

『日本経済新聞記事オープンコーパス』に対する メタデータと語義情報付与

加藤 祥
目白大学

浅原 正幸
国立国語研究所
masayu-a -at- ninjal.ac.jp

概要

本研究では『日本経済新聞記事オープンコーパス』に対して、メタデータとして新聞記事分類情報を付与した。さらに語義情報として分類語彙表番号を付与した。本研究で付与した情報は、『現代日本語書き言葉均衡コーパス』に含まれる新聞記事サンプルとの対照が可能である。分類情報の分布から『日本経済新聞記事オープンコーパス』は「経済」記事に特化していることが確認された。また、各新聞記事の「経済」記事との詳細な記事分類情報の分布や語義情報の分布を比較した。

1 はじめに

『日本経済新聞記事オープンコーパス』¹⁾は日本経済新聞の朝夕刊(2013年1~2月)96本の記事からなる、営利目的での利用を許可された新聞記事コーパスである。国立国語研究所にて、UniDic短単位形態論情報・長単位形態論情報・文節係り受け・Universal Dependenciesの付加情報を付与した。さらに、自己ペース読文法による文節単位の読み時間や、短単位動詞・長単位自立語・文節単位の印象評定情報も付与されている[1]。

本研究では同コーパスに、メタ情報として新聞記事分類情報と、語義情報として『分類語彙表』[2]に基づく分類番号を付与した。本稿では、『現代日本語書き言葉均衡コーパス』(BCCWJ)[3]の新聞記事分類情報[4]・語義情報[5]と対照し、それぞれの分布の異同について確認する。

2 メタデータの付与

新聞記事情報として分類情報や記事種別を付与した。これらの新聞記事の分類は、BCCWJの新聞記

事に付与された情報である[4]。まず、記事内容については、紀伊國屋書店『CD-HIASK』の分類を援用した記事情報(大)(小)を付与している。分類情報(大)として、総類・政治・経済・労働・文化・科学・社会・事件・スポーツ・国際・見出しの情報を付与した。さらに分類情報(小)として付録の表5に示すラベルを付与した。『日本経済新聞記事オープンコーパス』とBCCWJの各社の新聞記事分類情報(大)を表1に示す。また、一般記事とそれ以外を区別するために、記事種別1(付録表6)・記事種別2(付録表7)の情報を付与した。

サンプル数は少ないものの、『日本経済新聞記事オープンコーパス』は一般報道記事の分布がほとんどを占めており、案内や読み物記事などがほとんど含まれていない。なお、『日本経済新聞記事オープンコーパス』はランダムサンプリングを行ったが、公開する際に「人名の羅列からなる人事記事」・「死亡記事」は公開対象から外されている。

丸山ほか[6]によれば、BCCWNのPN(出版サブコーパス「新聞」)は、『全国新聞ガイド』(社団法人日本新聞協会発行)の「全国紙」5紙「ブロック紙」3紙および各地の有力な地方紙8紙の印刷総文字数の推計を母集団とし、新聞タイトルによる16分類×発行年(2001年から2005年)5分類の80層各層に含まれる全ページヘランダムに割り振った優先順位の高い順に、指定されたページに含まれる文章が抽出されている。サンプルには「全国紙」4紙(日本経済新聞は含まれない)「ブロック紙」3紙「地方紙」7紙が収録されている。BCCWJに含まれる新聞記事サンプルは、各新聞記事における記事分布の実態に近いものと考えられる。表1の分布から、『日本経済新聞記事オープンコーパス』の記事分布が経済に特化していることが確認された。

表2には「経済」記事の小分類について示す。『日

1) <https://nkbb.nikkei.co.jp/alternative/corpus/>

表1 記事分類情報(大)に基づく分布(記事数)

