

キーワード： 今昔物語集 巻 漢字 類似度

## 要旨

今昔物語集に出現する漢字異なり 3千余・延べ 488千余を 28巻の巻ごとに整理して、巻の間での宮島達夫の類似度を算出し、巻巻の関係を見る。巻間 378対のうちで、類似度が小さく、あまり似ていないものは、本朝世俗部諸巻と天竺震旦部および本朝仏法部前半の諸巻との間である。類似度が大きく、似通うものは、本朝仏法部内の諸巻の間であり、ほかに隣接 2巻の間などあちらこちらに散在する。また、ある巻について、他の全巻を一体とみなしてそれとの類似度を算出する、という操作をすると、本朝仏法部の諸巻で値が大きく、本朝世俗部で小さい。結果は意外なものではなく、漢字の類似度の有効性も知られることになる。

## 1 類似度一覧

新日本古典文学大系『今昔物語集 一～五』の本文に現れる漢字をすべて取り上げ、巻ごとに整理する。ただし、各巻・各語の標題は除外して、各語おおむね「今ハ昔、……トナム語り伝ヘタルトヤ。」のようである本文に就く。巻第8・巻第18・巻第21は本文を全く欠いて合計巻数28、その全巻で漢字異なり 3,100, 延べ 488,290 である。各巻の漢字の間で、宮島達夫の類似度を算出する。

二巻ごとの 378 (=27×28/2) 対に得られる数値は、一覧すると次ページの表となる。ただし、本来の表は次のような形であり、視認のしやすさを狙って分割している。すなわち、次の縦線で左中右に3分割し、かつ横線の下を削除して、【左】

【中】【右】と名付けた部分を次ページで上中下に配する。

	巻2	…	巻9	巻10	巻11	…	巻19	巻20	巻22	…	巻30	巻31	他巻
巻1	.7408	…	.5913	.6211	.6264	…	.6092	.6384	.5225	…	.5234	.5708	.6796
…	…	…	…	…	…	…	…	…	…	…	…	…	…
巻9	【左】		.6796	.6100	…	.6293	.6584	.5087	…	.5619	.6147	.6839	
巻10				.6621	…	.6969	.7190	.5880	…	.6337	.6946	.7525	
…					…	…	…	…	…	…	…	…	…
巻19					【中】		.7730	.6353	…	.6981	.7727	.8014	
巻20								.6020	…	.6472	.7415	.8218	
…								…	…	…	…	…	…
巻30									【右】		.7166	.6672	
	巻2	…	巻9	巻10	巻11	…	巻19	巻20	巻22	…	巻30	巻31	.7657

卷1	卷2	卷3	卷4	卷5	卷6	卷7	卷9	卷10	卷1
卷2	▲7408	▲7578	△7214	.6840	.6713	.6172	▽5913	.6211	卷2
卷3		▲7677	▲7348	△7052	.6778	.6364	.6312	.6507	卷3
卷4			▲7543	△7038	.6958	.6360	.6101	.6413	卷4
卷5				▲7577	△7189	.6832	.6518	△7115	卷5
卷6					.6615	.6484	.6394	△7257	卷6
卷7						▲7406	.6522	.6383	卷7
卷9							.6958	.6429	卷9

【左】

卷2	卷3	卷4	卷5	卷6	卷7	卷9	卷10	卷11	卷12	卷13	卷14	卷15	卷16	卷17	卷19	卷20	卷1		
卷2	卷11	卷12	卷13	卷14	卷15	卷16	卷17	卷19	卷20	卷1	卷2	卷3	卷4	卷5	卷6	卷7	卷9		
卷3	.6264	.6395	.6112	.6272	.5979	.5964	.6305	.6092	.6384	卷2	.6274	.6482	.6189	.6422	.6065	.6166	.6499	.6157	.6501
卷4	.6402	.6505	.6176	.6433	.6031	.6053	.6396	.6099	.6493	卷3	.6402	.6505	.6176	.6433	.6031	.6053	.6396	.6099	.6493
卷5	.6905	△7145	.6700	△7030	.6637	.6848	△7063	.6815	△7159	卷4	.6905	△7145	.6700	△7030	.6637	.6848	△7063	.6815	△7159
卷6	.6584	.6830	.6484	.6781	.6255	.6835	.6765	△6968	△7181	卷5	.6584	.6830	.6484	.6781	.6255	.6835	.6765	△6968	△7181
卷7	△7070	△7276	.6901	△7066	.6693	.6252	△7134	.6292	.6757	卷6	△7070	△7276	.6901	△7066	.6693	.6252	△7134	.6292	.6757
卷9	.6676	△7197	△7250	▲7427	.6717	.6336	△7088	.6323	.6750	卷7	.6676	△7197	△7250	▲7427	.6717	.6336	△7088	.6323	.6750
卷10	.6100	.6463	.6280	.6430	.6148	.6344	.6506	.6293	.6584	卷9	.6100	.6463	.6280	.6430	.6148	.6344	.6506	.6293	.6584
卷11	.6621	.6796	.6455	.6780	.6413	△6989	.6683	△6969	△7190	卷10	.6621	.6796	.6455	.6780	.6413	△6989	.6683	△6969	△7190
卷12		▲7461	.6782	.6878	.6589	.6462	△7013	.6664	△7005	卷11		▲7461	.6782	.6878	.6589	.6462	△7013	.6664	△7005
卷13			▲7624	▲7831	▲7396	△7215	▲7671	△7285	▲7632	卷12			▲7624	▲7831	▲7396	△7215	▲7671	△7285	▲7632
卷14				▲7684	△7337	.6666	▲7359	.6763	△6977	卷13				▲7684	△7337	.6666	▲7359	.6763	△6977
卷15					△7268	△7116	▲7537	△7210	▲7447	卷14					△7268	△7116	▲7537	△7210	▲7447
卷16						.6811	▲7340	△7052	▲7053	卷15						.6811	▲7340	△7052	▲7053
卷17							▲7199	▲7734	▲7545	卷16							▲7199	▲7734	▲7545
卷19								▲7118	▲7369	卷17								▲7118	▲7369
									▲7730	卷19									▲7730

