

助詞の使用実態 —BCCWJ・CSJ にみる分布—

丸山 直子 (東京女子大学現代教養学部) †

Usage of Postpositional Particles in BCCWJ and CSJ

Naoko Maruyama (Tokyo Woman's Christian University)

要旨

現代日本語の助詞について、現代日本語書き言葉均衡コーパス (BCCWJ) 及び話し言葉コーパス (CSJ) における用いられ方を観察し、書き言葉と話し言葉の違い、及びそれぞれのサブコーパス (レジスター) ごとの違いを明らかにした。BCCWJ はコアのみ (新聞、雑誌、書籍、白書、知恵袋、ブログ) を調査対象とし、CSJ は、同一話者による独話 (学会講演) と対話 (自由会話) 4 件ずつを対象として調査を行った。コレスポネンシ分析も行った。

BCCWJ も CSJ も、全語数の約 30% が助詞であり、助詞の中では格助詞が最も多い。BCCWJ においては、白書と知恵袋・ブログは、助詞の使用法に関して、様々な点で対極にある。白書はかなり特殊で、格助詞相当の複合辞が多く、短単位と長単位で大きく分布が異なる。新聞は多少白書に似た性質を持つ。知恵袋とブログは、終助詞が多い等の話し言葉的な性質を帯びているが、相互に異なる性質も持つ。CSJ は、講演の方が格助詞が多く、対話には副助詞・終助詞が多い。融合・縮約の多さも話し言葉特有の現象として指摘できる。

1. はじめに

現代日本語の助詞について、現代日本語書き言葉均衡コーパス (BCCWJ) 及び話し言葉コーパス (CSJ) における用いられ方を観察することで、書き言葉と話し言葉の違い、及びそれぞれのサブコーパス (レジスター) ごとの違いを明らかにする。BCCWJ はコアのみ (新聞、雑誌、書籍、白書、知恵袋、ブログ) を調査対象とし、CSJ は、同一話者による独話 (学会講演) と対話 (自由会話) 4 件ずつを対象とする。

2. 調査対象

BCCWJ、CSJ の、調査対象としたものを表 1、表 2 に記す。BCCWJ は、コアすべてで、短単位で約 100 万語、長単位で 80 万語である。CSJ は、4 名の学会講演・自由会話 1 件ずつで、計 8 件である。こちらは短単位で計 3 万語という小さなサンプルである。

表 1 BCCWJ の調査対象

	短単位総数	長単位総数
出版・新聞コア	308,504	224,140
出版・雑誌コア	202,268	159,883
出版・書籍コア	204,050	169,730
特定目的・白書コア	197,011	129,646
特定目的・知恵袋コア	93,932	78,770
特定目的・ブログコア	92,746	75,242
計	1,098,511	837,411

† maruyama@lab.twcu.ac.jp

表2 CSJの調査対象

講演者 ID	性別	生年代	基にした学 会講演 ID	短単位 数	長単位 数	自由会話 ID	短単位 数	長単位 数
1185	女	70to74	A11F0703	5,634	4,697	D03F0034	3,021	2,699
19	女	65to69	A05F0043	3,512	2,655	D03F0058	2,330	2,039
471	男	75to79	A11M0369	3,119	2,246	D03M0004	2,491	2,178
373	男	45to49	A11M0469	6,763	5,379	D03M0038	3,638	3,278
計				19,028	14,977		11,480	10,194

3. 助詞の分類

本稿では、BCCWJ は中納言オンライン版の短単位・長単位分割及び品詞分類に基づき、CSJ は、DVD に収められている、短単位・長単位データに基づき集計した。BCCWJ・CSJ とも、格助詞・副助詞・係助詞・接続助詞・終助詞・準体助詞の六分類である。

