

# 国立国語研究所学術情報リポジトリ

## Addressive Expressions in Written Japanese

メタデータ	言語: jpn 出版者: 公開日: 2015-10-30 キーワード (Ja): キーワード (En): 作成者: 加藤, 祥, 柏野, 和佳子, 立花, 幸子, 丸山, 岳彦, KATO, Sachi, KASHINO, Wakako, TACHIBANA, Sachiko, MARUYAMA, Takehiko メールアドレス: 所属:
URL	<a href="https://doi.org/10.15084/00000544">https://doi.org/10.15084/00000544</a>

## 語りかける書きことばの表現

加藤 祥<sup>a</sup> 柏野和佳子<sup>b</sup> 立花幸子<sup>c</sup> 丸山岳彦<sup>b</sup>

<sup>a</sup>国立国語研究所 コーパス開発センター プロジェクト研究員

<sup>b</sup>国立国語研究所 言語資源研究系/コーパス開発センター

<sup>c</sup>国立国語研究所 コーパス開発センター 技術補佐員

### 要旨

書籍テキストに見られる「語りかける」という文体の特徴を報告する。調査対象には、『現代日本語書き言葉均衡コーパス』(BCCWJ)に収録されている図書館サブコーパスを使用した。コーパスを用いた文体分析を行うにあたっては、語や文脈的な語の結びつきなどの頻度情報のほか、コーパスに付与された書誌情報やアノテーターによる作業コメントなどを用いた。「語りかける」という文体は、エッセイやブログなどのくだけたテキストにのみ出現しやすく、直接的に読み手へ呼びかけや問いかけを行うなどの表現を有すると考えられてきた。しかし、書籍においては、いわゆるハウツー本をはじめとするような教示的な態度を示すテキストに出現しやすい傾向があり、必ずしも直感的に「語りかける」ととらえられる表現が多く含まれるばかりではないことがわかった。本稿は、テキストが「語りかける」と読み手が判断した際に、文脈に依存した表現や、テキストに向かう読み手の前提的態度などが影響していたことを示す\*。

キーワード：BCCWJ, 文体, 語りかける, アノテーション, 文書分類

### 1. 「語りかける」文体とは何か

まず、以下に『現代日本語書き言葉均衡コーパス』(BCCWJ)に収録された書籍テキストの例を示す。

- (1) お年寄りや赤ちゃんを連れて人、重い荷物を持った人には、席をゆずりましょう。席をゆずるのは、あなたが立っていても平気なときだけよ。病気やケガをしているときは無理をしなくてもいいの。また、ゆずられた方は、「ありがとう」とお礼を言い、会釈をします。また自分が降りるとき、あるいは替わってくれた人が降りるときも、もう一度軽く「ありがとう」と言うのをおすすねなくね。

(LBb5\_00010：バーバラ寺岡「魅女ってみませんか」)<sup>1</sup>

- (1) は、書き手が読み手に対して直接語りかけているような印象を受けるテキストであろう<sup>2</sup>。

\* 本稿は国立国語研究所萌芽・発掘型共同研究プロジェクト「テキストの多様性を捉える分類指標の策定」(プロジェクトリーダー：柏野和佳子)の研究成果である。また、2012年9月～2014年3月に国立国語研究所で開催された第1回～第4回「コーパス日本語学ワークショップ」における発表内容(保田ほか2012a, 2012b, 2013a, 2013b)をまとめ、大幅に加筆修正したものである。とくに本稿では、図書館サブコーパスの全書籍サンプル(10,551サンプル)を対象とした調査をするとともに、再調査結果に基づき新たな分析と考察を行っている。

<sup>1</sup> 用例の出典は、(サンプルID：執筆者/著者「出典」)として示す。以下同様。

<sup>2</sup> テキストを読んだ複数(3名)のアノテーターが、「語りかける」と判断した典型例である。

このような印象は、読み手に対する人称（「あなた」）や丁寧さを示す接頭辞（「お」）、文末の終助詞（「ね」「よ」）や勧誘（意志推量形「ましょう」）のような特徴的表現によって形成されると考えられる。本稿は、このように書き手が読み手に対して直接語りかけているような印象を与えるテキストを「語りかける」テキスト<sup>3</sup>と呼び、「語りかける」文体を有するテキストの言語的特徴について分析を行う。

BCCWJ に収録されている図書館サブコーパスの書籍サンプル（全 10,551 サンプル）に対し、柏野（2013）は文書分類の観点から人手で情報を付与するアノテーション作業を行った。ここで付与された指標の一つに、「語りかけ性（とてもある・どちらかといえばある・とくにない:3 段階）」がある。この作業分類結果を用い、アノテーターがどのような根拠によってテキストを「語りかける」と判断したのかの分析し、「語りかける」という文体がどのような特徴を有しているか考察する。

## 2. 関連研究と本研究の目的

小磯ほか（2011）は、調査者から得た評定語を指標としてテキスト分析を行う際、「書きことば的一話しことば的」という尺度に、「読み手に語りかける一語りかけの少ない」という尺度を含む複数の観点に関与する可能性があると述べている。すなわち、「話しことば」的であることと「語りかける」文体の間には一定の関係性の存在が考えられる。一方、近代小説における言語表現を分析した安藤（2012）は、小説における再現的提示の手法とは、二人称的世界が顕在しないことであると、し、「読み手に語りかける言文一致の形がありえたならば、「言」に近い文体が創出されたかもしれない」と述べている。読み手に語りかける印象を与える文体とは、既存の「言文一致」の範疇にない表現を伴うということになるだろう。また、三宅（2005）は、「話しことば」の典型としてイメージされるものが「おしゃべり」であるとする。保田ほか（2012a）の調査では、「話しことば」的なテキストと「書きことば」的なテキストという観点で文書分類を実施した結果を用い、「語りかけ性」との関連について分析を行っている。この際、話しことば的であると判断されるテキスト<sup>4</sup>には、「おしゃべり」に頻出するフィルターや言いよどみ、音声的变化に関わる融合などが現れたのに対し、「語りかけ性」が「ある」とされるテキストにはその種の特徴が現れにくいという差異が見られた。「話しことば的」と判断されるテキストは、戯曲調で地の文がト書きの場合や講演の書き起こし、一人称小説などに限定され、書籍サンプルの 0.6%<sup>5</sup>と僅少

<sup>3</sup> 読まれることを前提として書かれるものである以上、どんな書きことばであっても読み手に語りかけるともいえる。しかし、ここでは、読み手（テキストを読んだアノテーター）が、「語りかける」と判断した結果を用い、「語りかける」文体の特徴分析を行う。