【中】

卷11	卷12	卷13	卷14	卷15	卷16	卷17	卷19	卷20	卷22	卷23	卷24	卷25	卷26	卷27	卷28	卷29	卷30	卷31	他卷				
卷11	卷22	卷23	卷24	卷25	卷26	卷27	卷28	卷29	卷30	卷31	他卷	卷1	卷2	卷3	卷4	卷5	卷6	卷7	卷9				
卷12	▼5225	▼5053	▼5630	▼5420	▼5637	▼5338	▼5515	▼5347	▼5234	▼5708	卷1	▼5225	▼5053	▼5630	▼5420	▼5637	▼5338	▼5515	▼5347	▼5234	▼5708	.6796	
卷13	▼5301	▼5198	▼5794	▼5478	▼5892	▼5555	▼5516	▼5574	▼5437	▼5892	卷2	▼5301	▼5198	▼5794	▼5478	▼5892	▼5555	▼5516	▼5574	▼5437	▼5892	卷2	.6954
卷14	▼5186	▼5179	▼5663	▼5492	▼5612	▼5389	▼5478	▼5387	▼5283	▼5794	卷3	▼5186	▼5179	▼5663	▼5492	▼5612	▼5389	▼5478	▼5387	▼5283	▼5794	卷3	.6925
卷15	▼5639	▼5567	.6327	▼5825	.6149	.6011	.6042	.6009	▼5809	.6517	卷4	▼5639	▼5567	.6327	▼5825	.6149	.6011	.6042	.6009	▼5809	.6517	卷4	.7648
卷16	▼5746	▼5836	.6504	.6145	.6513	.6308	.6304	.6325	.5955	.6727	卷5	▼5746	▼5836	.6504	.6145	.6513	.6308	.6304	.6325	.5955	.6727	卷5	.7580
卷17	▼4953	▼5123	▼5792	▼5567	▼5534	▼5400	▼5523	▼5422	▼5078	.6001	卷6	▼4953	▼5123	▼5792	▼5567	▼5534	▼5400	▼5523	▼5422	▼5078	.6001	卷6	.7206
卷19	▼4888	▼5086	▼5900	▼5609	▼5741	▼5632	▼5627	▼5722	▼5266	.6176	卷7	▼4888	▼5086	▼5900	▼5609	▼5741	▼5632	▼5627	▼5722	▼5266	.6176	卷7	.7167
卷20	▼5087	▼5154	▼5853	▼5782	.6046	▼5779	▼5724	▼5894	▼5619	.6147	卷9	▼5087	▼5154	▼5853	▼5782	.6046	▼5779	▼5724	▼5894	▼5619	.6147	卷9	.6839
卷21	▼5880	▼5901	.6726	.6382	.6606	.6520	.6403	.6531	.6337	.6946	卷10	▼5880	▼5901	.6726	.6382	.6606	.6520	.6403	.6531	.6337	.6946	卷10	.7525
卷22	▼5418	▼5505	.6181	.5928	▼5889	▼5801	▼5894	▼5708	▼5463	.6429	卷11	▼5418	▼5505	.6181	.5928	▼5889	▼5801	▼5894	▼5708	▼5463	.6429	卷11	.7359
卷23	▼5663	▼5773	.6679	.6112	.6398	.6212	.6429	.6264	.5971	.6915	卷12	▼5663	▼5773	.6679	.6112	.6398	.6212	.6429	.6264	.5971	.6915	卷12	.7974
卷24	▼5136	▼5306	.6070	▼5600	▼5907	▼5809	▼5816	▼5805	▼5559	.6361	卷13	▼5136	▼5306	.6070	▼5600	▼5907	▼5809	▼5816	▼5805	▼5559	.6361	卷13	.7313
卷25	▼5545	▼5866	.6614	.6056	.6402	.6263	.6260	.6304	.6020	.6870	卷14	▼5545	▼5866	.6614	.6056	.6402	.6263	.6260	.6304	.6020	.6870	卷14	.7794
卷26	▼5405	▼5463	.6389	▼5741	.6083	.6060	.5971	.6007	.5958	.6554	卷15	▼5405	▼5463	.6389	▼5741	.6083	.6060	.5971	.6007	.5958	.6554	卷15	.7274
卷27	.5958	.6335	△7056	.6410	▲7472	△7140	.6884	△7258	△6972	▲7473	卷16	.5958	.6335	△7056	.6410	▲7472	△7140	.6884	△7258	△6972	▲7473	卷16	.7821
卷28	▼5539	▼5705	.6399	.6045	.6388	.6204	.6157	.6344	.6004	.6829	卷17	▼5539	▼5705	.6399	.6045	.6388	.6204	.6157	.6344	.6004	.6829	卷17	.7778
卷29	.6353	.6504	▲7437	.6646	▲7448	△7259	▲7402	△7282	△6981	▲7727	卷19	.6353	.6504	▲7437	.6646	▲7448	△7259	▲7402	△7282	△6981	▲7727	卷19	.8014
卷30	.6020	.6232	△7039	.6439	△7003	.6769	.6884	.6864	.6472	▲7415	卷20	.6020	.6232	△7039	.6439	△7003	.6769	.6884	.6864	.6472	▲7415	卷20	.8218
卷31		▼5609	.6356	▼5609	▼5869	.5988	.6170	▼5862	▼5813	.6149	卷22		▼5609	.6356	▼5609	▼5869	.5988	.6170	▼5862	▼5813	.6149	卷22	.6166
卷32			.6508	.6568	.6801	.6882	.6791	.6946	.6147	.6722	卷23			.6508	.6568	.6801	.6882	.6791	.6946	.6147	.6722	卷23	.6460
卷33				.6594	△7033	△7307	△7292	△7108	△6973	▲7475	卷24				.6594	△7033	△7307	△7292	△7108	△6973	▲7475	卷24	.7330
卷34					.6667	.6507	.6716	.6763	.6003	.6689	卷25					.6667	.6507	.6716	.6763	.6003	.6689	卷25	.6783
卷35						▲7656	▲7341	▲7874	△7056	▲7457	卷26						▲7656	▲7341	▲7874	△7056	▲7457	卷26	.7250
卷36							▲7294	▲7718	△7125	▲7604	卷27							▲7294	▲7718	△7125	▲7604	卷27	.7012
卷37								▲7411	.6628	△7341	卷28								▲7411	.6628	△7341	卷28	.7128
卷38									△7043	▲7491	卷29									△7043	▲7491	卷29	.7079
卷39									△7166		卷30									△7166		卷30	.6672
卷40											卷31											卷31	.7657