4. 調査で得られた助詞

以下に、それぞれのコーパスに含まれていた助詞の一覧を表にして示す。

表3 コーパス中の助詞一覧

	BCCWJ	CSJ
格助詞(短単位)	ガ、ヲ、ニ、ト、デ、ヘ、ヨリ、カラ、ノ、トテ、ニテ、サ	ガ、ヲ、ニ、ト、デ、ヘ、ヨリ、カラ、ノ、デハ(じゃ)
格助詞(長単位)	ヲ通ジテ、ヲハジメ、ヲメグル、ヲモツテ、ニアタツテ、ニアタリ、ニイタルマデ、ニオイテ、ニオケル、ニ関シテ、ニ関スル、ニ際シ、ニ際シテ、ニシテ、ニ対シ、ニ対シテ、ニ対スル、ニツイテ、ニツキ、ニトツテ、ニヨツテ、ニヨリ、ニヨル、ニヨルト、ニヨレバ、ニワタツテ、ニワタリ、ニワタル、際ニ、トイウ、トイッタ、トシテ、カラシテ、カラスルト、カラスレバ、タメノ	ヲモトニシタ、ヲモトニシテ、ニオイテ、ニオケル、ニ関シテ、ニ関シマシテ、ニ関スル、ニ比ベテ、ニ従ツテ、ニ対シテ、ニ対シマシテ、ニ対スル、ニツイテ、ニツキマシテ、ニトツテ、ニ伴ウ、ニ基ヅイタ、ニ基ヅイテ、ニ基ヅク、ニヨツテ、ニヨル、ニヨリマス、ニヨリマス、トイウ、トイッタ、トシテ、トイタシマシテ
副助詞(短単位)	ダケ、ノミ、バカリ、キリ、マデ、クライ、ナド、ナンカ、ナンテ、カ、ヤ、ヤラ、ホド、シカ、サエ、スラ、ツテ、タリ、シ、カシラ、ガニ、シモ、ズツ、ゾ、ダニ、タラ、ツ、デン、ドコロ、ナリ、ナンゾ、ナント	ダケ、ノミ、マデ、クライ、ナド、ナンカ、カ、ヤ、ホド、シカ、スラ、ツテ、タリ、シモ、ズツ、タツテ、モ、コソ
副助詞(長単位)	ダケデナク、ノミナラズ、ツウ、ニ限ラズ	トカ
係助詞(短単位)	ハ、モ、コソ、ゾ、バ、ヤ	ハ
係助詞(長単位)	トイエドモ、トイッテモ、トキタラ、ニイタツテハ	なし
接続助詞(短単位)	シ、テ、ト、バ、カラ、ガ、ケレド、トモ、ニ、タツテ、ツツ、ナガラ、ケン、サカイ、ド、トテ、ナリ	シ、テ、ト、バ、カラ、ガ、ケレド、ツツ、ナガラ、テハ(ちゃ)
接続助詞(長単位)	カラトイッテ、カラニハ、ウエデ、ウエニ、カト思ウト、タトコロ、タトコロデ、タメニ、トシタラ、トシテモ、トスレバ、トテ、ト同時ニ、トトモニ、トハイエ、ニ関ワラズ、ニシタガイ、ニシタガツテ、ニシテハ、ニシテモ、ニシロ、ニセヨ、ニツレ、ニツレテ、ニモカカワラズ、モノ、ヤイナヤ、ワリニ	テハ、テモ、ノデ、ノニ

終助詞(短単位)	カ、サ、ナ、ネ、ヨ、ゼ、ゾ、ワ、ノ、イ、カシラ、ヤ、ケ、モノ、ジャン、エ、カナ、クサ、チョ、デ、テン、ド、ネン、ノウ、バイ、ベイ、モガ	カ、ナ、ネ、ヨ、ゾ、ワ、カシラ、ヤ、ケ、モノ
終助詞(長単位)	なし	なし
準体助詞(短単位)	ノ	ノ
準体助詞(長単位)	なし	なし

長単位の欄は、短単位にない形のもの載せている。それぞれ、出現形が異なるものも含んでいる。特に話し言葉には、縮約・融合の形が多く含まれる。

5. BCCWJにおける助詞

5.1 全語数における助詞の割合と助詞内における各助詞の割合

助詞の数を以下に示す。

表4 BCCWJ 全語数における助詞の割合 (短単位)

	全語数	格助詞	副助詞	係助詞	接続助詞	終助詞	準体助詞	助詞総数 個数	%
新聞コア	308504	57052	4186	11838	8876	554	1207	83713	27.14
雑誌コア	202268	35722	3015	9108	8491	1483	2047	59866	29.60
書籍コア	204050	38083	3303	10426	10744	1369	2681	66606	32.64
白書コア	197011	36619	1985	4290	6084	59	120	49157	24.95
知恵袋コア	93932	14274	2086	4215	5457	2384	1976	30392	32.36
ブログコア	92746	14104	1533	4137	4405	1607	1309	27095	29.21
計	1098511	195854	16108	44014	44057	7456	9340	316829	28.84

