1・2」の結果から明らかになった量的側面からの特色を、先にまとめてみることにしよう。

まず、本研究の中心的な目的としての、学習者における漢字使用率についてである。この点について、最も明確な数字で表れている結果は、上記のどの図表でも右端に示されている「合計」の結果である。まずKの場合は、全く同様の8課題に対する10名ずつの合計が、下位群 (27.0) < 中位群 (27.6) < 上位群 (30.4) と、レベルが上がるに従って数値の上昇している様子が見てとれる。下位群と中位群の差はさほど大きいとは言えないが、上位群には際立った違いが表れ、一般的に推測できる非漢字圏学習者における予想 (⇒学習者のレベルの上昇に従い、漢字使用割合も上がる) を裏付けていると言える。

それに対して、Cの場合は、同じく全く同様の8課題に対する10名ずつの合

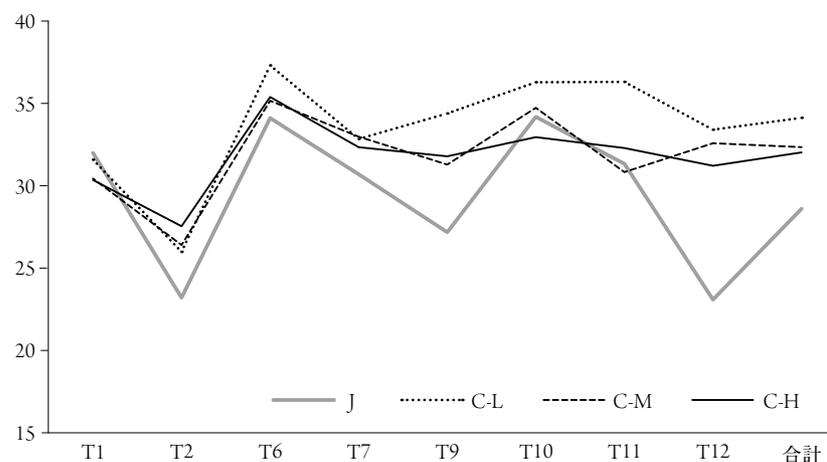


図2

計が、下位群 (34.1) > 中位群 (32.3) > 上位群 (32.0) と、Kの場合とは対照的に、レベルが上がるに従って数値の下降している様子が見てとれる。こちらの場合、上位群と中位群の差は必ずしも大きくはないが、下位群には際立った違いが表れている^[注5]。

そして、Jの場合も参照すると、Kの場合、上位群の数値はJの場合を上回ってしまっているが、全体としてはレベルが上がるに従って漢字使用の割合が上昇して母語話者のそれに近づいてゆき、一方、Cの場合には、レベルが上がるに従って漢字使用の割合が下降するという方向で、やはり母語話者のそれに近づいていると言えるのである。

次に、タスクごとの結果を見てみると、まずタスク1の「面識のない先生への依頼」というかなりハードルの高い課題では、先に示した「合計」の順位がほぼ反映されたものとなっている。一方、タスク1とは相手のみが換わるタスク2「友人への依頼」というかなり気が楽になる課題では、友人との親密さの度合を表現にどう反映させるかという点で個人個人による差が激しくなるため、全体としてまとまった傾向は窺えない。

タスク6は、新聞への意見投書という最も公式性の強い課題であったためか、特にKの場合にレベル差に対応する漢字使用率の違いが顕著に表れた。一方、Cの場合にはKの場合のような顕著な差は出なかったが、全体の場合と同様に、下位群の場合に最も高い使用割合の数値が表れている。

タスク7と9は、ともに自分の故郷の観光スポットや料理を紹介するという課題であるため、KとCのそれぞれのレベルにおいて、比較的似通った数値となっている。レベル差の点では、ここでもKの場合は下位群の数値が低めなのに対し、Cの場合には下位群の数値が高めとなっている。

タスク10と11は、先生及び友人に対して、研究についての意見を述べるという設定で、留学生であるインフォーマントたちにとっても比較的なじみの深い課題であったためか、全体にJの場合ととても似通った数値が示されているのが大きな特徴である。それらの中で、Cの下位群のみが、この二つの課題においても、特に高い数値を示していることが目立っている。