【右】

数値はすべて 1未満であり，小数第5位以下を切り捨てる。表では，整数部 0 を表示せず，また，すぐ下に記す意図で小数点の一部を ▲ △ ▼ ▽ に置き換えている。

表で例えば巻6について読むと，表【左】の横軸「巻6」を上から下にたどって巻1～巻5との類似度 .6713 ～ .6615 を得，縦軸「巻 6」を左から右にたどって巻7～巻10との類似度 .7406 ～ .6383 を得，表【中】の縦軸「巻 6」を左から右にたどって巻11～巻20との類似度 .7070 ～ .6757 を得，表【右】の縦軸「巻 6」を左から右にたどって巻22～巻31との類似度 .4953 ～ .6001 を得る。

表【右】の最右列「他巻」は，例えば，巻6に対して，他の全27巻を一体のものとして見なし，巻6と他全巻との間で類似度 .7206 を算出している。

類似度全378個の数値は，四分位に関係させて大きいほうから並べると，次のようである。

第 1位	0.7874 (巻26－巻29間)	第 94位	0.6968 (巻5－巻19間)
第 95位	0.6958 (巻3－巻6間)	第189位	0.6413 (巻10－巻15間)
第 96位	0.6958 (巻7－巻9間)	第283位	0.5955 (巻5－巻30間)
中央値	0.6411 (計算値)	第378位	0.4888 (巻7－巻22間)
第190位	0.6410 (巻16－巻25間)		
第284位	0.5928 (巻11－巻25間)		
第285位	0.5913 (巻1－巻9間)		

全378個のうちで，大きいほうの約10%＝38個，0.7341 (巻26－巻28間) 以上について，小数点を ▲ に置き換える。逆に小さい 38個，0.5534 (巻6－巻26間) 以下について，小数点を ▼ に置き換える。▼ は，表【右】上半分にまとまる。表【右】上半分は小さいもののまとまりでもあって，▼ に次ぐ56個も小数点を ▽ に置き換えてみている。▼ と ▽ との合計94個は，つまり小さいほうの25%であり，表【右】上半分に集中するほかには，表【右】下半分で巻17の2個，巻22の5個，表【左】で ▽ の最大値を示す巻1－巻9間に見えるにとどまる。

値が大きいほうはまとまりが緩い。10%～25%で小数点を △ に置き換えると，▲ および △ は，表【中】下半分に多い。また，表全体で，下の斜辺に沿って多く見え，配置に近い巻の間で似通うということを表すようである。

今昔物語集の部と関係させて類似度の数値を見ると，次のようである。

表【左】で	天竺震旦部内 (巻1～巻10)	0.7677 ～ 0.5913
表【中】上1/2で	本朝仏法部×天竺震旦部	0.7427 ～ 0.5964
表【中】下1/2で	本朝仏法部内 (巻11～巻20)	0.7831 ～ <u>0.6462</u>
表【右】上1/3で	本朝世俗部×天竺震旦部	<u>0.6946</u> ～ <u>0.4888</u>
表【右】中1/3で	本朝世俗部×本朝仏法部	<u>0.7727</u> ～ 0.5136
表【右】下1/3で	本朝世俗部内 (巻22～巻31)	0.7874 ～ 0.5609

表【中】下1/2の本朝仏法部内で，範囲の下限を 0.6462 (上記中央値より上の第179位) として，まとまりがよい。巻16・巻19・巻20は，他の部の諸巻との間でも ▼ ▽ を見せない。逆に，表【右】上1/3の本朝世俗部と天竺震旦部とで，上限を 0.6946

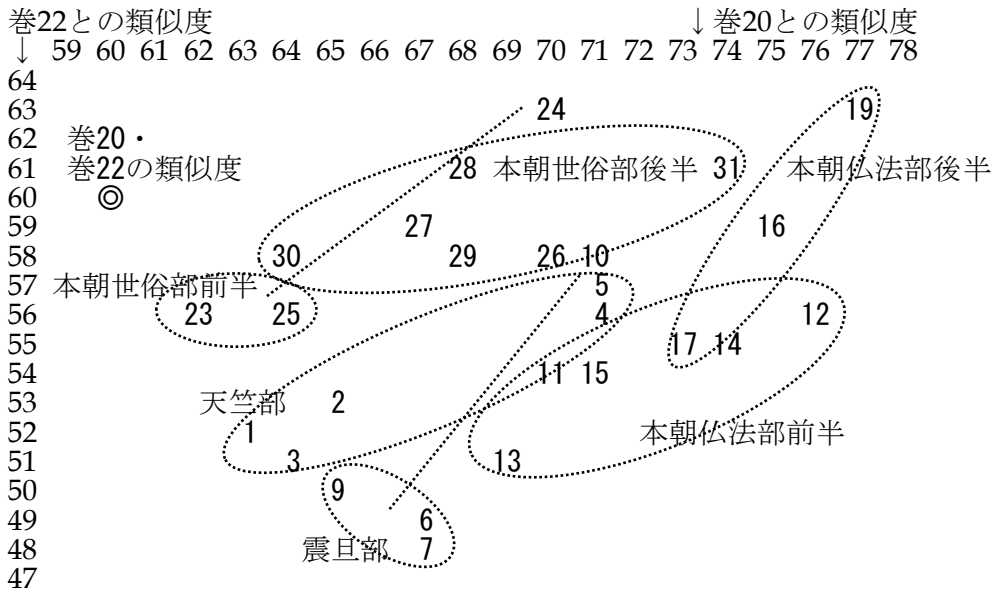
(上の四分位に近い第97位)として隔たりが大きく、巻22・巻23・巻25は他巻との間でも ▲ △ を見せない。