この調査から、以下のことがわかる。

- 1) 全語数の約30%が助詞である。
- 2) 助詞の中では格助詞が最も多い。助詞のうち47%~74%が格助詞。
- 3) 知恵袋・ブログは、他に比べて、格助詞が少なく、終助詞が多い。
- 4) 白書と、知恵袋・ブログは、対極にある。白書はかなり特殊である。新聞は多少白書に似た性質を持つ。

全体： 格助詞 > 接続助詞・係助詞 > 副助詞 > 準体助詞 > 終助詞

新聞： 格助詞 > 係助詞 > 接続助詞 > 副助詞 > 準体助詞 > 終助詞

雑誌： 格助詞 > 係助詞 > 接続助詞 > 副助詞 > 準体助詞 > 終助詞

書籍： 格助詞 > 接続助詞 > 係助詞 > 副助詞 > 準体助詞 > 終助詞

白書： 格助詞 > 接続助詞 > 係助詞 > 副助詞 > 準体助詞 > 終助詞

知恵袋： 格助詞 > 接続助詞 > 係助詞 > 終助詞 > 副助詞 > 準体助詞

ブログ： 格助詞 > 接続助詞 > 係助詞 > 終助詞 > 副助詞 > 準体助詞

(上記二重下線は、他のレジスターに比べて相対的に多いもの、一重下線は少ないもの。以下同様。)

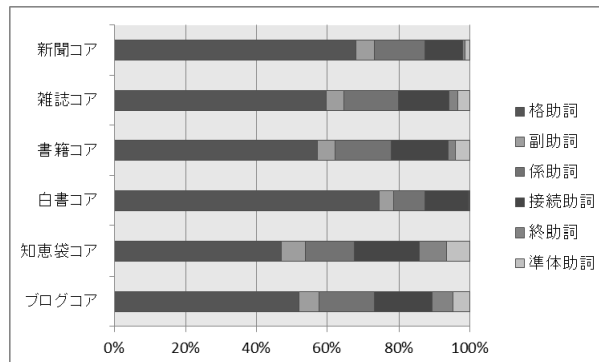


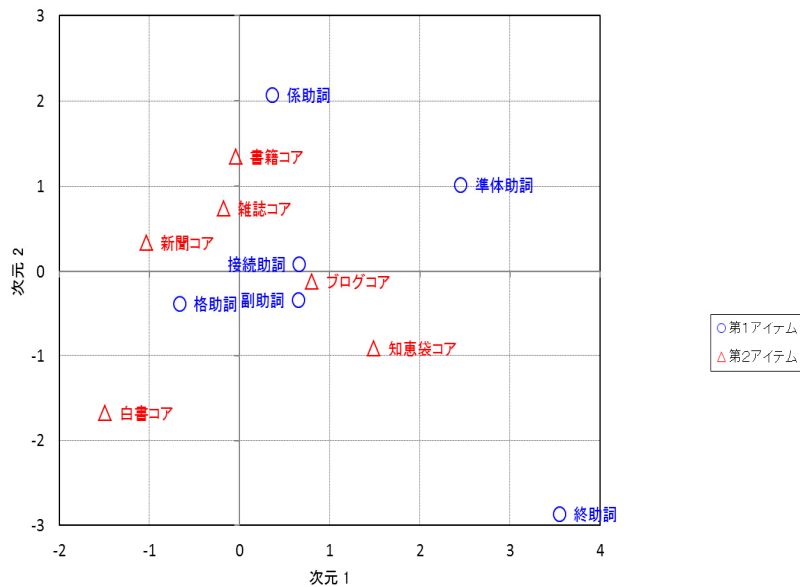
図1 BCCWJ レジスターごとの助詞の割合 (短単位)

長単位でも、全体の傾向は変わらない。格助詞が最も多い。
 短単位の場合の1万語当たりの数は、表5の通りである。

表5 BCCWJ 1万語当たりの助詞の数(短単位)

	格助詞	副助詞	係助詞	接続助詞	終助詞	準体助詞	助詞全体
新聞コア	1849	136	384	288	18	39	2714
雑誌コア	1766	149	450	420	73	101	2960
書籍コア	1866	162	511	527	67	131	3264
白書コア	1859	101	218	309	3	6	2495
知恵袋コア	1520	222	449	581	254	210	3236
ブログコア	1521	165	446	475	173	141	2921
計	1783	147	401	401	68	85	2884