分類(大)	総類	政治	経済	労働	文化	科学	社会	事件	スポーツ	国際	見出し	その他	総計
日経新聞	3 3.1%	20 20.8%	47 49.0%	2 2.1%	4 4.2%	3 3.1%	2 2.1%	5 5.2%	10 10.4%	0 0.0%	0 0.0%	0 0.0%	96 100.0%
朝日新聞	12 2.3%	65 12.5%	67 12.9%	15 2.9%	89 17.1%	24 4.6%	91 17.5%	39 7.5%	61 11.7%	46 8.8%	5 1.0%	1 0.2%	520 100.0%
読売新聞	22 4.2%	83 15.9%	52 9.9%	16 3.1%	45 8.6%	39 7.5%	101 19.3%	28 5.4%	102 19.5%	28 5.4%	7 1.3%	0 0.0%	523 100.0%
毎日新聞	47 9.8%	44 9.2%	54 11.3%	7 1.5%	55 11.5%	11 2.3%	79 16.5%	53 11.1%	78 16.3%	44 9.2%	3 0.6%	4 0.8%	479 100.0%
産経新聞	25 5.5%	36 7.9%	51 11.2%	5 1.1%	59 13.0%	23 5.1%	89 19.6%	57 12.6%	74 16.3%	34 7.5%	0 0.0%	1 0.2%	454 100.0%
ブロック紙	68 5.9%	167 14.4%	131 11.3%	18 1.5%	154 13.3%	55 4.7%	220 18.9%	123 10.6%	142 12.2%	74 6.4%	6 0.5%	4 0.3%	1162 100.0%
地方紙	166 6.8%	303 12.4%	311 12.7%	45 1.8%	302 12.3%	129 5.3%	464 19.0%	202 8.3%	364 14.9%	138 5.6%	3 0.1%	20 0.8%	2447 100.0%
(BCCWJ-WLSP)	33 7.9%	64 15.2%	57 13.6%	7 1.7%	57 13.6%	12 2.9%	40 9.5%	46 11.0%	65 15.5%	30 7.1%	9 2.1%	0 0.0%	420 100.0%

表2 記事分類情報(小)に基づく「経済」記事の分布(記事数)

分類(小)	経済(小)	金融	貿易	商業	鉱工業	エネルギー	農林漁業	運輸交通	情報通信	土木建設	経済(大)
日経新聞	5 11%	14 30%	2 4%	8 17%	8 17%	4 9%	1 2%	2 4%	1 2%	2 4%	47 100%
朝日新聞	20 28%	19 27%	3 4%	6 8%	8 11%	4 6%	2 3%	6 8%	3 4%	0 0%	71 100%
読売新聞	4 7%	18 32%	3 5%	7 13%	3 5%	2 4%	2 4%	6 11%	7 13%	4 7%	56 100%
毎日新聞	6 11%	24 42%	5 9%	7 12%	3 5%	1 2%	3 5%	2 4%	3 5%	3 5%	57 100%
産経新聞	6 11%	19 36%	2 4%	8 15%	5 9%	2 4%	3 6%	5 9%	1 2%	2 4%	53 100%
ブロック紙	51 33%	34 22%	2 1%	18 12%	3 2%	7 4%	8 5%	21 13%	7 4%	5 3%	156 100%
地方紙	53 16%	79 24%	9 3%	58 17%	36 11%	7 2%	35 11%	19 6%	25 8%	11 3%	332 100%
(BCCWJ-WLSP)	13 22%	13 22%	4 7%	5 9%	6 10%	3 5%	0 0%	10 17%	4 7%	0 0%	58 100%

本経済新聞記事オープンコーパス』は若干「金融」「商業」「鉱工業」などが多く、「農林漁業」「情報通信」が少ないという傾向は見られているが、「経済」記事の具体的な内容については、各新聞記事と『日本経済新聞記事オープンコーパス』において特徴的な差異があるというよりは、サンプルの取得された際の時事的な影響が考えられる。