<sup>4</sup> 「話しことば的」とアノテーターに判断された典型例（保田ほか 2012a）を以下に示す。特徴的と考えられる表現に下線を施した。

（補 1） え。やだなあ。こんな、みっともないの。しかたないから、あきらめて思いっきり口開けた。あーん。先生が調べていって。「奥歯だけ処置しちゃおうね」手にしてる麻酔の注射針から、透明な液体がチュッとこぼれた。ぞっつ。ヴィーン。ドリルの唸り声。いやーッ！！ こわーい！！  
（LBh9\_00139：青山えりか「好きから始まる kiss 物語」）

<sup>5</sup> 図書館サブコーパスからランダムに選出した 1,890 サンプル中、3 名のアノテーター全員が「話しことば的」と判断したサンプルは 12 サンプルにすぎなかった。

で、「語りかけ性」とは異なるものと判断されていた。「話しことば的」と「語りかける」文体とは、必ずしも一致するものではないという結果が得られた。

それでは、「語りかける」文体には、どのような特徴が見られるのであろうか。

まず、「語り（物語）」の特徴として、「歴史的現在形（historical-present）」の出現頻度が高いということがいわれている（Schiffrin 1981, Silva-Corvalán 1983, 池上 1986）。また今村（2007）は、終助詞の「よ」「ね」などは語りかけの度合いが高いものの代表であるが、「のだ」や「わけだ」が文末につく文も、それらがつかないものより「語りかけ度」は高いと考察している。このような特徴的な表現が出現することにより、「語りかける」テキストとして読み手に受け取られることが考えられる。

特徴的な表現の頻度調査では、岸本（2005）が、「ネット日記」の分析において「読み手めあて」で用いられる丁寧体（調査結果における頻度順位 1 位）や伝達態度を示す終助詞（「よ」「ね」：同順位 2 位）が現れやすいことを示した。先の（1）における直感的な「語りかける」印象を与える表現とは、そのような「読み手意識表現」としての特徴を備えたものと考えられよう。

同様に、野田（2012）は、「ブログや軽い文体のエッセイの文章」を分析している。そのような種類の文章では、典型的な書きことばとは異なり「書き手と読み手とのコミュニケーションが意識されている」とし、エッセイ末の「読み手を意識した表現」を取り上げて調査している。この結果、エッセイ末には「読み手の存在を特に意識した表現（丁寧体・終助詞・疑似独話・余韻を感じさせる表現など）」が多く現れることを示している。一方、表現の種類や頻度には個人差も大きいことを指摘している。詳細は後述するが、実際に、（1）で見たような「語りかけ性」が「ある」と判断されたテキストに出現する語レベルの特徴的表現（「読み手」を意識した表現とされる丁寧体や終助詞といった特定の表現）の出現頻度は、サンプルによってはとりたてて高いとも限らず、ばらつきがあった。

以下の（2）は、先行研究で指摘された「語りかける」という印象を受ける表現、または直感的に「読み手を意識した」と考えられる表現を含まないが、アノテーターに「語りかけ性」が「ある」と判断された例である。

- (2) カップリングコンデンサが大きい場合、オレンジ色の側の配線が同じように IC ソケットの足にハンダ付けできればどのように付けても構わない。完成図を見てもらえれば分かると思うが、コンデンサの左の部分は大きくスペースが残してあるので、アキシャルリードのものも基板上に取り付け可能だ。また、大きすぎて基板からはみ出したとしても、特に問題はない。なお、後で説明するが、このコンデンサは無しにも出来る。その場合はこのステップを省き、IC ソケットの足を R C A ジャックに何かの配線で直結する。

(LBt5\_00030：酒井智巳「はじめてつくるプリアンプ」)

本稿では、（2）のような、先行研究で指摘されてきた「語りかける」表現を含まないにもかかわらず、「語りかける」（「語りかけ性」が「ある」）と判断されたテキストを含め、テキストが持つどのような特徴によって「語りかける」という印象が形成されているのか分析を行う。

### 3. 調査対象データ

『現代日本語書き言葉均衡コーパス』(BCCWJ)の図書館サブコーパスに含まれる書籍の10,551サンプル(可変長サンプル)をランダムに並べ替え、人手で文書分類を行った(柏野2013)。作業は、延べ9名で行った<sup>6</sup>。付与した情報は、専門度(5段階:専門家向き、やや専門的な一般向き、一般向き、中高生向き、小学生・幼児向き)、客観度(4段階:とても客観的～とても主観的)、硬度(4段階:とても硬い～とても軟らかい)、くだけ度(3段階:とてもくだけている、どちらかといえばくだけている、くだけていない)、語りかけ性度(3段階:とてもある、どちらかといえばある、とくにない)である。

この情報付与を行った全10,551サンプル(可変長サンプル合計:35,732,431語・サンプル平均:約3,400語)のうち、外国語、数式やプログラミング言語、データの羅列などにより、一定量の本文が認められず判断対象外とされたサンプルを除く8,821サンプル(30,479,196語)について、「語りかけ性」の有無が付与されている。柏野(2010)では、「あなた」や「みなさん」などの呼びかけ表現や、「でしょう」「ではないでしょうか」といった問いかけや相づちを求めるような文末表現など、「直接的な語り」と呼べるような表現が含まれるテキストを「語りかけ性を有するテキスト」と呼んでおり、原則としてアノテーターはこの定義に従っている。

なお、本稿の分析に際しては、会話文の影響を考慮するため、アノテーターが「小説」であると判断したサンプルを除外することとした。8,821サンプル中、小説ではないと判断された5,901サンプル(19,223,693語)を本稿の調査対象データとする。

データの内訳を表1に示す。

表1 調査対象データ

「語りかけ性」がとてもある	738 サンプル	2,213,991 語
「語りかけ性」がどちらかといえばある	1,087 サンプル	3,411,718 語
「語りかけ性」はとくにない	4,076 サンプル	13,597,984 語

テキストの形態素解析には、「MeCab 0.993 + UniDic 2.1.0」を用いた。また、特徴的な表現の検索に「中納言 1.1.0<sup>7</sup>」を用いた。本稿の分析結果に示す品詞情報や語彙素等の要素は、解析結果に基づく。

そのほか、情報付与を行ったアノテーターには、判断に際して参考にした表現などについて、適宜コメントを記述するよう指示した。コメントは、アノテーターごとに記入量が異なるものの、それぞれの作業サンプル数の2%～5%に付されていた。

### 4. 語りかけるテキストに頻出する語

まず、調査する書籍データにおいて、「語りかける」文体に頻出する語を調査する。

<sup>6</sup> 判断のゆれの検証と修正のために、全体の1/3は1サンプルにつきアノテーター3名が判断した。残り2/3は、1サンプルにつきアノテーター1名が判断した。