最後に、少なくとも三国に共通する説話である「七夕」の物語を紹介するというタスク12では、個人個人の自由な記述方法が反映して、文章の長さその

ものにおける長短の違いが顕著であった。KとCにおけるレベル差も一応は見られたが、それ以上に際立ったのが、Jにおける全体的な数値の低さであった。これは、推測するに、学習者たちが物語の記述そのものに集中したと見られるのに対して、母語話者の場合は、内容の記述に加えて、対象が小学生新聞であることがある程度考慮に入れられていたと考えられ、その結果が数値の相対的な低さに関係したと思われる。

以上、漢字の使用に関する全体の状況を見てみると、数値的な差異としてはさほど大きいものではないが、「はじめに」の部分で仮説的に述べたように、非漢字圏であるKの場合とは逆に、漢字圏であるCの場合に、レベルが上がるにつれて却って漢字の使用割合が減少していくという傾向が、ここに窺えるものと思われる^[注6]。次項では、そうした傾向を生み出したと考えられる要因について、実態をより詳しく、具体的に検討するような質的な面からの調査によって、そのアプローチを進めてみたい。

3.2 質的調査

前項で述べた量的な調査を参考にしつつ、ここではさらに詳しく、タスクを絞った上で、具体的かつ質的な調査を進めてみる。

以下では、タスクごとに対象となる全ての漢字語彙を調査対象とする必要があるため、選択するタスクをさらに少数に絞ることとした。その他のタスクを対象から外した理由は、それぞれ次の通りである。

タスク7・9⇒この二つは、上記の通りに個々のインフォーマントの故郷の違いによって、観光スポットや料理の種類が異なるため、必然的に使用される語彙も異なりが大きくなることから、対象から除外した。

タスク10⇒先生へのメール連絡という点でタスク1と共通するところがあるが、1の方では相手が面識のない先生であるのに対し、10では面識のある先生が対象であり、先生との距離感や意識の部分で個人差が出やすい点を考慮して、タスク1の方を採用した。

タスク2⇒これも、友人へのメール連絡という点でタスク11と共通するものであるが、タスク2の場合、特に日本人学生において、(親しい友人を想定しているときに)内容が極端に短くなっている傾向が目立ち、この場合には調査対象となる語の数も極端に少なくなってしまう^[註7]ことから、こちらを除外した。

タスク12⇒これは「七夕」の内容を紹介するという興味深いタスクであるが、一つには、全体の長さが圧倒的に長くなりやすいことと、基本的には三国でほぼ同様の内容とはなっているが、固有名詞などの具体的な名称の部分で違いが見られる^[註8]ということから、対象からは外した。

その結果、実際に調査対象としたタスクの数は三つで、その内容は、次の通りである。

タスク1: 面識のない先生に図書を借りる

(メール、自発型・長さA・特定〈疎/目上〉)

タスク6: 市民病院の閉鎖について新聞に投書する

(投書、自発型・長さB・不特定)

タスク11: 友人に早期英語教育についての意見を述べる

(メール、頼まれ型・長さB・特定〈親/友人〉)

3.2.1 具体的方法

本調査における具体的な方法と手順は、次の通りである。

①量的調査の部分で示した、10名ずつの七つのグループ(「J (I)」、「K (L)」、「K (M)」、「K (H)」、「C (L)」、「C (M)」、「C (H)」)について、それぞれのグループを一まとまりとして、三つのタスクにおいて使用された漢字語彙を全て調べた。なお、この調査では、それぞれのグループが全体として使用している漢字(語彙)のバリエーションを見ながら比較検討を進めるため、以下では全ての点において「異なり語」の単位での実態を分析することにする。

②対象となった三つの課題(タスク1・6・11)の作文において、使用された漢字の全てを、「単語」のレベルで抽出した^[註9]。その後、その単語が「訓(くん)よみ」で使用されたものか「音(オン)よみ」で使用されたものかを区別して^[註10]、用例数を数えた。

③韓国語母語話者〔K〕と中国語母語話者〔C〕については、レベル別に三つのグループに分かれている(上位群、中位群、下位群)ため、そのグループごとに集計を行い、それぞれの母語の場合別に比較した。

④それらの結果と日本人の場合〔J (I)〕とを比較対照した。

3.2.2 分析の要点

この質的な面での調査では、先に行った量的な調査の結果を参考にして、主に次に挙げる二つの点に絞って、経過を報告する。

その一つ目は、学習者の一つの目標として日本人学生の場合を想定するもので、「日本人学生の全使用語彙に対する共通度」を、K・Cそれぞれの母語についてレベル別の違いを調べてみた。ここに焦点を当てた理由は、具体的なタスクに適切に対応する、個々の課題での基本的ともいえる語彙を対象として、そこでの学習者による使用状況を見てみようとしたものである。