3 語義情報の付与

『日本経済新聞記事オープンコーパス』96記事・33346語に対し、『分類語彙表』の分類番号を手により付与した。比較対象として、『現代日本語書き言葉均衡コーパス』[3]に『分類語彙表』の分類番号を付与したBCCWJ-WLSPの語義分布(類・部門:分類番号の小数点二桁まで)の頻度を示したものを表3に示す。語義情報の分布は、BCCWJのPNと『日本経済新聞記事オープンコーパス』でほぼ同様といえる。経済記事の割合に差は見られるものの、

経済記事であることによる語義の分布に差はない。そこで、中項目(分類番号の小数点二桁まで)についても確認を行った。表4に示す。

まず、1.37(体-活動-経済)に差が大きいのは「経済」分野の記事量により、「経済」分野の記事量との関連により、1.15(体-関係-作用)で1.1540(上下)が影響しているほか、1.19(体-関係-量)のような数値において差が生じている。同様に、1.24(体-主体-成員)は「社長」「部長」などの役職名、1.25(体-主体-公私)は地名、1.26(体-主体-社会)は企業名、1.27(体-主体-機関)は「政府」「連合」「組合」などの機関名の差ともなる。なお、「経済」分野に限定して対照しても、.37(経済)と.19(量)は『日本経済新聞記事オープンコーパス』がBCCWJのPNを上回り(1.37ではBCCWJ:3.7%,日経:4.8%、1.19ではBCCWJ:11.3%,日経:15.5%)、経済と数量に関する語義の語彙が多く用いられているといえる。

なお、2.57(用-自然-生命)における差は『日本経

表3 語義情報の分布(類・部門: 分類番号の小数点以下一桁まで)

	日経 (割合)	BCCWJ-WLSP (割合) PN 経済のみ	BCCWJ-WLSP (割合) PN	BCCWJ-WLSP (割合) PB, PM, PN
1.1 体-関係	6750 20.20%	3967 21.65%	21812 18.56%	50863 14.70%
1.2 体-主体	2963 8.90%	1504 8.21%	11010 9.37%	24261 7.00%
1.3 体-活動	4120 12.40%	2414 13.17%	12743 10.84%	28198 8.10%
1.4 体-生産物	506 1.50%	416 2.27%	2063 1.76%	7418 2.10%
1.5 体-自然	312 0.90%	159 0.87%	1691 1.44%	6757 1.90%
2.1 用-関係	1352 4.10%	755 4.12%	4893 4.16%	17559 5.10%
2.3 用-活動	1565 4.70%	849 4.63%	5998 5.10%	20357 5.90%
2.5 用-自然	14 0.00%	10 0.05%	126 0.11%	584 0.20%
3.1 相-関係	1138 3.40%	623 3.40%	3873 3.29%	15625 4.50%
3.2 相-主体	2 0.00%	0 0.00%	0 0.00%	7 0.00%
3.3 相-活動	174 0.50%	80 0.44%	648 0.55%	2682 0.80%
3.5 相-自然	14 0.00%	7 0.04%	226 0.19%	541 0.20%
4. 他	82 0.20%	67 0.37%	408 0.35%	2497 0.70%
ラベルなし	14354 43.00%	7476 40.79%	52199 44.41%	169745 48.90%
合計	33346 100.00%	18327 100.00%	117543 100.00%	347094 100.00%

『日本経済新聞記事オープンコーパス』において「死亡記事」が削除されたことによる影響である。

また、BCCWJのPNと『日本経済新聞記事オープンコーパス』には、文体的な違いが見られている。語義的な違いではないが、たとえば、文長は大差がない(「。」の割合はBCCWJ3.1%、日経3.6%、以降BCCWJと『日本経済新聞記事オープンコーパス』の経済記事における語彙の出現割合を示す)印象であるが「、(読点)」はBCCWJで4%であるのに対し、日経は3%であり、読点の使用割合の違いが散見される。特に、BCCWJで見られる(1)のような文は、日経では句点で2文にされている。

- (1) 平たく言えば、少なくとも10社か15社、市場で喧伝されている問題企業に産業再生委が思い切ったメスを入れ、産業再編の具体像を示す、それによって日本が本格的に変わり始めたというシグナルを市場に与えることが必要だ。(BCCWJ, サンプルID: PN1b.00004)