<sup>7</sup> <https://chunagon.ninjal.ac.jp/>, 短単位データ 1.0・長単位データ 1.0

「語りかける」かどうか（「語りかけ性」が「ある」か「ない」か）の分類結果において、語の出現頻度を調査することで、語レベルの特徴的表現を抽出した。

「語りかけ性」が「とてもある」・「どちらかといえばある」・「とくにない」の3分類について、調査対象とした全テキストにおけるすべての語の出現頻度を算出し、それぞれの分類群の頻度上位300語を取得したところ、異なりで合計366語を得た（延べ26,144,828語であり、全テキストにおける69%にあたる）。これら366語各々において、「語りかけ性」の有無群（「とてもある」「どちらかといえばある」を「ある」群、「とくにない」を「ない」群と二分した）でカイ二乗値を求めて $p$ 値を計算し（イエーツの補正あり）、有意水準0.1%以下（ $p < 0.001$ ）で有意差を見たところ、366語のうち302語で有意差が現れるという結果となった。すなわち、助動詞をはじめとする一般的な語が多くを占める上位頻度語の8割以上（全テキストにおいて58%以上にあたる）が、「語りかけ性」の有無に関係していると考えられることになる。「語りかけ性」の有無において、語彙は全般に異なる可能性があるということである。

そこで、とくに差の見られる語に着目し、次頁の表2にカイ二乗値の高い語<sup>8</sup>を上位から有無群別の出現数とともに示す。表内の「語りかけ性」が「ある」群に特徴的と考えられる語は、該当行を網掛けした。

表2に見られる有意差の大きな語は、「語りかけ性」の判断において、特徴的な表現としてとらえられる可能性が高い。「語りかけ性」が「ある」群においてとくに頻出している助動詞の「ます」「です」は先行研究で指摘された「読み手めあて（岸本2005）」表現であり、同様に助動詞の「た」が少ないのは、「ある」群で現在形文末が多く用いられていることを示唆する。また、準体助詞の「の」は「のだ」文末の多いことを示している。そのほか、接頭辞「御」や代名詞「貴方」も、直感的に「語りかける」という印象につながると認識しやすい語であろう<sup>9</sup>。

<sup>8</sup> 語の集計は語彙素で行っているため、表記の違いや活用形の違いは語彙素に統合されている。たとえば「ます」は「ましよう」「ませ（ん）」を含む。

<sup>9</sup> 語の活用形についても「語りかけ性」有無群などで同様の調査を行ったが、詳述するほどの成果がみつからなかったため、本稿では省略する。なお、終止形撥音便（「ません」に限定されるため、表2における「ます」+「ず」と一致する）でカイ二乗値が4695.8と大きく「ある」群に優位であったほか、意志推量形（「でしょう」「あろう」のような推量が多いが、文脈上勧誘と読み取れる表現を含む）で同676.2、仮定形（5.3で後述する文脈と関連する）で同615.0と「ある」群に優位であった。命令形などについても有意差は見られなかったが、意志推量形で有意差が見られたことは、勧誘の「ましよう」のような表現が多いことと関連すると考えられよう。



表2 「語りかけ性」有無群で有意差の見られる語（上位のみ）

語彙素	品詞	ある	ない
ます	助動詞	60,424	22,541
です	助動詞	43,964	17,211
た	助動詞	117,238	347,253
の	助詞-準体助詞	55,288	98,321
ず	助動詞	15,747	21,705
御	接頭辞	11,168	13,894
の	助詞-格助詞	247,173	665,862
人	名詞-普通名詞-一般	10,862	14,743
年	名詞-普通名詞-助数詞可能	10,454	41,412
千	名詞-数詞	5,105	23,980
ね	助詞-終助詞	4,063	4,046
有る	動詞-非自立可能	53,074	156,869
言う	動詞-一般	47,513	91,905
自分	名詞-普通名詞-一般	7,989	10,962
良い	形容詞-非自立可能	8,798	12,597
九百	名詞-数詞	3,226	15,835
貴方	代名詞	2,083	1,565
	総語数	5,625,709	13,597,984

以下に(1)(2)を再掲する。表2に見られる「語りかけ性」が「ある」群に多い(とくに有意差がある)語を下線で示した。

- (1) お年寄りや赤ちゃんを連れた人、重い荷物を持った人には、席をゆずりましょう。席をゆずるのは、あなたが立っていても平気なときだけよ。病気やケガをしているときは無理をしなくてもいいの。また、ゆずられた方は、「ありがとう」とお礼を言い、会釈をします。また自分が降りるとき、あるいは替わってくれた人が降りるときも、もう一度軽く「ありがとう」と言うのをおわすれなくね。
- (2) カップリングコンデンサが大きい場合、オレンジ色の側の配線が同じようにICソケットの足にハンダ付けできればどのように付けても構わない。完成図を見てもらえれば分かると思うが、コンデンサの左の部分は大きくスペースが残してあるので、アキシシャルリードのものも基板上に取り付け可能だ。また、大きすぎて基板からはみ出したとしても、特に問題はない。なお、後で説明するが、このコンデンサは無しにも出来る。その場合はこのステップを省き、ICソケットの足をRCAジャックに何かの配線で直結する。

表2に見る「語りかけ性」の「ある」群に優位な語が、「語りかけ性」が「ある」と判断された(1)には見られるが、(2)には含まれていない。

また、たとえば「語りかけ性」が「ある」群に頻出する助動詞の「です」「ます」が用いられていれば「語りかけ性」が「ある」と判断されやすいテキストと考えられるが、丁寧体が用いられていれば常に「語りかけ性」が「ある」と判断されるわけでもない。(3)は、「です」「ます」が用いられていても「語りかけ性」が「ある」とは判断されていない例である(丁寧体表現に下線を施した)。

- (3) 3列目火山列を構成する尻別は、浸食された安山岩質成層火山です。その南方の洞爺カルデラは、約10万年前に大量の流紋岩質火砕流(洞爺火砕流)が噴出して形成されました。約4～5万年前に洞爺カルデラの中央部に安山岩質溶岩ドーム群からなる中島火山が噴出しました。その後、約1万5000～2万年前ごろから、玄武岩～安山岩質の小型成層火山である有珠火山がカルデラ南西縁で噴火をはじめ、数千年前に活動をいったん終了しました。  
(LBm4\_00030: 中川光弘・高橋正樹「北海道の火山」)