各巻が他全巻との間に示す類似度を、大きいほうから並べて四分すると、次のようである。類似度そのものは四分位相当のところのみに挙げ、他は巻次で示す。

0.8218	巻20	巻19	巻12	巻16	巻14	巻17	巻31	0.7657
0.7648	巻4	巻5	巻10	巻11	巻24	巻13	巻15	0.7274
0.7250	巻26	巻6	巻7	巻28	巻29	巻27	巻2	0.6954
0.6925	巻3	巻9	巻1	巻25	巻30	巻23	巻22	0.6166

本朝仏法部はこの上半分にまとまり、特にその後半(巻16～巻20)が上位に並ぶ。逆の最小は、本朝世俗部のうちの巻22・巻23などである。

いま、他全巻との類似度の最大・最小の巻20・巻22を軸として、それぞれとの類似度で各巻を位置づけると、次の図のようである。類似度は小数第1・2位による。



◎ は巻20—巻22間の類似度であり、この図の趣旨からすると、巻20の位置は ◎ の右の図外、巻22の位置は ◎ の上の図外、それぞれ目盛り100のところである。軸とする巻を別に採ると、諸巻の位置が違って見えるが、根本的でなく、立ち入らない。

今昔物語集の研究を回顧することは省略するが、諸巻の間の近さ・遠さとしてここに見たところに、意外性はないであろう。その穏当な結果によって、漢字の類似度という見かたも無益なものではないことになると思われる。

本稿の本体は以上である。以下、注釈を加える。宮島の類似度そのもの、新日本古典文学大系『今昔物語集』の本文の処理、また、特に漢字の処理、の3点である。注釈のあとに、部の中の類似度を算出してみる。

## 2 宮島達夫の類似度

### 2. 1 基本

宮島達夫の類似度は、宮島（1970）で提起された概念である。二作品の類似度は、単位を標目で整理し、標目ごとに作品それぞれにおける出現比率を比べて、大きくないうほうを累計した数値である。——これが宮島の規定である。

例えば、次の二作品A・Bにおいて、単位を漢字とするとき、

A 物ナドへ行〈ゆ〉クニモ、妻子〈めこ〉ニテモ強〈あなが〉チニ不審〈いぶか〉シトハ不思〈おもふ〉マジキ也。 ……単位数10, 標目数 9

B 身ヲ棄テテ行〈おこな〉フト云ヒ乍ラモ、無下ニ、不知〈しら〉ザラム所ニハ不可行〈ゆくべから〉ズ。 ……単位数13, 標目数11

作品双方に出現する標目の単位行について出現比率を算出すると、

「行」Aで  $1/10 < B$ で  $2/13$  類似度に採用する数値  $1/10=0.1000$

「不」Aで  $2/10 > B$ で  $2/13$   $2/13=0.1538$

であり、この二作品の類似度は、数値の合計 0.2538 である。

ここに作品としたものは、ともに今昔物語集の一節であるが、数値の算出にあたり、いま、訓みもカナとの関係も顧みていない。「行」が「ゆく」か「おこなふ」かを区別し、また、「不」が、熟字「不審」の一部であるか、「まじ」や「ず」に対応するか、ということを区別して、つまり単位・標目を別に設けるならば、数値は当然に異なってくる。なお、〈 〉内は、直前の漢字幾つかに対して校注者が施した訓である。

類似度は、直感的には、単位総数（延べ）が等しい二作品において、標目を同じくする単位をできるだけ多く、ただし二作品で同数、採り上げ、採り上げた単位数を単位総数で除した結果である。次の作品Cは、単位総数10であって作品Aと等しく、3箇所で重なるが、

C 騒シク、物念ジ不為〈せ〉ザラム人ハ、必ズ此〈か〉ク弊〈つたな〉キ也。

Aの側の「不」2箇所のうちの1箇所は、Cと重ならないと扱う。作品A・Cの類似度は、 $3 = 「物」1 + 「不」1 + 「也」1$  を単位総数 10 で除した 0.3 である。

単位総数が異なる二作品の間では、単位総数を等値にすればよい。例えば公倍数まで拡張するならば、上の作品A・Bのばあい、Aを13回繰り返し、Bを10回繰り返して、単位総数をそれぞれ130とし、「行」A13 対 B20, 「不」A26 対 B20 から、採り上げる単位数  $33 = 「行」13 + 「不」20$  を得、それを単位総数 130 で除して、数値 0.2538 が算出される。「行」でAよりBに多い単位数  $7 = 20 - 13$ , 「不」でAに多い単位数  $6 = 26 - 20$  は、類似度のためには採り上げられない。

公倍数と称するのはためられるものの、そのような一つとして、あらゆる作品に成立する 1 がある。すなわち、どのような作品も単位総数（延べ）を 1 として正規

化し、出現頻度を出現比率とするならば、宮島の最初の規定に至ることになる。

二作品の単位総数を等値にするには、単位総数が大きいほうから、小さいほうと同数になるまで単位を抽出する、などということも考えられる。石井（2019）に実例を示した。

類似度の数値は 0 以上 1 以下の範囲にあり、それが大きいほど、二作品は似ていると考えられる。ただし、ある数値を超えれば類似しているといった、絶対的な基準があるわけではない。三つ以上の作品の間で数値を算出して、この作品間のほうが他の間より似ている・似ていないと、相対的に見ることができるにとどまる。

宮島は、類似度の概念に至る前に、標目数および単位数による共通性を計算している。上の作品 A・B に即して見るならば、次のようである。

標目で  $0.1111 = \frac{\text{共通する標目}2}{(A\text{の標目}9 + B\text{の標目}11 - \text{共通分}2)}$

単位で  $0.4375 = \frac{\text{共通する標目の単位数つまり上の } \underline{\text{数}}7}{(A\text{の単位}10 + \overline{B\text{の単位}13} - \text{共通分}7)}$