2.12(用-関係-存在)の差では「居る」「有る」「成る」などの頻度差の影響が見られる。「ている」「て」BCCWJ:1.5%、日経:1.2%、「居る」BCCWJ:0.7%、日経:0.5%をはじめ、「となる」「成る」BCCWJ:0.6%、日経:0.4%などの現在や未来のアスペクトは日経で使われにくく、「受け取る」「参入した」のような言い切りの文末が多い。同様に、「とされる」「れる」BCCWJ:0.4%、日経:0.2%「という」「言う」BCCWJ:0.4%、日経:0.1%などの婉曲表現の割合も異なることから、日経サンプルではBCCWJに含まれる新聞サンプルよりも断定的な印象が強まると考えられる。関連して、1.20(体-主体-人間)の差は、敬称の「さん」が『日本経済新聞記事オープンコーパス』において用例がなく「氏」に限定され

ていたことによる。語彙的な頻度差として明確に現れるほどではないが、表現の違いから、専門的な印象が生じている可能性がある。

- (2) 収支見込みも、当初は数年間の赤字が必至とみられていたが、今月三十一日に予定している初年度決算で「経常利益は数億円単位の黒字になる」(山下社長)という。(BCCWJ, サンプルID: PN1d.00001)
- (3) トヨタ自動車が2割近い減少を見込む一方で、軽が好調な本田は12年の74万5000台から「できるだけ伸ばしたい」(岩村哲夫副社長)とした。(日経, サンプルID: 0201NKM0307)

このほか、記事内容との関係ではなく、表4において最も違いが現れるのは1.10(体-関係-事柄)と3.10(相-関係-真偽)であるが、これらは1.1010または3.1010(こそあど・他)の「此れ」「其れ」「此の」「其の」「こう」「そう」の差による。日経サンプルでは「それほど」「それだけ」「これまで」「これから」「このほど」などの限定的な表現で用いられる例(例(4))がほとんどであり、BCCWJサンプルに見られる指示詞として機能しているこそあど(例(5))の出現はまれであった。

- (4) これまでは世界販売を伸ばしても単独営業赤字が収益の足を引っ張ってきた。今後は円安を追い風に競争力を高めて韓国・現代自動車などライバル勢に対抗する。(日経, サンプルID: 0206NKM0075)
- (5) 経団連は「債権放棄は資本主義のルールをゆがめる。本来、採用すべきでない」との考えだ。これは、債権放棄を受けた企業に経営責任や株主責任を求め、再建計画に数値基準を設けた中間報告の中核的な考えに反映されている。(BCCWJ, サンプルID: PN1g.00002)

表4 語義情報の分布(類・中項目: 分類番号の小数点以下二桁まで)