(2)や(3)を見ると、「語りかける」という印象が、必ずしも頻出する語によって形成されているのではないと考えられる。それでは、何によって読み手(アノテーター)はテキストから「語りかける」という印象を受け取ったのであろうか。

「語りかけ性」は「とくにない」と判断された群において数詞と助数詞が頻出することや、「自分」「人」(これらの名詞は(1)にも含まれている)などの名詞の出現数に差異が見られることは、「語りかけ性」の有無により話題の違いがあることに関係すると推測される。「語りかけ性」の有無群別で有意差の見られる語には、表2以下でも各種数詞や「経済」「国」「政府」「社会」「歴史」「教育」(このような語は「語りかけ性」が「ない」群に多い)、「子供」「大切」「易い」「食べる」(同、「語りかけ性」が「ある」群に多い)などの名詞がある。これらはテキストの扱う話題の違いを反映する語であると考えられる。この点についてはあらためて後述する。

## 5. 書籍における「語りかける」文体とはどのようなものか

語の頻度のみでは、「語りかける」という文体がどのようなものかはとらえ難いといえる。それでは、書籍において「語りかける」文体とはどのようなものであるのだろうか。本節は、「語りかけ性」が「ある」という判断がどのような根拠で行われたのかについて、アノテーターが作業時に記述したコメントを収集し、アノテーターが判断する際にどのような表現を判断の根拠として用いたのかを分析する。

判断を行ったアノテーターは、「語りかけ性」が「とてもある」「とくにない」よりもむしろ「どちらかといえばある」という判断の際、コメントを記述しやすい傾向があった。まれに「とてもある」と判断されたサンプルにコメントがある場合は、「明らかにあなたに語りかけている文体」などの記述にとどまっている。「とてもある」とまではいい難いが、「語りかける」と感じた場合に「どちらかといえばある」を選択し、その判断根拠を積極的に示そうとしていたのだと考えられる。そのため、アノテーターコメントから得られた表現は、「とてもある」群よりもむしろ「どちらかといえばある」群のテキストで判断の根拠とされた可能性がある。



そこで、まず「どちらかといえばある」群に特徴的な表現の調査を行い (5.1)、実際にアノテーターが意識した表現を分析する (5.2)。さらに、「教示的」というアノテーターコメントをもとに、語レベルにとどまらない表現について調査する (5.3)。

### 5.1 「どちらかといえば」語りかける表現

アノテーターのコメントが「語りかけ性」が「どちらかといえばある」という判断に多く記述されていたことから、「語りかける」表現は「語りかける」程度の差により違いの見られる可能性がある。そこで、「語りかけ性」が「どちらかといえばある」群にのみ特徴的な表現をとらえるため、「語りかけ性」が「とてもある」群と「どちらかといえばある」群に大きな差があり、かつ「とくにない」群とも異なる出現率を示す表現に着目する。

図1は、上位頻度語（前掲366語）について、分類群ごとに出現率の偏差を標準偏差で割った値のうち「とてもある」群と「どちらかといえばある」群間で距離が1.8以上（外れ値）となった語を示し、「とてもある」群に出現率の高い場合はプラス、「どちらかといえばある」群に出現率の高い場合はマイナスで表したものである。外れ値の差が大きいほど、「とてもある」群と「どちらかといえばある」群での出現傾向が大きく異なることを意味する。但し、「とてもある」群と大きな差があっても「どちらかといえばある」群が「とくにない」群に類似した傾向である場合は、語りかける程度の差が「とてもある」群にのみ突出する特徴の可能性が考えられる。そこで、「どちらかといえばある」群と「とくにない」群間距離が0.5以上のものに限定する。

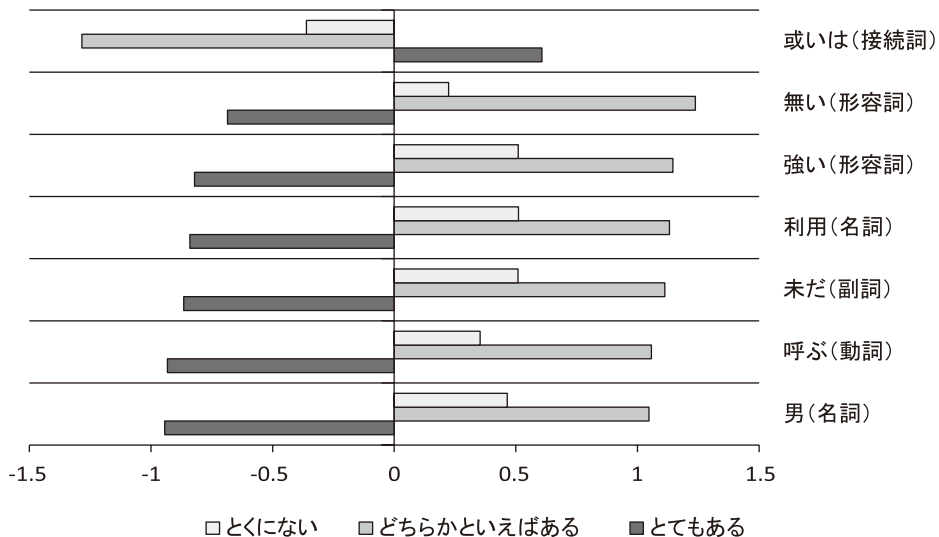


図1 「語りかけ性」分類群における外れ値調査結果の抜粋

「どちらかといえばある」群においてのみ出現率が顕著な（「とてもある」群と差があり、かつ「とくにない」群とも差がある、「どちらかといえばある」群固有の特徴）語には、図1に見える「強い」「無い」をはじめ形容詞が多いほか、「利用」「会社（距離 1.6）」のような名詞、そのほか「或いは（接続詞）」「未だ（副詞）」「呼ぶ（動詞）」「ながら（接続助詞：距離 1.5）」などがある。

これらを見ると、表現によって語りかける程度の差が感じられるというよりも、テキストの話題に違いのある可能性が推測される。アノテーター（読み手）は、「どちらかといえばある」群のほうが、直感的に「語りかける」と受け取られる語が見当たらずとも、テキストの話題から「語りかける」印象を受けた可能性もあろう。

また、ここで得られたような語（形容詞や動詞など）は、語のみでは「語りかけ性」の有無群で差があるとはとらえ難いが、文脈上「語りかける」と感じることに影響を与える場合があると考えられる。