単位による計算で、除数から共通分を減じない方法もあるであろう。また、そもそも、一方からすれば共通性が高く見えるが、他方からはそうでもないといった、非対称的に捉える方法もある。二作品の共通性あるいは距離を測る方法は、多様に提案されている。類似度に関連する概念に立ち入ることは、ここでは避けたい。

## 2. 2 拡張

作品群のうちで、ある一作品を特徴づけるのに、他の作品すべてが一体であるとして、一作品と他全作品との間の様相を見る、という手法がある。宮島・近藤（2011）は、そのような手法であるキルガリフの式によって、古典文学諸作品の特徴語というものを論じている。

宮島（2015）は、万葉集全20巻の巻間で用語の類似度を算出し、それとともに、各巻について、その巻と他19巻との類似度を算出している。万葉集全20巻が、上に言う作品群の全体であり、各巻が各作品であるということである。

本稿の1「類似度一覧」の表に「他巻」として設けた欄は、その数値の左の各巻と他の27巻全体との間の類似度を、示している。

### 3 今昔物語集

#### 3. 1 本文の設定

本稿は、宮島・他(2014)『日本古典対照分類語彙表』に関係する。この語彙表は、宮島(1971)『古典対照語い表』を拡張したものであり、拡張の過程で、今昔物語集を取り上げることも検討した。諸作品を扱うには、先行する総索引を利用して、今昔物語集でも、馬淵・有賀(1982)『今昔物語集自立語索引』および馬淵(1984)『今昔物語集漢字索引』に至る馬淵和夫監修・編集の一連の各巻文節索引があった。しかし、今回は索引を自ら作成することを試みた。

今昔物語集の本文は、新日本古典文学大系(岩波書店)のいずれも第1刷による。

今昔物語集 一～五 新大系第33～37巻 1993年 5月28日～1999年 7月28日

校注・今野達・小峯和明・池上洵一・森正人

調査対象とする部分は、『日本古典対照分類語彙表』の全般の方針に従って、各語おおむね「今ハ昔、……トナム語り伝ヘタルトヤ。」とある言わば本体である。各巻・各語の標題は、採り上げない。また、本文中の注記も調査対象としない。次の《 》内であり、\ を改行位置とする割り書きになっている。

(巻1語10=1冊35面10行) 蓮花比丘〈れんぐゑびく〉《又華色〈くゑしき〉トモ云〈いふ〉。\\一人二名。》ト云フ羅漢〈らかん〉ノ比丘尼〈びくに〉ノ

( )内の=以下は、新日本古典文学大系における所在であり、行は空白も含めて数えて、面当たり第1・2冊で14行、第3・4・5冊で16行である。

本文における空白は、単純に無視し、長さを見たり文脈・文字を想定したりすることをしない。本文が中断して未完結になったもの(巻7語43=2冊159面8行など)も、あるままに扱う。要するに、文字として残されたものはすべて検討し、しかも、欠損によって結果に歪みが生ずるであろうが、その歪みは考慮しない。

本文は、凡例に次のように言われ、底本との間に距離がある。摘要する。

- ・ 底本では付属語や活用語尾を小字の片仮名で二行に割書きしているが、通常の手書き方に改めた。
- ・ 漢字は原則として通行の字体に改め、常用漢字表にある文字は新字体を用いた。
- ・ 特殊な合字や連字体等は通行の文字に改めた。「廿」「卅五」等は「二十」「三十五」等に改めた。
- ・ 明らかな誤字・衍字等は訂正し、必要に応じて注を付した。

今日、受け容れられやすく、流布しやすい本文である、ということであろう。

この本文のうちの漢字を取り上げ、漢字とカタカナとの区別は、今日感覚に従う。万葉仮名としての出現は、見た目判断して漢字として扱う。

(巻11語27=3冊75面7行) 「志之〈こころざし〉」の捨仮名

(巻24語15=4冊410面14行) 「父古萱〈こそ〉」の助詞

本文について、たまたま気づいた箇所があるが、改めない。底本との関係を精査するよりは、流布しているままの姿を受け止めようとするからである。

(巻12語34=3冊171面4行) 仏教ヲ信ジテ出家ノ心有リ。

別巻索引でも(396面左欄)、抄出された岩波文庫でも(本朝部上冊159面10行)、「仏教」であるが、京都大学本で(影印下冊96面11行)「仏法」である。校訂は記されず、恐らく「仏法」であるべきである。影印については、本稿末尾に付記する。

(巻7語30=2冊143面8行=京都大学本影印上冊349面10行) 煖〈あたた〉カ

(巻12語30=3冊157面13行=京都大学本影印下冊79面1行) 柔煖〈こうなん〉

(巻15語36=3冊434面15行=東京大学本影印4冊565面2行) 柔煖〈こうなん〉

この3件は、底本で字形が「左火・右上而・右下而」であるが、「煖」との関係は未詳である。次の「◇」は、字形「左火・右上而・右下而」で本文が作成されている。

(巻5語20=1冊450面6行=京都大学本影印上冊246面4行) ◇〈やは〉ラ

### 3. 2 繰返し符号の処理

本文の漢字を集計することにつき、本文そのままではないところもある。一つは、繰返し符号「と」「\」「ゞ」「ゝゝ」「ゞゝ」を、そのままとせず、漢字・カナに置き換えることである。ここで、「ゝゝ」「ゞゝ」は、くの字点である。

「と」「\」は大半が直前の文字を繰返し、「ゞ」は濁のカナで繰返す。そのようには簡単でない事例を、幾つか挙げる。本文を → 以下のように扱い、繰返し符号を下線部のように変える。