	日経 (割合)	BCCWJ (割合)	日経 (割合)	BCCWJ (割合)
1.10 体-関係-事柄	261 0.783%	5325 1.534%	2.13 用-関係-様相	24 0.072% 234 0.067%
1.11 体-関係-類	398 1.194%	3303 0.952%	2.14 用-関係-力	0 0.000% 7 0.002%
1.12 体-関係-存在	134 0.402%	1160 0.334%	2.15 用-関係-作用	631 1.892% 8037 2.316%
1.13 体-関係-様相	244 0.732%	2071 0.597%	2.16 用-関係-時間	29 0.087% 225 0.065%
1.14 体-関係-力	87 0.261%	354 0.102%	2.17 用-関係-空間	14 0.042% 133 0.038%
1.15 体-関係-作用	751 2.252%	4193 1.208%	2.19 用-関係-量	9 0.027% 107 0.031%
1.16 体-関係-時間	779 2.336%	6460 1.861%	2.30 用-活動-心	332 0.996% 4438 1.279%
1.17 体-関係-空間	321 0.963%	4219 1.216%	2.31 用-活動-言語	136 0.408% 3172 0.914%
1.18 体-関係-形	35 0.105%	535 0.154%	2.32 用-活動-芸術	7 0.021% 112 0.032%
1.19 体-関係-量	3740 11.216%	23243 6.696%	2.33 用-活動-生活	51 0.153% 1203 0.347%
1.20 体-主体-人間	349 1.047%	6371 1.836%	2.34 用-活動-行為	806 2.417% 8105 2.335%
1.21 体-主体-家族	67 0.201%	1150 0.331%	2.35 用-活動-交わり	82 0.246% 592 0.171%
1.22 体-主体-仲間	62 0.186%	441 0.127%	2.36 用-活動-待遇	47 0.141% 522 0.150%
1.23 体-主体-人物	73 0.219%	2604 0.750%	2.37 用-活動-経済	74 0.222% 1334 0.384%
1.24 体-主体-成員	425 1.275%	2725 0.785%	2.38 用-活動-事業	30 0.090% 879 0.253%
1.25 体-主体-公私	902 2.705%	6056 1.745%	2.50 用-自然-自然	5 0.015% 100 0.029%
1.26 体-主体-社会	632 1.895%	2825 0.814%	2.51 用-自然-物質	5 0.015% 115 0.033%
1.27 体-主体-機関	454 1.361%	2089 0.602%	2.56 用-自然-身体	0 0.000% 9 0.003%
1.30 体-活動-心	888 2.663%	8411 2.423%	2.57 用-自然-生命	4 0.012% 360 0.104%
1.31 体-活動-言語	386 1.158%	4345 1.252%	3.10 相-関係-真偽	141 0.423% 3130 0.902%
1.32 体-活動-芸術	66 0.198%	994 0.286%	3.11 相-関係-類	164 0.492% 1861 0.536%
1.33 体-活動-生活	292 0.876%	2640 0.761%	3.12 相-関係-存在	65 0.195% 1335 0.385%
1.34 体-活動-行為	238 0.714%	1476 0.425%	3.13 相-関係-様相	178 0.534% 2510 0.723%
1.35 体-活動-交わり	357 1.071%	2413 0.695%	3.14 相-関係-力	31 0.093% 223 0.064%
1.36 体-活動-待遇	457 1.370%	2158 0.622%	3.15 相-関係-作用	28 0.084% 311 0.090%
1.37 体-活動-経済	996 2.987%	2839 0.818%	3.16 相-関係-時間	213 0.639% 2023 0.583%
1.38 体-活動-事業	440 1.319%	2922 0.842%	3.17 相-関係-空間	10 0.030% 21 0.006%
1.40 体-生産物-物品	67 0.201%	680 0.196%	3.18 相-関係-形	0 0.000% 55 0.016%
1.41 体-生産物-資材	71 0.213%	711 0.205%	3.19 相-関係-量	308 0.924% 4156 1.197%
1.42 体-生産物-衣料	17 0.051%	582 0.168%	3.20 相-主体-人間	1 0.003% 7 0.002%
1.43 体-生産物-食料	53 0.159%	1210 0.349%	3.25 相-主体-公私	1 0.003% 0 0.000%
1.44 体-生産物-住居	30 0.090%	1148 0.331%	3.30 相-活動-心	94 0.282% 1784 0.514%
1.45 体-生産物-道具	56 0.168%	1221 0.352%	3.31 相-活動-言語	2 0.006% 64 0.018%
1.46 体-生産物-機械	195 0.585%	1359 0.392%	3.33 相-活動-生活	5 0.015% 158 0.046%
1.47 体-生産物-土地利用	17 0.051%	507 0.146%	3.34 相-活動-行為	33 0.099% 353 0.102%
1.50 体-自然-自然	45 0.135%	1171 0.337%	3.35 相-活動-交わり	0 0.000% 16 0.005%
1.51 体-自然-物質	83 0.249%	791 0.228%	3.36 相-活動-待遇	24 0.072% 126 0.036%
1.52 体-自然-天地	37 0.111%	680 0.196%	3.37 相-活動-経済	16 0.048% 179 0.052%
1.53 体-自然-生物	24 0.072%	107 0.031%	3.38 相-活動-事業	0 0.000% 2 0.001%
1.54 体-自然-植物	13 0.039%	569 0.164%	3.50 相-自然-自然	9 0.027% 420 0.121%
1.55 体-自然-動物	10 0.030%	682 0.196%	3.51 相-自然-物質	0 0.000% 22 0.006%
1.56 体-自然-身体	51 0.153%	1914 0.551%	3.52 相-自然-天地	1 0.003% 3 0.001%
1.57 体-自然-生命	49 0.147%	843 0.243%	3.53 相-自然-生物	3 0.009% 9 0.003%
2.10 用-関係-事柄	0 0.000%	4 0.001%	3.56 相-自然-身体	0 0.000% 21 0.006%
2.11 用-関係-類	198 0.594%	1375 0.396%	3.57 相-自然-生命	1 0.003% 66 0.019%
2.12 用-関係-存在	447 1.340%	7437 2.143%	4. 他	81 0.243% 2497 0.719%
タグなし	14354 43.046%	169745 48.905%	総計	33346 100.000% 347094 100.000%