## 5.2 読み手に意識される表現

では、次にアノテーターのコメント内容を見てみたい。具体的には、以下のようなコメントの記述があった。表記等は原文の通りである。

- ・「べきである」「ほしい」「求められる」などが、どちらかといえば語りかけている印象。
- ・「考えていただきたい」などあるので語り性があるとすべきか迷う。
- ・「…だろうか」といった問いかけ表現が多めなので、少し語り性がある印象は受ける。
- ・「われわれ」が複数回出現するため、若干語りかけているようにも感じられる。
- ・一か所「私たち」が使用されているが、「人間一般」のことを表す使い方。
- ・「～思います」「～思うのです」が多く、どちらかといえば語りかけている印象がある。
- ・「～ましょう」「～のです」「ことです」が多く、どちらかといえば語りかけている印象になる。
- ・「～読んだ方は」「と思うはずだ」や、「～ということだ」の繰り返しなど、どちらかといえば語りかけている感じがある。
- ・可能形「楽しめます」「与えます」などが幾分語りかける感じである。
- ・独り言めいた「どうであろう。」があるので、少し語り性があるような印象を受ける。

アノテーターコメントから得られた「語りかける」印象を与える表現を表3にまとめる。

表3 アノテーターコメントから得られた「語りかける」印象を与える表現

<ul style="list-style-type: none"> <li>・昔話に特徴的な表現</li> <li>・教示的表現： <ul style="list-style-type: none"> <li>- 希望：「～てください」「～てほしい」「～したい」</li> <li>- 注意・禁止：「～ように」「～に注意」「～はいけない」</li> <li>- 勧誘：「～でござらん」「～しよう」</li> <li>- 可能：「～はず」「～できます」</li> <li>- 評価：「～が大切」「～を要する」「～が便利」「～が狙い目」</li> </ul> <p style="text-align: center;">(そのほか「よい」「悪くない」などの形容詞もコメントあり)</p> </li> <li>・人称：「私たち」「我々」 <p style="text-align: center;">(「が多い」のような傾向や「単純な人称でない」というコメントもあり)</p> </li> <li>・文末表現： <ul style="list-style-type: none"> <li>- 「のである」「わけです」「からです」「ものです」など</li> </ul> <p style="text-align: center;">(「が多い」のような傾向もコメントあり)</p> </li> <li>- 読み手を想定した表現： <ul style="list-style-type: none"> <li>「いただく」「申し上げる」「～の方 (かた)」「ございます」</li> <li>「あげる」「させてください」「お聞かせすることにしよう」</li> </ul> </li> <li>- 婉曲表現： <ul style="list-style-type: none"> <li>「～と思う (見える／感じ)」「～でしょうか」「～だろう」</li> </ul> </li> <li>- 読み手の判断を想定した表現： <ul style="list-style-type: none"> <li>「思うかもしれません」「おわかりのことでしょう」</li> </ul> </li> </ul>
--

表3にまとめた表現はいずれも「語りかけ性」が「ある」群において出現率の高い表現ではないか、あるいはコーパス全体において出現数が少ないため、「語りかけ性」が「ある」と判断されたテキストからの機械的な抽出は難しい表現であるといえる。これらは、「文脈上意識される表現」、「表現群としてまとまると意識される表現」の2種類に分類できると考えられる。以下、得られた表現の分析を行う。

### 5.2.1 文脈上意識される表現

表3に挙げた表現の一例として、以下の(4)のように、アノテーターに指摘された「私たち」「我々」のような人称代名詞が含まれている場合がある。これらは文脈によって、読み手との一体感を生じさせることがあると推測される。

- (4) 暑さが増して、私たちはできるだけ涼しくなれる洋服を着ることでしょう。しかしマルチーズのように長毛犬は、自分の着ている毛を取ってしまうような調節はできません。毛玉だらけにさせてしまわないように、毎日クシを入れてあげましょう。

(LBa6\_00006：高橋宏美「マルチーズの飼い方」)

なお、「私たち」「我々」の出現率を見ると、「とてもある」群で0.018%、「どちらかといえばある」

群で0.018%、「とくにない」群で0.029%となっており、「とくにない」群のほうがむしろ高い。語の出現頻度のみからは看取しにくい表現であろう。

また、以下の(5)は小説の例であるが、それまでのテキストを語りかけていなかったように読んでいたとしても、著者と読み手を含む（一体感を生じさせると考えられる）「わたしたち」が現れることによって、「語りかけ性」が「ある」と判断された例と考えられる。

- (5) スーツはフックのどうしようもなく趣味のわるい衣装を材料にしてウエンディがいよいよやながら作っているものだった。あとになって、みんなのあいだではこんなうわさ話がささやかれた。新しいスーツをはじめて着た晩、ピーターは長いことキャビンにいて、フックの葉巻用パイプをくわえ、人さし指だけ突き出した片手を下向きに曲げて鉤爪に似せると、それを高く上げて人をおどすかっこうをしていた、と。でも、わたしたちはこの船ばかりながめているわけにはいかない。もうずいぶん前に三人の作中人物が薄情にも空を飛んで逃げていってしまい、すっかりさびしくなったあの家にもどらなくては。

(LBjn\_00026: ジェイムズ・M・バリ (著) / 佐伯泰樹 (訳) 「ピーター・パン」)

(5) のような例は、読み手と著者を含む「わたしたち」という語の出現によるというだけではなく、「わたしたち」の強い印象によって、「語りかける」と受け取られた可能性がある。

「わたしたち」に類似した例は以下のようなものもある。次の(6)は、現在形の文末が頻出している(前掲4節参照)ことや「やめた方がよい」という勧誘表現(前掲5.2参照)によって「語りかけ性」が「ある」と判断された可能性もあるが、「筆者は思う」(下線を施した)という表現によって著者が直接前面に出てくることにより、「語りかける」印象を強く与えていると考えられる。

- (6) ここでいう濃度比を「濃縮係数」と称している本も多いが、上述のように濃度比単独では濃縮を意味しないから、これを濃縮係数と呼ぶのをやめた方がよいと筆者は思う。英語の‘concentration’には、「濃度」という意味の他に「濃縮すること」という意味もあるので、おそらく混同されて使用され始めたのであろう。(LBf4\_00036: 西村雅吉「環境化学」)

松本(2008)は、現実世界の読者に作者が直接語りかけるテキストに着目している。本稿の調査において「作者」「筆者」「著者」を含む用例は、「語りかけ性」が「ある」群における出現率は10%、「ない」群では13%であり、「語りかけ性」が「ない」群により多く出現する傾向が見られた。また、取得されたこれらの用例のうち筆者(類)が語り手であると読み取れる例は6%にすぎない。しかし、(6)に見られるように、テキスト内に筆者(類)が直接的に「顔を出す」と読み取れる例においては、「語りかけ性」が「ある」との判断がなされている傾向があった。