このクロス表の内容をもとに、ジャンルと助詞の関係をより詳細に把握するため、助詞タイプを第1アイテム、コーパス種別を第2アイテムとしてコレスポネンス分析を行った。その結果、下記の散布図を得た(図2)。なお、第1次元の寄与率は90.64%、第2次元の寄与率は6.95%、2つの次元による累計寄与率は97.59%であるため、2つの次元に基づく解釈に一定の妥当性があると判断した。第1次元の寄与率が圧倒的である。軸解釈を行うと、第1次元はブログや知恵袋などのくだけた話し言葉的ジャンル(+)と、新聞・白書のようなかたい書き言葉的ジャンル(-)を区分している軸と考えられる。また、第2次元は書籍・雑誌のような一般的内容を扱ったジャンル(+)と白書のような特定内容を扱ったジャンル(-)を区分する軸と考えられる。このことから考えると、第1象限、つまり、くだけた言語と一般的内容を特徴とする領域には係助詞、準体助詞が多く、第2象限、つまり、くだけた言語と特定内容のジャンルには終助詞が多い。第3象限、つまり、かたくて一般的なジャンルに特徴的な助詞は存在せず、第4象限、つまり、かたくて特定内容のものには格助詞が多い。



第1、第2アイテムによるスコア散布図

図2 BCCWJ コレスポネンス分析の散布図(短単位)

5. 2 それぞれの助詞における語の割合

5. 2. 1 格助詞

BCCWJ コアにおける格助詞の内訳は以下の通りである。

表 6 BCCWJ 格助詞の数 (短単位)

	ガ	ヲ	ニ	ト	デ	ヘ	ヨリ	カラ	ノ	その他
新聞コア	7235	10294	9826	6165	5102	522	141	1362	16390	15
雑誌コア	5000	6192	6645	4168	2728	228	137	900	9714	10
書籍コア	5372	6576	7361	5079	2530	277	140	861	9880	7
白書コア	3545	6265	7948	3615	1456	341	129	675	12636	9
知恵袋コア	2375	2029	2641	1931	1552	57	97	326	3261	5
ブログコア	2018	1965	2643	1710	1371	118	82	410	3773	14
計	25545	33321	37064	22668	14739	1543	726	4534	55654	60

新聞： ノ>ヲ>ニ>ガ>ト>デ>カラ>ヘ>ヨリ
 雑誌： ノ>ニ>ヲ>ガ>ト>デ>カラ>ヘ>ヨリ
 書籍： ノ>ニ>ヲ>ガ>ト>デ>カラ>ヘ>ヨリ
 白書： ノ>ニ>ヲ>ト>ガ>デ>カラ>ヘ>ヨリ
 知恵袋： ノ>ニ>ガ>ヲ>ト>デ>カラ>ヨリ>ヘ
 ブログ： ノ>ニ>ガ>ヲ>ト>デ>カラ>ヘ>ヨリ

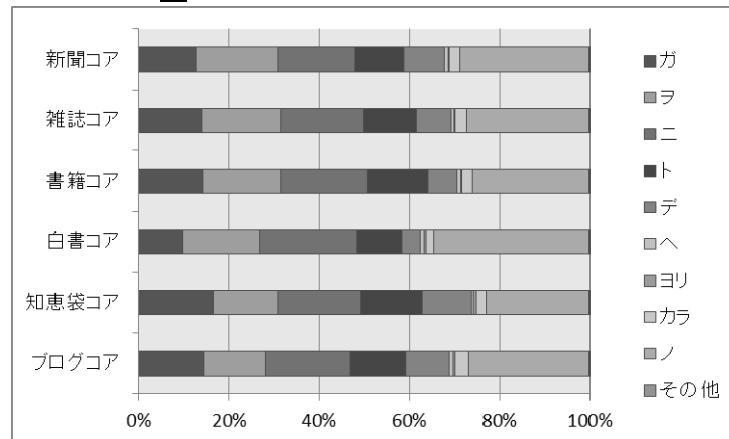


図 3 BCCWJ レジスターごとの格助詞の割合 (短単位)

長単位で調査すると、だいぶ値が異なる。ニを伴う複合辞、トを伴う複合辞の数が多いことがわかる。特に白書には、「により」「における」等、ニを伴う複合辞が多い。

表 8 BCCWJ 格助詞の数 (長単位)