4 おわりに

『日本経済新聞記事オープンコーパス』に付与した新聞記事分類情報(メタデータ)と分類語彙表番号(語義情報)について報告した。

BCCWJの新聞記事サンプルにおける記事分布や語義分布は新聞一般の実態に近いものと考えられる。本研究の情報付与により、『日本経済新聞記事オープンコーパス』では、記事の種別において一般報道記事以外の割合がきわめて低く、「経済」記事

の割合が半数に及び、記事分布において「経済」分野の専門性の高いことが確認された。しかし、経済記事の内容の内訳(分類情報(小))では大差がなく、語義分布においても新聞一般との違いはほとんどなかった。『日本経済新聞記事オープンコーパス』は、記事内容の詳細な分布や語義の分布においては、概ねBCCWJの新聞記事サンプルと同様の傾向があるといえる。但し、語義分布に現れた一部の違いには、文体的な違いが見られ、読み手に専門性を強める印象が生じている可能性が考えられる。

謝辞

本研究は国立国語研究所共同研究プロジェクト「実証的な理論・対照言語学の推進：アノテーションデータをを用いた実証的計算心理言語学」、JSPS 科研費 JP18K00634, JP22K18483 の助成を受けた。

参考文献

- [1] 浅原正幸, 加藤祥, 高松純子. 『日本経済新聞記事オープンコーパス』に対する言語受容情報付与. 2023 年度日本認知科学会第 40 回大会, 2023.
- [2] 国立国語研究所 (編). 分類語彙表増補改訂版. 大日本図書, 2004.
- [3] Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. Balanced Corpus of Contemporary Written Japanese. **Language Resources and Evaluation**, Vol. 48, pp. 345–371, 2014.
- [4] 加藤祥, 浅原正幸. 『現代日本語書き言葉均衡コーパス』新聞サブコーパスの記事情報. 日本語の研究, Vol. 19, No. 2, pp. 206–214, 2023.
- [5] 加藤祥, 浅原正幸, 山崎誠. 分類語彙表番号を付与した『現代日本語書き言葉均衡コーパス』の書籍・新聞・雑誌データ. 日本語の研究, Vol. 15, No. 2, pp. 134–141, 2019.
- [6] 丸山岳彦, 柏野和佳子, 田中牧郎. 『現代日本語書き言葉均衡コーパス』利用の手引 第 1.1 版, 第 3 章 サンプルリング. 国立国語研究所コーパス開発センター, 2015. <https://doi.org/10.15084/00003228>.