(巻17語47=4冊92面2行) 取レドモとレドモ → 取レドモ取レドモ

(巻28語4=5冊200面8行) 君達ヤ有と → 君達ヤ有君達ヤ有

(巻5語27=1冊464面9行) 今ヤ喰とフ → 今ヤ喰今ヤ喰フ

(巻24語31=4冊443面13行) 参ヌヤとと → 参ヌヤ参ヌヤ

(巻26語5=5冊21面15行) 押廻シとと → 押廻シ押廻シ

(巻12語12=3冊119面4行) 我レヲ取レ、ととと

→ 我レヲ取レ、我レヲ取レ

(巻26語8=5冊37面16行) 我モ得ムゝゝ → 我モ得ム我モ得ム

(巻28語29=5冊246面6行) 何ノ音ヅ、ゝゝヅ → 何ノ音ヅ、何ノ音ヅ

(巻26語24=5冊89面13行) 今ヤ、物語シ畢テ出ル、ゝゝ

→ 今ヤ、物語シ畢テ出ル、今ヤ、物語シ畢テ出ル



## 4 漢字

### 4. 1 漢字を扱うこと

漢字を単位・標目として類似度を見た調査に、浅野（1994）および石井（2014）がある。浅野は幾つかの作品の間を見、石井は一作品の巻の間を見る、という違いはあるものの、ともに、漢字が表語文字であって、その出現の様相が作品なり巻なりの内容の異同を反映するということを期待している。

漢字の表語と言っても、今昔物語集では「多語」を担うと見てよく、種類の要素に対応する。例えば、漢字「許」は次のような現れかたをする。

- ・ 助詞「ばかり」および派生副詞「しばしばかり」「いかばかり」
- ・ 名詞「もと」および複合名詞「みもと」・動詞「おもとだつ」など
- ・ 動詞「ゆるす」および転成名詞「ゆるし」「ゆるされ」
- ・ 漢語「御裁許〈ごさいきよ〉」の構成要素

また、漢字「蜜」は、花や実が含む液を指すものは、漢語「蘇蜜」などにとどまり、他は、普通は漢字「密」で表される「頭蜜〈けんみつ〉」「蜜〈きびし〉」「蜜〈ひそか〉」「蜜音〈しのびね〉」などである。こうした、多語あるいは通用の性格は、他の漢字でも見られるところである。しかし、例えば「許」を「計」「元」「免」との間で調整したり、「蜜」を「密」に改めたりすることは、しない。漢字は、相応の混沌を抱えたものとなる。

### 4. 2 異体関係の整理

漢字の集計に設定する単位は、文章に現れている一字一字である。問題があるようにはほとんど感じられないが、本文の凡例に、「廿」「卅」を「二十」「三十」に改めたと記され、また、（巻30語5=5冊410面6行）「血肉〈ちじし〉」は、注に、底本および諸本の一字「上一・中血・下白」を訂したという。本文そのままでは集計しないところとして、本文での一字を分字することもある。（巻5語13=1冊426面12行）「◇〈あなに〉ク」の「◇」は字形「上穴・下憎」であり、2字「穴」「憎」とする。また、（巻20語33=4冊286面6行）「◇部〈くさかべ〉ノ真眞〈まとじ〉」の「◇」は字形「上日・下下」であり、2字「日」「下」とする。こうした一方で、「曆」は「麻」「呂」に分字しない。

なお、分字されているものは合字しない。（巻5語21=1冊451面3行）「穴井〈ししのあな〉」、（巻29語36=5冊379面7行）「草馬〈めうま〉」は、「辵」「左馬・右草」を分字したと注で解説する。本稿では、本文のとおりに2字ずつとする。

漢字の集計での標目は、本文そのままではなく、異体関係を整理した結果になる。

本文の凡例で「通行の字体」「新字体」によると言われるものは、ある漢字について底本で旧字体ばかりであるならば、本文での単純な置き換えである。本文の漢字「予」は、旧字体に「予」「豫」があるものの、専ら震旦・日本の地名・人名を表し、本文の底本で恐らくすべて「豫」の形である。底本に従って標目の形を「豫」としても、見た目が異なるが、標目の数は一つのみで変わらない。漢字「着」「著」は、底本で形を判別することは難しいが、新字体における「著」の意味のものは現れず、本文は「着」で通している。漢字「弁」は、常用漢字表で旧字体「辨」「瓣」「辯」が示され、そのほかに字体「辨」もあるが、本文ではすべて「弁」であり、実は恐らく底本でもすべて「弁」の形である。つまり、漢字「弁」の関係では、底本ですでに異体の関係が整理され、本文でもそれを踏襲していることになる。

底本に新字体も旧字体もあるならば、標目の整理に相当することが、本文ですで行われていることになる。漢字「無」は、諸巻の底本で多く「无」の形で現れ、京都大学本では、全体にわたって専ら「无」である。他の底本では「無」の形も紛れ込み、(巻26語8=5冊36面9行=東京大学本影印6冊232面7行)「無〈なき〉物〈もの〉無〈な〉ク」は、底本東京大学本で、前が「無」、後が「无」である。また、漢字「余」は、旧字体に「余」「餘」があるものの、自称の「余」は現れず、残余を意味するものが、本文の底本に「余」「餘」双方の形で現れる。底本に従って標目を二つにしようところを、本文で一つに整理している。本文で底本の漢字が整理されていれば、集計で分離することはできない。

本文の漢字の異体関係は夥しくあって、ここに詳細を記すことはできない。上の3.1末尾に挙げた「煖」「左火・右上而・右下而」は、異体関係にあると見て、一つの標目のもとに統合する。次なども統合する。所在は一例である。

- ・ 「園」(巻2語2=1冊104面9行)「菌」(同=106面10行)
- ・ 「圧」(巻7語31=2冊148面1行)、垂が「广」(巻25語5=4冊505面2行)、  
「𠂔」(巻5語17=1冊436面4行)など
- ・ 「槌」(巻17語4=4冊10面2行)、右部分が「遣」(巻12語21=3冊138面8行)、  
「送」(巻29語39=5冊386面11行)「迷」(巻5語3=1冊399面13行)
- ・ 「嗟」(巻12語20=3冊133面12行)、右下部分が「日」(同=133面13行)
- ・ 「煙」(巻5語13=1冊427面3行)「烟」(巻2語23=1冊148面10行)
- ・ 「筋」(巻10語1=2冊284面14行)「上卍・下卍」(巻12語20=3冊133面13行)
- ・ 「総」(巻12語11=3冊117面11行)「捻」(巻9語28=2冊227面9行)