このほかにも、テキスト内における出現頻度が低くとも、文脈上「語りかける」という印象を多く与えやすい特徴的な表現はほかにも指摘できる。

- (7) すべると危険です。釣り上げられて跳ねた魚や自分の釣りバリでケガをするかもしれません。ベテランは、釣り物によっては真夏でも釣り用ブーツで足を保護しています。小物釣

りでも滑らないスニーカーなどを履きたいものです。

(LB07\_00034：井田玲子「超明解！船釣り入門 ABC」)

この(7)には「ベテラン」という語が現れている。ノウハウを知りたいという目的を有する素人の読み手にとって、目指すべき姿としての「ベテラン」の言動に関する叙述は、読み手に対して規範的であるといえる。結果として、「だからあなたも倣いましょう」というように、(具体的にそのような記述がなくとも)読み手に対する指示や勧誘を喚起するため、「語りかける」という印象を与える要素になる可能性がある。(6)の「筆者」との類似性も考えられる語であろう。

このように、ごく少数の要素であっても「語りかける」印象に強い影響を与える特徴的な表現が現れることで、「語りかけ性」が「ある」という判断のなされることが考えられる。

### 5.2.2 表現群としてまとめると意識される表現

アノテーターが「語りかけ性」の有無を判断する際に根拠とした表現は、文脈上意識されるか文脈上少数であっても印象が強いという語に限らない。

アノテーターに着目されやすい表現において、単独で挙げられるのではなく、複数の種類(例：「のである」「からです」「ものだ」など)が並列的に例示され、「この種の表現が多い」ため、語りかけられている感じがした旨が記述されている場合があった。これらの表現群がまとまったために、読み手の意識に上ったものと考えられる。

これらの表現群の出現率(「のだ」群<sup>10</sup>などの文末例)を、図2に示す。

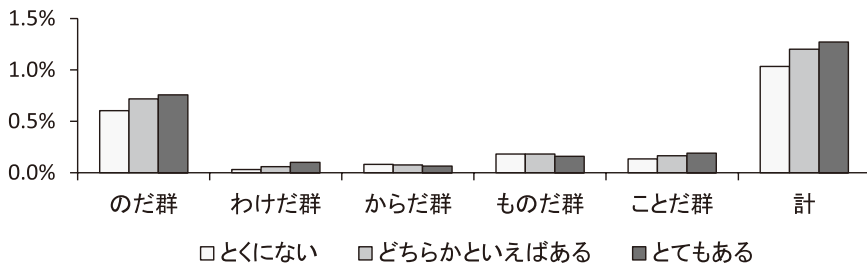


図2 「語りかけ性」の分類群においてアノテーターが多いと感じた表現群の出現率(形態素)

図2に示したうち、準体助詞の「の」を含む「のだ」群については、今村(2007)で「語りかけ度」が高い文末に「のだ」が多いと指摘されており、また4節で見た通り「語りかけ性」有無群における出現頻度でも準体助詞の「の」には差がある。とはいえ、読み手の感覚としては「のだ」群のみがテキストにおける出現割合の高い表現とは感じにくいようであり、コメント内の記述においても「からだ」群「ものだ」群などと並列的かつ複合的な扱いであった。また、「からだ」群や「ものだ」群では、「語りかけ性」の有無群の違いがさらに少ないか、むしろ出現率の高低が逆転している場合もある。但し、これらの表現をまとめると、「語りかけ性」が「ある」群に

<sup>10</sup> ここでは、「のである」「のです」「のだ」などをまとめて「のだ」群(他も同様)とする。

おける出現率が高くなる傾向が強まっている。そのため、読み手が「この種の文末が多い」と感じることで「語りかけ性」が「ある」と判断したことが推測される。

同様に、表3で「教示的な表現」として挙げられた希望や禁止、勧誘のような表現についても、「この種の表現が多い」ため、「語りかけ性」が「ある」と判断したとのアノテーターコメントが得られている。すなわち、表3に挙げたような、昔話に特徴的な表現、教示的表現、人称、文末表現、婉曲表現などのカテゴリーに属する（と文脈上アノテーターの読み取った）表現群が「多い」という印象が、「語りかけ性」が「ある」との判断になったものと考えられる。

直感的に「語りかける」印象を与えると考えられる表現や「語りかけ性」が「ある」群に有意に見られる上位頻度語の含まれていなかったテキスト（2）を再掲する。アノテーターのコメントから得られた表現に下線を施した。さらに、「どちらかといえばある」群に固有（5.1参照：ここでは「とてもある」群との距離 1.5 以上、「とくにない」群との距離 0.4 以上）の特徴語（「ない」「大きい」「過ぎる」）はゴチック体にした。

- (2) カップリングコンデンサが**大きい**場合、オレンジ色の側の配線が同じように I C ソケットの足にハンダ付けできればどのように付けても構わない。完成図を見てもらえれば分かると思うが、コンデンサの左の部分は**大きく**スペースが残してあるので、アキシャルリードのものも基板上に取り付け**可能**だ。また、**大きく**過ぎて基板からはみ出したとしても、特に**問題はない**。なお、後で説明するが、このコンデンサは無しにも出来る。その場合はこのステップを省き、I C ソケットの足を R C A ジャックに何かの配線で直結する。

この（2）には、「どちらかといえばある」群にのみ固有な表現と、アノテーターによって「語りかけ性」が「ある」という判断の根拠とされた表現が含まれていることがわかる。とくに、アノテーターが「教示的な表現」とコメントした表現を複数含んでいる。これらによって、直接的に「語りかける」表現を持たない（2）のようなテキストも、「語りかける」と受け取られたのだといえよう。

### 5.3 文脈として語りかける表現

表3に見たアノテーターに着目された表現は、文脈上意識された表現と考えられるが、いずれも語や句として挙げられる表現であった。それでもなお、どのような文脈によってそれらの表現が意識されたのかという問題が残る。そこで、頻度の高い表現について、語よりも大きなレベルで調査した。

読み手（アノテーター）が「語りかける」と感じた（「語りかけ性」が「ある」と判断した）根拠として「教示的な表現」を挙げる場合があった。実際、前節で（2）にはその種の表現群が含まれていることもわかった。しかし、「教示的な表現」は語レベルのものだけではない。たとえば、読み手の求める条件（例示）があって、その対応方法（結果や評価など）が示されるといふ文脈が考えられる。



これまでに見た (1) では、「病気やケガをしているとき」「無理をしなくてもいい」という「条件」と「対応方法」の関係が述べられている。(2) でも、「カップリングコンデンサが大きい場合」「構わない」をはじめ、「はみ出したとしても」「問題はない」などの対応関係が挙げられている。以下、該当部分に下線を引いて再掲する。