二つ目は、学習者のレベル別グループそれぞれの使用漢字の特性を見てみたいことから、それぞれの母語において、他のレベルでは使われていない「そのグループのみの漢字使用語彙」に焦点を当てて、その特徴を調べてみた。ここに焦点を当てた理由は、前者の場合とは反対に、回答者一人一人の自由な個性や表現の幅・広がり、個々のグループ特有の使用語彙の面から検証してみようとしたものである^[註11]。

3.2.3 調査の結果

本調査の過程と結果を、上に挙げた二つの要点に絞って説明する。

(1) 日本人学生の場合との共通度

ここでは、上の調査方法の部分で述べた通り、対象とする三つのタスクについて、J (I) の10名による使用漢字語彙の全てを基準として、それと共通する、

各々のグループにおける使用漢字語の数を、訓読みのものと音読みのものとに分けた形で示すことにする。その結果は次に示す、表2～4の通りである。なお、それぞれのグループの合計の部分の()内の数字は、Jに対する共通度の割合である。

表2 タスク1の結果

	J (I)	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	40	21	26	26	25	27	28
音読み	49	18	25	31	21	25	26
合計 (%)	89(100)	39(44)	51(57)	57(64)	46(52)	52(58)	54(61)

表3 タスク6の結果

	J (I)	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	119	41	51	56	47	54	54
音読み	150	46	59	60	51	55	66
合計 (%)	269(100)	87(32)	110(41)	116(43)	98(36)	109(41)	120(45)

表4 タスク11の結果

	J (I)	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	70	28	40	41	30	36	39
音読み	84	28	42	46	37	38	43
合計 (%)	154(100)	56(36)	82(53)	87(56)	67(44)	74(48)	82(53)

また、この結果のポイントをより分かりやすく示すために、三つのタスク別に、中位群であるMの場合を中心とするKとCの共通度の推移を、左から右へL→M→Hの順で、次の図3～5に掲げてみる。

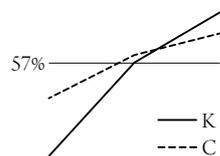


図3 タスク1

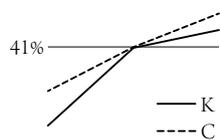


図4 タスク6

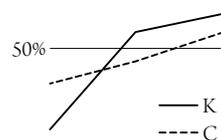


図5 タスク11

まず、三つの図それぞれの基準線として示した数値の高低(タスク1=57>タスク11=50>タスク6=41)の差は、全体の字数そのものの長短(一般に、短いと高く・長いと低い)、及び、定型性の高低、によるものと考えられる。タスク1は、タスクの内容(⇒依頼)が明確であることから、全体の語数は少なく、また文の定型性も高くなる。それに対してタスク6は、新聞への意見投書という内容から容易に想像できる通り、全体の語数が長くなるし、表現も多様なものとなりやすく、定型性は低くなる。タスク11は、両者の中間的な性格を有するものと言えよう。

結果のグラフを見ると、K・Cともに、L→M→Hの順で数値は確実に右肩上がりになっていることは6例に共通する。ただし、その割合(=傾き)は、全体としてKの方が大きく着実な伸びを見せているのに対して、Cではそれほど大きな差はない。その中でやや例外的なのは、タスク6・11における、M→Hの場合であるが、ここでもKとCの両者に大きな差が見られるわけではない。

なお、結果が煩瑣になるために、ここで敢えて図化することはしていないが、訓読みと音読みの別に分けた結果についても、L→M→Hの順に従って、数字はそのほとんどが増加しており、2か所(タスク1の訓、KのMとH、タスク6の訓、CのMとH)に同一の数字が見られるが、程度の大小は別として、訓・音別の数値でも、全て右肩上がりの数値を示していることが分かる。

以上のような結果から、ここで注目しておきたいポイントは、全体としてはKの場合ほど顕著な形ではないが、Cの場合にも確実に右肩上がりの状況が生まれているという点である。前項の量的調査の結果に表れていた、漢字の使用語彙数における状況(⇒右肩下がり)に対して、こうした「Jと共通する語」というタスク内容に沿った基本的な語彙の場合には、Cの場合にも、Kと逆の形となるような特異な状況は生まれていない、と考えられるのではないだろうか。