なお「惣」は統合せず、別字とする。

統合しないものものもある。その一例は次である。

- ・ 「悪」(巻9語3=2冊182面10行)「慄の右下が心」(巻9語4=2冊184面14行)

## 5 漢字の分量

今昔物語集の漢字の分量は、次のとおりである。巻ごとに、「文字」欄は、漢字延べとカナ延べとの合計である。句読点・鉤括弧等を含まない。「延べ」欄は漢字延べであり、「比率」欄は、文字延べに占める漢字延べの比率である。「異り」欄は、漢字異なりである。「高頻度上位20」は、出現頻度が高い1~20位の漢字を順位に従って並べ、最右欄「我然」は、高頻度20位のうちに「我」があれば「a」を記し、「然」があれば「b」を記す。最下行は、全巻の合計である。

	文字	延べ	比率	異り	高頻度上位20	我然
巻1	35507	19828	.5584	1262	給此云人仏事子其王我出不無有家也太時道奉	a
巻2	38957	21119	.5421	1173	此人其云仏王事有給也無我不子一家国見時生	a
巻3	29964	16488	.5502	1216	此給仏云人其王有不我事也大見子無時一羅国	a
巻4	33276	17558	.5276	1196	此人云其事給有不無也我王思見仏大国子一可	a
巻5	33310	17215	.5168	1245	此人云其事有王也不給国子我無見大一可思行	a
巻6	29859	16511	.5529	1232	此人云其有給事不也仏見我法無奉経一時来大	a
巻7	27819	15040	.5406	1263	此人云其事有経僧一不也見法我無所時思行亦	a
巻9	35443	18502	.5220	1367	云此人事其有不我也見無時母一然子語死君来	ab
巻10	35346	17683	.5002	1186	此云人事其有不也無給我見国思然子可王来行	ab
巻11	38359	21723	.5663	1399	其云人此給也有事大天寺見国無行是子仏然我	ab
巻12	40605	21779	.5363	1380	此人云其事有也不経給無見思僧法行来寺大我	a
巻13	33864	18398	.5432	1254	此人云経法其事花有不也無誦見来僧聖我所聞	a
巻14	37812	19888	.5259	1282	此云人事其有経僧不思也法見我無行花来心聞	a
巻15	36205	18947	.5233	1125	此云人事有其無也不仏見来聞生思僧心楽行間	
巻16	43426	21331	.4912	1263	此云人有事思其無不也見音来觀給女男行我然	ab
巻17	37985	20200	.5317	1234	此人其云事有僧也不藏地無我見思行善来薩聞	a
巻19	50241	24388	.4854	1367	此人有云事思其不見無也給行御心子然出来共	b
巻20	38916	20410	.5244	1436	此人云事有其不也無見思我给行然可来僧師大	ab
巻22	9868	5009	.5076	706	大御給此臣人其事有思申也無不子見家言納云	
巻23	15091	7471	.4950	863	此云有其人不思也事見様被来共行成女取大出	
巻24	38703	18168	.4694	1253	此人云有其事思不給無見也御行然聞女可今大	b
巻25	22896	12040	.5258	1107	此云其不有人事思守也然等無者国馬見可共兵	b
巻26	38101	18700	.4908	1214	此云人有其思事不見也無来然出男何行共物程	b
巻27	35696	16470	.4613	1108	此有人云其不思事見然無也行女来様物程何許	b
巻28	42471	20534	.4834	1317	此人云有事其不也見共然物思給無上出御何咲	b
巻29	42437	19712	.4645	1148	此有人云其不思事然見無也者共女物男行様取	b
巻30	17856	7892	.4419	768	此思云有不人其事見然無心女行也妻様何男物	b
巻31	31906	15286	.4790	1179	此人有其云不事思見也然行無来者女様共何可	b
全巻	951919	488290	.5129	3100	此人云其有事不也思見無給我然行来大可聞子	ab

漢字の比率は、天竺震旦部および本朝仏法部で高く、本朝世俗部で低い。この、天竺震旦部および本朝仏法部と、本朝世俗部との対立は、a と b との対立でも、おおむね見ることができる。

こうした数値も、それに支えられた類似度の数値も、漢字の単位・標目の設定に従った結果である。単位・標目が異なれば結果も異なり、端的に知られるものとして、浅野 (2011 p. 257) が示した数値を対比してみる。浅野は、底本は本稿と同じく京都

大学本ながら、日本古典文学大系の本文により、各語の標題を含む。浅野は他の数値も示しているが、ここでは省く。本稿というのは、上の再掲である。

		本稿	浅野		本稿	浅野
卷 2	延べ	21119	21573	異なり	1173	1190
卷 5	延べ	17215	17616	異なり	1245	1247
卷27	延べ	16470	17309	異なり	1108	1119
卷29	延べ	19712	20210	異なり	1148	1160

なお、浅野は、この4巻全体における高頻度文字を列挙する。本稿の28巻全体の上位20字の、浅野での順位を見ると、次のようである。9位までは一致する。

此 人 云 其 有 事 不 也 思 見 無 給 我 然 行 来 大 可 聞 子  
 1 2 3 4 5 6 7 8 9 11 10 13 15 12 14 19 22 25 38 16

単位・標目が異なり、数値が異なっても、大勢は細部まで似るかもしれない。

## 6 部間の類似度

部間の類似度を計算する。天竺・震旦・本朝を枠としながら、最初に挙げた巻間類似度を考慮して、諸巻をまとめる。二通りとする。値が大きいもの・小さいものは、小数点を ▲ △ ▼ ▽ に置き換える。漢字の分量も添える。