- (1) お年寄りや赤ちゃんを連れた人、重い荷物を持った人には、席をゆずりましょう。席をゆずるのは、あなたが立っていても平気なときだけよ。病気やケガをしているときは無理をしなくてもいいの。また、ゆずられた方は、「ありがとう」とお礼を言い、会釈をします。また自分が降りるとき、あるいは替わってくれた人が降りるときも、もう一度軽く「ありがとう」と言うのをお忘れなくね。
- (2) カップリングコンデンサが大きい場合、オレンジ色の側の配線が同じように IC ソケットの足にハンダ付けできればどのように付けても構わない。完成図を見てもらえれば 分かると思うが、コンデンサの左の部分は大きくスペースが残してあるので、アキシャルリードのものも基板上に取り付け可能だ。また、大きすぎて基板からはみ出したとしても、特に問題はない。なお、後で説明するが、このコンデンサは無しにも出来る。その場合はこのステップを省き、IC ソケットの足を R C A ジャックに何かの配線で直結する。

ここでは、上例に見られたような表現パターンを探索的に取得し、頻度調査を行った。

「語りかけ性」の有無群ごとに、前文脈に条件を提示する部分があり、後文脈にその回答が示されるという表現パターン数種を抽出した。後文脈は 10 語以内に検索語の出現する例を取得したが、取得した用例は人手で確認し、後文脈につながりがない場合、条件+回答になっていないと考えられる場合は対象外とした。調査した表現パターンと結果(表4・表5)を事例とともに示す。

- ①条件(例: ~すると, ~するなら, ~すれば, ~するとき, ~する場合など)+特徴的表現  
前文脈が条件で、後文脈に特徴的表現例(名詞)を含む場合を検索した。
- 前文脈: 動詞+「と」(接続助詞)または動詞・形容詞・助動詞の假定形または動詞や「時<sup>11</sup>」  
「場合」  
後文脈: 「必要」「可能」「大切」「便利」を含む
- (8) 放置自転車禁止と書かれた道路に置く場合は短時間で用足しをすることが必要です。  
(LBb7\_00009: 草森紳一「コンパクトカメラの大冒険」)
- (9) デザイン事務所といえば, 場所のイメージもやはり大切です。  
(LBm3\_00070: 高橋慈子・伊藤華子「家庭をオフィスにする SOHO 読本」)
- (10) 大豆をすぐに使いたいなら, 真空パックのゆで大豆が便利。  
(LBm5\_00032: あんしんネット 21「安心キッチンクッキングノート」)

<sup>11</sup> 語彙素による検索を行ったため、各種表記を含む。以下の語についても同様である。

表4 「語りかけ性」有無群別①出現数  
(調整頻度 1,000 サンプル<sup>12</sup>, 条件+後文脈内要素)

語りかけ性	〈必要〉	〈可能〉	〈大切〉	〈便利〉
ある	11	48	18	43
ない	7	39	6	12

②前文脈が条件で、後文脈に動詞・形容詞を含む場合を検索した。

パターン A 前文脈：動詞+「と」(接続助詞)，後文脈：形容詞(終止形)

パターン B 前文脈：動詞+「時」，後文脈：動詞

パターン C 前文脈：動詞+「場合」，後文脈：動詞

パターン D 前文脈：仮定形，後文脈：動詞

(11) ゆっくり名前を考えたいときなどは、とりあえず出生証明書を書いてもらい、いつでも届出ができるようにしておくと よいでしょう。

(LBk1\_00004：日本文芸社編「赤ちゃんの名前実例全書」)

(12) 壁に御影石などを貼るときは、必ず石の裏側に防水を施しましょう。

(LBg5\_00034：浜口和博「プロも見落とす家づくりの急所」)

(13) 4回以上の支給がある場合は、月々の報酬に含まれます。

(LBr3\_00006：高橋徹編「社会保険・労働保険のすべてがわかる事典」)

(14) 来店時には明確な購買意図はもっていないが、漠然とした形で特定の商品の必要性を認識しつつ、価格その他の条件を満たしていることがわかれば購買に至るケースである。

(LBs6\_00010：竹村和久「価格・プロモーション戦略」)

表5 「語りかけ性」有無群別②出現数  
(調整頻度 1,000 サンプル, 条件別)

語りかけ性	パターン A	パターン B	パターン C	パターン D
ある	353	208	757	5,435
ない	200	121	519	4,122

表4・5から、条件+回答となるような表現が、「語りかけ性」の「ある」群に多く現れていることがわかる。語レベルの表現に限らず、このような構文や文脈が現れていることも、「語りかけ性」が「ある」と判断された要因の一つといえよう。同時に、教示的な表現を多く有することが書籍テキストの「語りかける」文体に関わっているものと考えられる。

<sup>12</sup>「語りかけ性」が「ある」群と「ない」群でサンプル数に差があるため、ここでは1,000サンプルあたりの出現数に調整した数を示す。

## 6. 語りかけるテキストの特徴的な書籍の種類

広く文脈を考えると、テキストの属する書籍自体の性質もまた手がかりになり得るだろう。

ここまでで見た「語りかける」文体に関わる表現には、「教示的」という特徴的な表現があった。それらは、教示的な書籍（いわゆるハウツー本）に含まれていることが推測される。いわゆるハウツー本に「語りかける」表現は多く現れているのであろうか。

### 6.1 語りかけるテキストを有する書籍

BCCWJ に収録された書籍のテキストには、図書分類コードのうち、NDC（日本十進分類法）と C コード（図書分類コード）が付与されている。これらはそれぞれテキストのジャンル情報の一つとみなすことができるだろう。「語りかける」という文体がどのような話題の書籍に出現する傾向があるのか、NDC と C コードに基づいたジャンル情報との関係から確かめる。

まず、NDC を見てみたい。ここでの NDC は、3 桁の数値で表され、各桁でそれぞれ 10 ずつに分類されるものとする。図 3 に、「語りかけ性」の有無群における NDC ごとの分布（1 桁目「類」）を示した。「語りかけ性」が「とてもある」「どちらかといえばある」群において比較的高い比率であるのは、「1（哲学）」と「4（自然科学）」、「5（技術・工学）」、「8（言語）」とわかる。