A 付録

表5 新聞記事に付与した記事分類

分類(大)	分類(小)						
総類	00_総類	01_皇室	02_言論報道	03_新聞	04_放送	05_出版	06_新メディア
	07_大戦						
政治	10_政治	11_国会	12_選挙	13_政党	14_行政	15_地方行政	16_司法
	17_財政	18_外交					
経済	20_経済	21_金融	22_貿易	23_商業	24_鉱工業	25_エネルギー	26_農林漁業
	27_運輸交通	28_情報通信	29_土木建設				
労働	30_労働	31_雇用	32_労働条件	33_賃金	34_労働運動	35_労働組合	36_官公労組
	37_民間労組	38_職業					
文化	40_文化	41_文学	42_教育	43_宗教	44_美術	45_芸能	46_演劇
	47_映画	48_音楽					
科学	50_科学	51_地球	52_天文	53_宇宙	54_原子力	55_生物	56_医療
	57_生理疫病	58_保健衛生	59_公害				
社会	60_社会	61_福祉	62_世代	63_家庭	64_生活	65_趣味娯楽	66_旅行観光
	67_世相風俗	68_行事					
事件	70_事件	71_災害	72_火災	73_事故	74_交通事故	75_凶悪犯	76_人質犯
	77_経済犯	78_公安事件	79_その他犯罪				
スポーツ	80_スポーツ	81_総合競技	82_陸上	83_水上	84_冬季	85_球技	86_野球
	87_武道	88_諸競技					
国際	90_国際	91_国際政治	92_国際軍事	93_平和運動	94_被爆	95_国際経済	96_国連
見出し							

表6 記事種別1に基づく記事分布

	A_評論	B_投書	C_インタビュー	D_ハウツー	E_催し案内	F_人	G_色	H_人事	I_募集	J_賞	一般	総計
日経新聞	4	0	3	0	2	4	0	2	0	0	81	96
朝日新聞	57	3	20	3	22	19	46	2	9	4	335	520
読売新聞	48	8	11	13	28	27	17	9	0	0	362	523
毎日新聞	60	1	10	4	42	27	33	2	2	2	296	479
産経新聞	40	3	9	1	29	41	23	2	1	0	305	454
北海道新聞	25	3	20	2	25	12	23	1	1	3	249	364
中日新聞	43	9	12	0	32	20	30	1	8	1	270	426
西日本新聞	33	9	16	0	30	27	34	3	1	1	218	372
河北新報	18	0	8	1	22	17	11	5	3	4	275	364
新潟日報	26	0	7	0	45	15	16	0	0	0	207	316
京都新聞	28	0	15	2	49	15	24	2	3	3	221	362
神戸新聞	31	18	8	0	42	18	20	2	3	2	236	380
中国新聞	7	0	6	0	37	18	15	2	3	0	253	341
高知新聞	26	3	4	0	34	9	10	2	1	0	231	320
琉球新報	22	0	2	2	17	15	2	4	1	7	292	364

表7 記事種別2に基づく記事分布

	1_連載	2_解説	3_テキスト	5_略歴	6_日誌	7_用語	8_データ	9_死亡	その他	総計
日経新聞	0	6	87	0	0	0	3	-	0	96
朝日新聞	53	92	354	0	5	0	6	3	7	520
読売新聞	36	97	378	0	5	0	3	2	2	523
毎日新聞	24	88	347	0	8	1	6	1	4	479
産経新聞	55	81	308	0	2	0	0	7	1	454
北海道新聞	25	56	269	0	5	2	1	1	5	364
中日新聞	32	59	328	1	5	0	1	0	0	426
西日本新聞	48	71	240	2	4	3	0	2	2	372
河北新報	10	54	293	0	3	0	2	0	2	364
新潟日報	17	55	237	0	3	0	3	0	1	316
京都新聞	38	66	247	0	3	0	1	4	3	362
神戸新聞	25	61	286	0	5	1	2	0	0	380
中国新聞	21	44	268	0	3	0	0	3	2	341
高知新聞	27	32	254	0	1	0	3	3		320
琉球新報	18	24	318	0	3	0	0	0	1	364