	ガ	ヲ	ニ	ニを伴う複合辞	ト	トを伴う複合辞	デ	ヘ	ヨリ	ノ
新聞コア	7168	10177	8402	1180	5160	910	4998	522	133	16290
雑誌コア	4895	6156	5860	479	3126	882	2579	228	128	9633
書籍コア	5155	6534	6297	658	3651	1187	2375	277	126	9826
白書コア	3473	6148	3843	3676	2688	636	1417	341	128	12450
知恵袋コア	2310	2027	2390	130	1586	261	1450	57	93	3247
ブログコア	1967	1957	2386	136	1375	257	1197	118	77	3756
計	24968	32999	29178	6259	17586	4133	14016	1543	685	55202

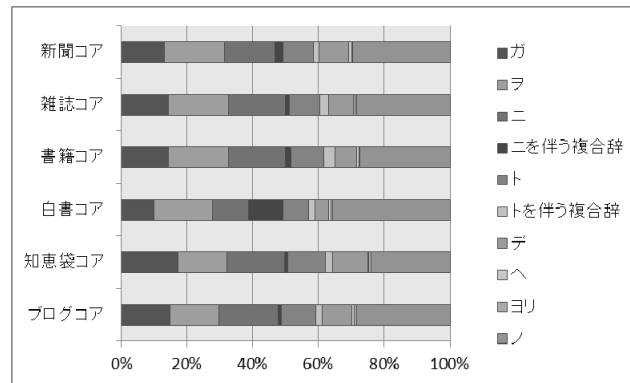


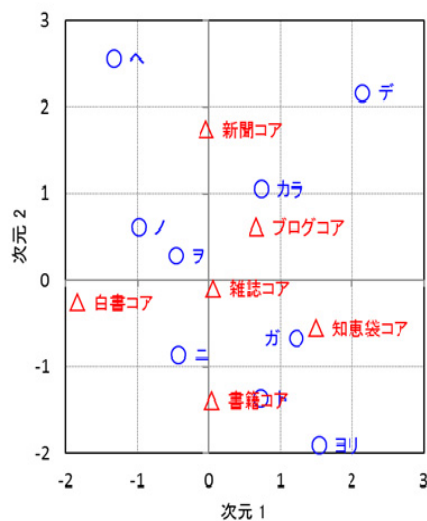
図4 BCCWJ レジスターごとの格助詞の割合 (長単位)

短単位の1万語当たりの数は、表7の通り。

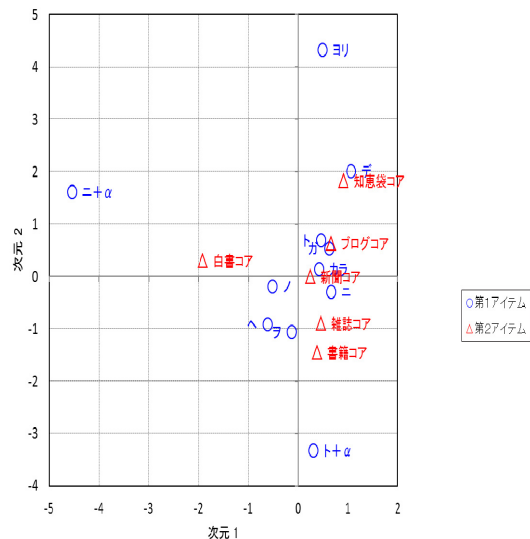
表7 BCCWJ 1万語当たりの格助詞の数 (短単位)

	ガ	ヲ	ニ	ト	デ	ヘ	ヨリ	カラ	ノ
新聞コア	235	334	319	200	165	17	5	44	531
雑誌コア	247	306	329	206	135	11	7	44	480
書籍コア	263	322	361	249	124	14	7	42	484
白書コア	180	318	403	183	74	17	7	34	641
知恵袋コア	253	216	281	206	165	6	10	35	347
ブログコア	218	212	285	184	148	13	9	44	407

このクロス表の内容をもとに、ジャンルと助詞の関係をより詳細に把握するため、助詞タイプを第1アイテム、コーパス種別を第2アイテムとしてコレスポネンス分析を行った。その結果、下記の散布図を得た(図5)。なお、第1次元の寄与率は80.21%、第2次元の寄与率は10.03%、2つの次元による累計寄与率は90.23%であるため、2つの次元に基づく解釈に一定の妥当性があると判断した。長単位についても同様の分析を行い、散布図を得た(図6)。軸解釈は、図5,6とも図2と同様でよいと思われるが、長単位の方が、より、白書及び $\text{ニ}+\alpha$ (ニを伴う複合辞) の位置が特徴的となっている。