(2) グループごとの独自の使用語の特色

ここでは、まずはKとCの母語別に分けた上で、三つのグループの使用語彙について、他の二つのグループ、及び、日本人学生(J)の場合と重ならないという、いわば「そのグループのみの使用の漢字語彙」を、これも訓読みのものと音読みのものとに分けて、その数字を示してみたい。三つのタスク順に、

表5～7となる。なお、表2～4では、合計の部分において、()内にJの値に対する相対的割合を示したが、ここでは基準となるJのような数値がないため、とりあえず省略する。

表5 タスク1の結果

	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	10	17	17	11	12	19
音読み	12	18	33	23	20	21
合計	22	35	50	34	32	40

表6 タスク6の結果

	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	23	54	64	27	50	54
音読み	35	77	95	100	63	72
合計	58	131	159	127	113	126

表7 タスク11の結果

	K (L)	K (M)	K (H)	C (L)	C (M)	C (H)
訓読み	6	19	32	17	22	26
音読み	26	38	39	46	36	38
合計	32	57	71	63	58	64

さてこちらでも、この結果のポイントをより分かりやすく示すために、三つのタスク別に一つずつの図を掲げてみたい。今回は、前の場合とは異なり、基準となるJの数値が存在しないため、母語別にL→M→Hの変化の様子を見るために、仮にLの場合の数字を「1」とした場合のMとHの数値を計算して示す。次に掲げる図6～8は、三つのタスク別に、そのように行ったKとCの合計の部分の値を図化したものである。

この三つの図を見たとき最も目立つ特色は、言うまでもなく、KとCのグラフの形の相違である。Kの場合はいずれもかなり急激な右肩上がりとなり、Hの場合には2.2～2.7というLの2倍以上の数値となっているのに対し、Cの場合には、Mにおいて少し下がることも含めて、HにおいてもLよりほとんど増

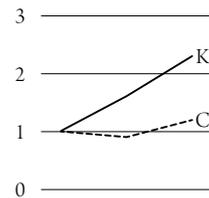


図6 タスク1

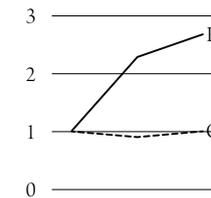


図7 タスク6

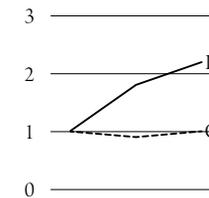


図8 タスク11

加することがない。

この顕著な差の要因をさらに検証するにあたり、今回は実態に近い姿を反映させるためいささか極端な図にはなるが、KとCについて、それぞれの訓読みの場合と音読みの場合を分けた上で、今回はLの場合を「1」と仮定する方法ではなく、語数の数字そのままの、いわば「生の語数」によって変化の状況を表してみよう。三つのタスク順に、図9～11となる（なお、タスクによって語数の数字が大きく異なっているため、それぞれの図のスケールは変えている）。

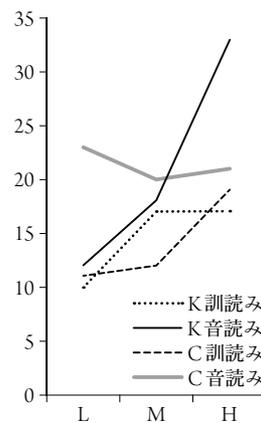


図9 タスク1

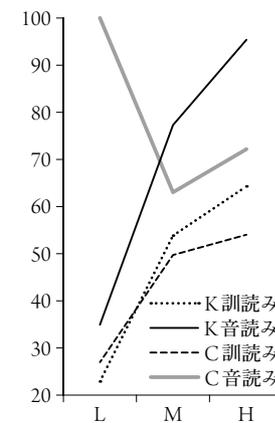


図10 タスク6

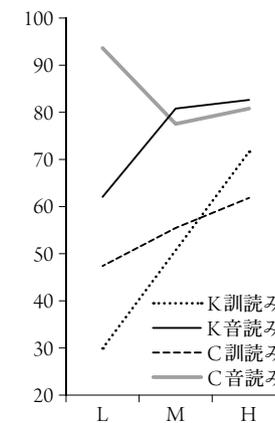


図11 タスク11

この三つの図から分かる通り、この場合でもKではレベルが上がるに従い、訓読みでも音読みでも、語数は上昇している。一方Cの場合、訓読みの例では

ある程度語数は右肩上がりであり上昇している。それに対して音読みの場合は、語数そのものが全体に多いのに加えて、Lの場合の語数が際立って高いため、右側が相対的に下がる状況となっている。