	震旦	仏前	仏後	世前	世後	他全	漢字延べ	比率	異り
天竺 卷 1～5	.7553	.7302	.7258	▼6523	▼6466	.7376	92208	.5391	1963
震旦 卷 6～10		△7864	.7658	▼6661	▼6733	.7958	67736	.5272	2096
本朝仏法前半 卷11～15			▲8106	.6872	.6861	.7998	100735	.5391	2050
本朝仏法後半 卷16～20				.7525	△7895	.8621	86329	.5061	1962
本朝世俗前半 卷22～25					▲8029	.7578	42688	.4931	1666
本朝世俗後半 卷26～31						.7443	98594	.4729	1917
	本朝仏法	本朝世俗		他全	漢字延べ	比率	異り		
天竺震旦 卷 1～10	▲7899	▼6895		.7681	159944	.5340	2484		
本朝仏法 卷11～20		.7556		.8336	187064	.5233	2357		
本朝世俗 卷22～31				.7369	141282	.4788	2139		

本朝仏法部はよくまとまり、前半は震旦部と近く、後半は世俗部後半と近い。本朝世俗部もよくまとまるが、天竺部・震旦部と遠い。この理解は、巻間類似度の一覧の後に示した図の解釈でありうる。

初めに記したことの繰り返しになるが、今昔物語集の研究は、説話の出典でも語句・語法の様相などでも緻密に行われていて、本稿のようなとらえかたは、大雑把に過ぎて不快に感じられるかもしれない。しかし、そうしたところに立ち返ってみることも、時にはよいであろう。

漢字の研究からすれば、漢字を簡便に見るのみでも、相応の結果は得られるということである。宮島の類似度は、その結果を得るのに貢献する。

- 参照論著 配列は第一著者の五十音順，同一著者では年代順，ただし一部不同。
- 浅野 敏彦 (1994) 「平安時代漢字文献対照漢字表」作成の試み  
——平安時代識字層が用いた漢字調査のてだてとして。  
大阪成蹊女子短期大学研究紀要 31 pp. 1-14。  
=浅野 (2011) pp. 146-168。
- 浅野 敏彦 (2011) 平安時代識字層の漢字・漢語の受容についての研究。  
和泉書院，研究叢書 415。  
本稿に特に関係するのは，浅野 (1994) の収録のほか，第三章付論であり，第二章付論とともにこの著作で初めて活字化された。二つの付論の標題等は次のようである。  
(第二章付) 『今昔物語集』天竺部 (巻一～五) の漢語語彙。  
同書 pp. 119-135。  
(第三章付) 『今昔物語集』巻二・五・二七・二九の漢字。  
同書 pp. 256-261。  
引用した数値は，著者のご教示によって一部を改める。
- 石井 久雄 (2014) 航米日録の漢字，その定量的概要。  
立命館言語文化研究 25.3 pp. 153-173。
- 石井 久雄 (2019) 『日本古典対照分類語彙表』および『雑誌用語の変遷』  
付論 宮島達夫の語彙の類似度について。  
国立国語研究所創立70周年・移管10周年公開講演会  
「言語史の計量的研究 宮島達夫の言語史研究」発表原稿追補。  
<http://kotonoha.art.coocan.jp/>。
- 馬淵 和夫・有賀嘉寿子 (1982) 今昔物語集自立語索引。  
笠間書院，笠間索引叢刊 39。
- 馬淵 和夫 (1984) 今昔物語集漢字索引。笠間書院，笠間索引叢刊 40。
- 宮島 達夫 (1970) 語いの類似度。国語学会，国語学 82 pp. 42-64。  
=宮島 (2019) pp. 31-62。  
類似度の説明は宮島 (2004) にもある。
- 宮島 達夫 (1971) 古典対照語い表。笠間書院，笠間索引叢刊 4。
- 宮島 達夫 (2004) 英語作品の用語の類似度。京都橘女子大学紀要 30 pp. 232-209。  
=宮島 (2019) pp. 63-91。
- 宮島 達夫・近藤明日子 (2011) 調査報告 古典作品の特徴語。  
計量国語学会，計量国語学 28.3 pp. 94-105。  
=宮島 (2019) pp. 134-150。
- 宮島 達夫・鈴木 泰・石井 久雄・安部 清哉 (2014)  
日本古典対照分類語彙表。笠間書院。

宮島 達夫 (2015) 万葉集巻別対照分類語彙表。笠間書院。

宮島 達夫 (2019) 言語史の計量的研究。笠間書院。

本文を確認するなどの必要に応じて、次などを参照する。

今昔物語集索引 新日本古典文学大系別巻 編・小峯和明 2001年 岩波書店

今昔物語集 天竺・震旦部 本朝部上 本朝部中 本朝部下 編・池上洵一

2001年 岩波文庫

鈴鹿本今昔物語集 ——影印と考証 上巻 下巻 序・安田章

1997年 京都大学学術出版会

京都大学貴重資料デジタルアーカイブ

<https://rmda.kulib.kyoto-u.ac.jp/item/rb00000125>

実践女子大学学術機関リポジトリ <http://id.nii.ac.jp/1157/00001357/> 以下

今昔物語集 一～六 東京大学国語研究室資料叢書1～6 解題・山口明穂

1984年～1986年 汲古書院

今昔物語集 一～五 日本古典文学大系22～26

山田 孝雄・山田 忠雄・山田 英雄・山田 俊雄 1959年～1963年 岩波書店

## 科学研究費補助金

今昔物語集の索引を作成するに当たり、科学研究費補助金の交付を受けた。

- ・ JSPS11610440 『今昔物語集』語彙の計量的研究

(1999～2001年度 基盤研究(C) 研究代表者・宮島達夫)

- ・ JSPS17520293 『今昔物語集』のKWIC作成の原理的問題とその意味・文法論的

利用の可能性 (2005～2007年度 基盤研究(C) 研究代表者・鈴木泰)

索引は未完成であり、成果を『日本古典対照分類語彙表』に取り込むことはできなかったが、本稿はその途中で得られた結実の一つである。

## 謝辞

本稿をまとめたのには、浅野敏彦氏・彦坂佳宣氏・辻加代子氏の慇懃に負うところが大きい。当初の目的は、宮島の類似度を算出することであった。しかし今昔物語集の内容と関係づけることに進めというのが、その慇懃であり、一步なりと踏み入ることができたとは思ふ。出来栄えはともかくとして、三先輩に感謝します。