第1、第2アイテムによるスコア散布図



第1、第2アイテムによるスコア散布図

図5 コレスポネンス分析の散布図(短単位) 図6 コレスポネンス分析の散布図(長単位)

格助詞に関しては、以下のことがわかった。

- ・格助詞の分布は、短単位と長単位でだいぶ異なる。特に、白書において違いが顕著である。
- ・短単位で白書にニが多い理由はニを伴う複合辞が多いからである。このことは、長単位の調査を行うとわかる。白書は、格助詞ニの46.25%が複合辞である。新聞が12%、あとのレジスターは一ケタである。「により」「における」「において」が多い。「により」「によって」は、白書以外は「により」より「によって」が多い。「に対し」「に対して」は、新聞のみ「に対し」が多い。複合辞に関わる格助詞はニとトが主である。接続助詞テが格助詞相当の複合辞を作ることが多いため、白書は短単位で調べると接続助詞のテが多い。
- ・デは白書には少ない。デは話し言葉的であり、デの代わりに複合辞を用いるためであると思われる。
- ・知恵袋にはガ・デが多い。

5. 2. 2 副助詞

BCCWJ コアにおける副助詞の数は以下の通りである。合計数が多い順に並べた。副助詞以降は、紙幅の関係で図を省略する。

表9 BCCWJ 副助詞の数 (短単位)

	ヤ	ナド	カ	マデ	ダケ	ツテ	タリ	ホド	クライ	シカ	バカリ	ナンテ	ノミ	サエ	その他	計
新聞コア	1106	1358	400	552	260	26	142	73	34	66	51	17	19	20	62	4186
雑誌コア	574	420	574	338	232	156	136	134	89	67	62	48	43	36	106	3015
書籍コア	506	387	830	354	313	132	143	144	101	81	59	37	24	48	144	3303
白書コア	881	613	70	224	42	0	53	26	3	3	3	0	45	2	20	1985
知恵袋コア	169	170	613	148	139	318	121	55	160	51	22	44	12	7	57	2086
ブログコア	129	82	474	172	127	169	74	45	88	49	23	32	17	6	46	1533
計	3365	3030	2961	1788	1113	801	669	477	475	317	220	178	160	119	435	16108

- ・新聞と白書にはヤ・ナドが多い。
- ・知恵袋・ブログにはカ・ツテが多い。

5. 2. 3 係助詞

BCCWJ コアにおける係助詞の数は以下の通りである。

表10 BCCWJ 係助詞の数 (短単位)

	ハ	モ	コソ	その他	計
新聞コア	8956	2840	41	1	11838
雑誌コア	6350	2706	47	5	9108
書籍コア	7316	3069	38	3	10426
白書コア	3577	711	2	0	4290
知恵袋コア	2780	1425	9	1	4215
ブログコア	2655	1466	13	3	4137
計	31634	12217	150	13	44014

- ・どのレジスターも、ハ>モ>コソの順である。
- ・白書は係助詞が全体的に少ないが、特にモが少ない。

5. 2. 4 接続助詞

BCCWJ コアにおける接続助詞の数は以下の通りである。合計数が多い順に並べた。

表 11 BCCWJ 接続助詞の数 (短単位)

	テ	ガ	ト	バ	カラ	ケレド	ナガラ	シ	ツツ	その他	計
新聞コア	6652	922	473	380	172	36	145	43	35	18	8876
雑誌コア	6104	692	483	398	343	150	148	114	28	31	8491
書籍コア	7803	811	540	607	514	123	159	126	23	38	10744
白書コア	5416	207	193	87	37	1	62	1	74	6	6084
知恵袋コア	3304	890	361	307	265	150	35	132	6	7	5457
ブログコア	2893	437	237	184	219	217	57	127	19	15	4405
計	32172	3959	2287	1963	1550	677	606	543	185	115	44057

- ・白書は、テを用いた複合辞が多いので、テが多い。
- ・白書は、テ以外の接続助詞は少ない。

5. 2. 5 終助詞

BCCWJ コアにおける終助詞の数は以下の通りである。合計数が多い順に並べた。

表 12 BCCWJ 終助詞の数 (短単位)