このように詳細に見てくると、前項の量的調査における際立った特徴となっていた、Cの漢字使用数におけるL > M > Hという現象に関しては、「グループごとの独自の使用語」のうち、特に音読みとなる語の場合に、Lの使用例の多さが大きな要因として挙げられることが分かってきた。特に漢字の音読み語の場合は、一般に二文字以上の熟語を形成する確率が高いこともあって、結果的に使用される漢字文字数が多くなるものと考えられる^[註12]。ただし、今回の調査では、全体数としてMの方がHの場合よりわずかにしても多くなるという点については、必ずしも明らかにすることはできていない。この点を探究するためには、今回考察の対象とすることができなかった、語の重複使用と関わる「延べ語数」の面からのアプローチなどが、さらに必要となるかもしれない。

この点については、今後の課題となるだろう。

4 おわりに

これまで見てきた通り、漢字圏である中国語を母語とする日本語学習者による書きことばの習得に関して、そのある時期には、漢字語のうち日本語的な性格が強いと考えられる訓読みの語を、少しずつ着実に学んで増やしていくと考えられる一方で、漢字語のうち漢語との共通性を少なからず有する音読みの語については、その表意的性格や語形の共通性から、日本語の語彙として一般的なものであるかどうかにかかわらず、とりあえずそうした熟語を使用することによって文章表現を続けながら、語義の相違やコロケーションを学ぶことを通じて、より適切な用法を洗練させてゆくというような戦略を採用しているらしい方向性が考えられるようになった。本稿で調査対象とした「YNU コーパス」という書きことば資料には、本人にとって意識的であるか無意識的であるかどうかは別としても、そうした様相の一端が、奇しくも表れていたと言えるように思われる。

最後に、本稿における調査の過程から、そうした傾向の一端が垣間見られた

具体的な例を一つ挙げて、紹介してみようと思う。

本稿における、漢字語彙の質的な面からの悉皆調査では、使用された語を、訓読みのものと音読みのものに分けた上で、五十音順に分類を行っていった。その中で圧倒的に用例が多かったのはサ行の「し/シ」の項目であったが、それに次いで用例が多かったのは、ともにカ行の「か/カ」と「け/ケ」の項目であった。その中で、後者の「け/ケ」の場合には「し/シ」や「か/カ」と大きく異なる特徴があり、訓読みの「け」で始まる語の方はたった1語（「削る」）しか用例がなかったのに対し、音読みの「ケ」で始まる語は非常に多く、三つのタスクの場合を併せて73語が出現していた。

次に挙げるのは、新聞への投書課題であるタスク6（市民病院の閉鎖に関する意見）の例であるが、C（L）をはじめとする使用語のパリエーションは、中国語母語話者の語彙的な豊かさを象徴している感がある。

以下に、JとCの3グループ分の具体的な語例を示す。L→M→Hの流れの中に、語彙的な面での一種の洗練のようなものを感じ取ることはできないだろうか（順番は基本的に出現順。Cの途中の「/」の記号は、その前の「Jと共通する例」と、その後の「そのグループ独自の例」の間の境界を示す）。

- J : 現在、経営難、経営、健康、検討、現行、結果、現象、元気、原因、現状
- C (L) : 現在、経営、健康、検討/経営者、経済的、嚴重、経費、掲載、決断者、現地、権益
- C (M) : 現在、経営難、経営、健康、検討、現行、原因/検査、健全、警察、怪我人
- C (H) : 現在、経営難、経営、健康、検討、結果、原因/建設、決断、継続

〈目白大学/横浜国立大学〉

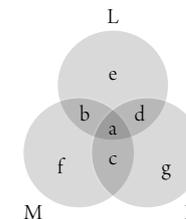
注

- [注1] …… その他の具体的な状況は、次のようなものである。
- ・ 指示文については、必要に応じて母語訳が与えられている。
 - ・ 一人あたりの調査時間は4時間程度で、二日に分けて行う者と、休憩を取りながら一日で行う者があった。
 - ・ 作文を書く時間には制限を設けず、個人個人のペースで書き進められている。
- [注2] …… 収録されたデータには、「オリジナルデータ」と「補正データ」の2種類がある。そのうち「オリジナルデータ」とは、調査協力が手書きで書いたものを、できる限りそのまま再現したものであり、一方「補正データ」とは、その後のデータベースでの検索の利便性を主な目的として、「オリジナルデータ」を適宜補正したものである。
- [注3] …… 三つのグループにおいて、それぞれに使用された全体の文字数、及び、漢字の文字数（使用割合）は次の通りである。
- | | 全体 | 漢字(同%) |
|-----------|-------|---------------|
| * I グループ: | 22858 | — 6518 (28.5) |
| II グループ: | 29291 | — 8396 (28.7) |
| III グループ: | 22554 | — 6182 (27.4) |
- [注4] …… この点に関わるものとしては、「オリジナルデータ」の方を利用して、「貸す」と「借りる」における漢字表記の混同を分析した金庭（2014）の論文が、前掲書に掲載されている。同論文によれば、この種の混同は学習者の場合ばかりではなく、母語話者の場合にも多数見られるとのことである。
- [注5] …… 個別に見ても、タスク2（友人に図書を借りる）とタスク7（ゼミの先生に観光スポット・名物を紹介する）を除く六つのタスクでは、明らかに下位群の値が高くなっており、他の2群との違いは明らかである（なお、タスク7でも、下位群と中位群は似通った数値で、いずれも上位群よりは高い）。
- [注6] …… ただし詳細に検討すると、タスク2などでは必ずしもこの傾向に当てはまらない例も見られる。このタスク2においては、親しい友人への「くだけた表現」という点で、学習者の作文にはある種の不自然さも窺われており、そうした面が数値的な結果に影響を及ぼしている可能性も考えられる。
- [注7] …… 参考までに、最も短かった具体例（全42文字）を次に掲げておく。
- ・ 『環境学入門』っていう本を卒論で使いたいんだけど、持っている？
取りに行くから貸してくれない？ (J-004)
- [注8] …… 分かりやすい具体的な例を挙げると、主人公である二人の名前が、Jでは「おりひめ」「ひこぼし」など、Kでは「ジクニョ」「ギョソウ」など、Cでは「織姫」「牛郎」など、となっている。
- [注9] …… 単語は、基本的に辞書の見出し語のレベルで抽出した。ただし、課題の内容から必然的に要請される固有名詞（『環境学入門』、市民病院、など）や、複合語（経営難、国際化、など）については、全体で1語と認定した。また、

活用語についても、同一の意味を表している場合には、これも1語と認定した。

- [注10] …… いわゆる、重箱読みや湯桶読みのように、1語の中の複数の漢字を音訓取り混ぜて読むような場合には、便宜的に、最初の文字の読み方の方に分類した。例えば、「気軽（キがる）」「役割（ヤクわり）」などは音読みの方に、「場所（バシヨ）」「手数（テスウ）」などは訓読みの方に含めた。ただし、それらの例は併せても10語程度であり、全体（約1800語）から見れば、0.6%ほどに過ぎない。

- [注11] …… 下位群〔L〕・中位群〔M〕・上位群〔H〕における漢字の使用語彙の関係を図示すれば、右の図のようになり、ここで調査の対象とするのは、e（下位群）・f（中位群）・g（上位群）の部分に含まれる語についてである。



- [注12] …… C (L) の場合には、次に挙げるような、一般に国語辞書に掲載されていないか、あるいは、一般には使用されない熟語を使用するケースが多かった。この傾向は、C (M) や C (H) にはほとんど見られないものである。

- タスク1： 写作、自本、指導先生、など
- タスク6： 就病、居住地、病患、など
- タスク11： 行式、大数、勉強能力、など

参考文献

- 金澤裕之（編）（2014）『日本語教育のためのタスク別書き言葉コーパス』ひつじ書房
- 金庭久美子（2014）「手書き作文に見られる漢字表記の誤りについて—「貸す」と「借りる」の混同に注目して」金澤裕之（編）『日本語学習者のためのタスク別書き言葉コーパス』pp.419-429. ひつじ書房
- 劉瑞利（2017）「日本語学習者の「名詞+動詞」コロケーションの使用と日本語能力との関係—「YNU書き言葉コーパス」の分析を通して」『日本語教育』166, pp.62-76.