	カ	ネ	ヨ	ナ	ノ	ワ	サ	ゾ	その他	計
新聞コア	323	77	71	42	8	4	1	10	18	554
雑誌コア	508	237	330	165	75	47	53	26	42	1483
書籍コア	626	183	206	125	67	61	25	20	56	1369
白書コア	58	0	0	0	1	0	0	0	0	59
知恵袋コア	1286	435	449	136	36	7	3	5	27	2384
ブログコア	417	468	288	265	40	22	30	13	64	1607
計	3218	1400	1344	733	227	141	112	74	207	7456

- ・白書は、カ以外はノが1件あったのみ。
- ・知恵袋は、カが多い。
- ・ブログは、出現形の種類が多い。

例) ナ: 265 例中「なあ」67 例、「な～」20 例、「なー」5 例、「なァ」2 例、「ナー」1 例
 ヨ: 288 例中「よ～」14 例、「よお」10 例、「よー」8 例、「ヨ」5 例、「よん」2 例、
 「よう」1 例、「よ～ん」1 例

6. CSJにおける助詞

6.1 全語数における助詞の割合と助詞内における各助詞の割合

表 13 CSJ 全語数における助詞の割合 (短単位)

		全語数	格助詞	副助詞	係助詞	接続助詞	終助詞	準体助詞	助詞総数	個数 %
学会講演	A11F0703	5634	924	199	126	299	57	95	1700	30.17
	A05F0043	3512	567	39	76	141	9	34	866	24.66
	A11M0369	3119	556	28	89	97	0	7	777	24.91
	A11M0469	6763	1111	123	138	411	43	62	1888	27.92
	Aグループ計	19028	3158	389	429	948	109	198	5231	27.49
自由会話	D03F0034	3021	266	201	56	129	206	65	923	30.55
	D03F0058	2330	217	236	37	99	138	68	795	34.12
	D03M0004	2491	266	106	44	107	96	90	709	28.46
	D03M0038	3638	422	175	71	178	177	88	1111	30.54
	Dグループ計	11480	1171	718	208	513	617	311	3538	30.82
総計	30508	4329	1107	637	1461	726	509	8769	28.74	

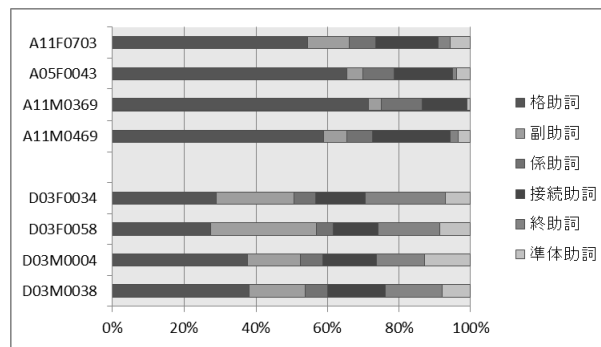
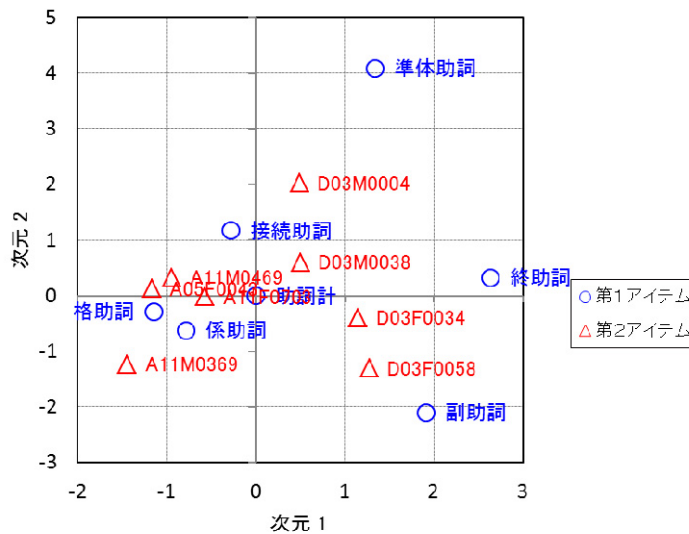


図 7 CSJ 独話（学会講演）・対話（自由会話）の助詞の割合（短単位）



第1、第2アイテムによるスコア散布図

図 8 CSJ コレスポネンス分析の散布図（短単位）

Aグループ（独話）とDグループ（対話）は、かなり異なる。図 8 を見ると、横軸の＋できれいに分かれているのがわかる。独話（学会講演）には格助詞が多く、対話（自由会話）には終助詞・副助詞が多い。

6. 2 それぞれの助詞における語の割合

6. 2. 1 格助詞

表 14 CSJ 格助詞の数（短単位）

		ガ	ヲ	ニ	ト	デ	ヘ	ヨリ	カラ	ノ	その他	計
学会講演	A11F0703	114	135	157	196	75	0	0	42	205	0	924
	A05F0043	110	60	94	92	20	0	3	7	180	1	567
	A11M0369	88	82	59	93	31	5	0	20	176	2	556
	A11M0469	144	191	154	210	60	2	0	79	271	0	1111
	Aグループ計	456	468	464	591	186	7	3	148	832	3	3158
自由会話	D03F0034	47	4	34	71	50	0	1	6	53	0	266
	D03F0058	36	19	28	60	38	0	5	1	30	0	217
	D03M0004	51	20	51	41	35	0	0	7	61	0	266
	D03M0038	100	60	63	87	50	3	0	15	43	1	422
	Dグループ計	234	103	176	259	173	3	6	29	187	1	1171
総計		690	571	640	850	359	10	9	177	1019	4	4329

6. 2. 2 副助詞

表 15 CSJ 副助詞の数 (短単位)

		カ	モ	ッテ	ナド	ダケ	クライ	マデ	タリ	ヤ	ナンカ	その他	計
学会講演	A11F0703	43	72	48	3	5	10	1	4	0	8	5	199
	A05F0043	4	20	1	1	2	0	3	0	1	0	7	39
	A11M0369	1	6	0	10	0	0	0	0	10	0	1	28
	A11M0469	25	52	1	22	4	4	8	0	2	0	5	123
	Aグループ計	73	150	50	36	11	14	12	4	13	8	18	389
自由会話	D03F0034	76	54	42	0	12	7	3	4		1	2	201
	D03F0058	127	44	49	0	3	4	4	4	0	0	1	236
	D03M0004	40	23	32	0	2	3	2	3	0	1	0	106
	D03M0038	59	42	63	0	2	2	4	2	0	0	1	175
	Dグループ計	302	163	186	0	19	16	13	13	0	2	4	718
総計		375	313	236	36	30	30	25	17	13	10	22	1107

- ・独話（学会講演）にナドを多く使う人がある。個人差がある。
- ・対話（自由会話）にはカ・ッテが多い。

6. 2. 3 係助詞

- ・CSJ ではハのみを係助詞としている。会話より講演の方が使用している。

6. 2. 4 接続助詞

- ・講演にテが多い。会話は縮約形「てる」が助動詞とされていることも影響していると思われる。
- ・講演にガが多く、会話にケレドが多い。（講演にケレドを用いている人も一人あり。）

6. 2. 5 終助詞

- ・会話の方が終助詞が多い。

7. 複合辞について

長単位として扱う場合と、長単位にしない場合がある。例えば、「ので」「のに」は、BCCWJ では、短単位分割でも長単位分割でも準体助詞「の」＋助動詞「だ」、準体助詞「の」＋格助詞「に」として扱うが、CSJ では、長単位分割では接続助詞として扱っている。複合辞の扱いは今後の課題である。

8. 複数の分類にまたがるものの扱いについて

例えば「って」は、少なくとも三種に分けられる。「手術って聞いてびっくりした」は格助詞、「～なんですって。」は終助詞、「人生って楽しいことばかりじゃないよ」は係助詞。しかし、BCCWJ 及び CSJ においては、すべて副助詞として扱っている。形態素解析としては副助詞として扱うとしても、その働きの違いを明らかにする必要がある。

9. まとめと今後の課題

- ・全語数の約 30%が助詞である。助詞の中では格助詞が最も多い。
- ・BCCWJ において、知恵袋・ブログは、他に比べて、格助詞が少なく、終助詞が多い。
- ・白書と、知恵袋・ブログは、対極にある。白書はかなり特殊である。新聞は多少白書に似た性質を持つ。知恵袋とブログは、ともに話し言葉的な性質を帯びているが、両者間で異なる性質も持つ。知恵袋の方がより独特である。
- ・CSJ においては、同一話者においても、学会講演と自由会話における違いが見られた。
- ・助詞の分類の仕方、認定の仕方には課題も残る。