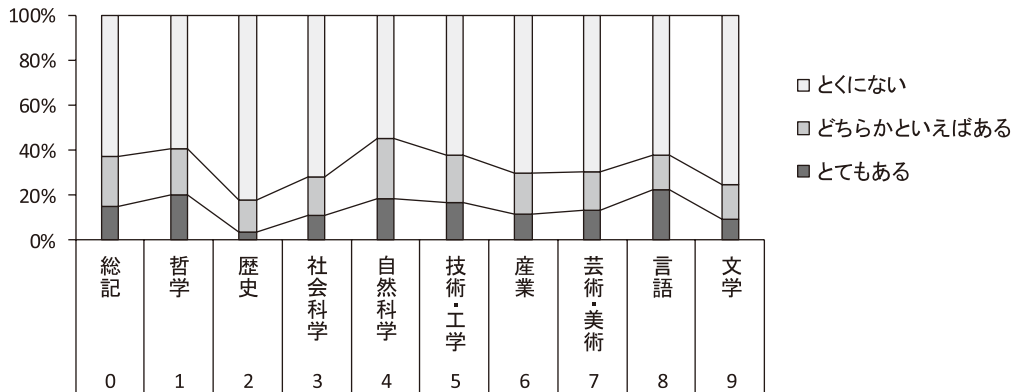


図3 分類群別 NDC 分布 (1 桁目)

図 4 には、図 3 で「語りかけ性」が「とてもある」「どちらかといえばある」群の比率が比較的高かった「4（自然科学）」について、下位区分の分布を示す（図 4 では、「とてもある」「どちらかといえばある」群と「とくにない」群で有無とした）。図 4（「4（自然科学）」の 2 桁目「綱」まで）では、「41（数学）」「49（医学・薬学）」で「語りかけ性」が「ある」群の比率が他分類に比べて高いことがわかる。また、「1（哲学）」では「15（倫理学）」（「ある」群 58.8%、「ない」群 41.2%）、「5（技術・工学）」では「59（家哲学・生活科学）」（「ある」群 53.9%、「ない」群 46.1%）で差が目立った。「8（言語学）」では、「8（言語学）」の 41.4% を占める「81（日本語）」<sup>13</sup> で、「ある」群 35.4%、「な

<sup>13</sup>『はがきの書き方』『敬語マニュアル』『文章術』などの書籍が「81（日本語）」に含まれる。

い」群 64.6%であった。これらのテキストを確認したところ、「語りかけ性」が「ある」と判断されたテキストは、啓蒙書や生活に関わる内容を扱う書籍に多く現れていることがわかった。

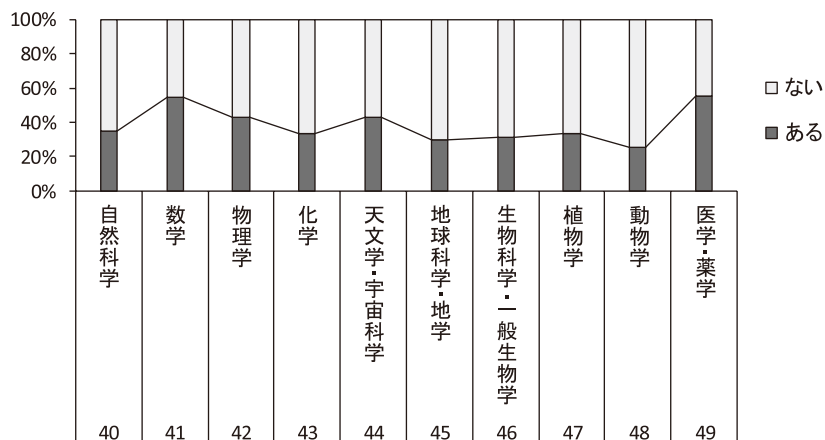


図4 分類群別 NDC 分布例 (「4 (自然科学)」2 桁目まで)

以下の (15) は「15 (倫理学)」の例であり, (16) は「41 (数学)」, (17) は「49 (医学・薬学)」の例である。

- (15) まず、この基本をしっかりと自覚しておかないと、仕事と家庭に対してどちらも中途半端になり、そのつもりもなかったのに家庭崩壊、という悲劇に陥ってしまいます。もちろん女性の中にも、男性的な人、経営能力の高い人、働くことが何よりも好きな人がいます。そうしたことを考えると、すべての女性が“家庭”を優先させるとはいいきれませんが、ただ、しっかりした夫がいる場合には、ごく自然な時期に、仕事から離れていくことは悪い選択ではありません。 (LBi1\_00025: 桜井秀勲「女が30代で自分を変える生きかた」)
- (16) まず、ちょうど半分の一五リットルの水を入れてかき混ぜてから、一五リットルの水を捨てます。これで古い水の割合は $2/3$ になりました。さらに、残りの一五リットルの新しい水を注いで、かき混ぜてから再び一五リットルを捨てます。こうすると、古い水の割合は $4/9$ となります。もっと、もっと古い水の割合を減らすことができないでしょうか。 (LBd4\_00018: 岡部恒治「絵でわかる微分と積分」)
- (17) 第2章でひざ関節は超精密機械であることを紹介しました。中でも、ひざ関節の軟骨と関節液は絶妙のコンビネーションで、魔術師的な働きをすることも、お話ししました。その強靱で柔軟な軟骨のすり減りが始まるというのです。何が起こったのでしょうか。 (LB04\_00034: 高山美治「ひざの痛みをとる・治す」)

同様に、先に挙げた (1) と (2) はともに「5 (技術・工学)<sup>14</sup>」に分類されていたテキストであり、

<sup>14</sup> 2 桁目まで見ると、(1) が「59 (家政学・生活科学)」, (2) が「54 (電気工学)」である。

いずれも「語りかけ性」が「ある」と判断されるテキストの多いジャンルであったことがわかった。ここで見たジャンルのテキストは、「語りかける」性格を持ちやすいことが考えられよう。但し、「語りかける」テキストが多く現れるジャンルの書籍が、ハウツー本であるとは限らない。そこで、「語りかけ性」が「ある」と判断されるテキストが、いわゆるハウツー本に多く現れるものであるかを調べるために、「Cコード（図書分類コード）」を確認する。

Cコードは4桁の数値によって表されるコードで、1桁目が販売対象、2桁目が発行形態、3・4桁目が内容を表す。図5は、販売対象と発行形態を「語りかけ性」の分類群ごとに示している。

図5から、「語りかけ性」の「ある」群では、販売対象として実用書、発行形態として新書の割合が高くなっていることがわかる。

実用書には、一般に啓蒙書や指導書類が多く含まれると考えられることから、「語りかけ性」が「ある」群のテキストが現れやすいのだと考えられる。先に見たNDCの「1（哲学）」、「4（自然科学）」などにハウツー本が多く含まれることとの並行した関係をとらえることができるであろう。

また、新書という形態は、「もと、解説的な教養書を中心とした（新明解国語辞典第7版）」ものであるため、ハウツー本の多い可能性が考えられる。

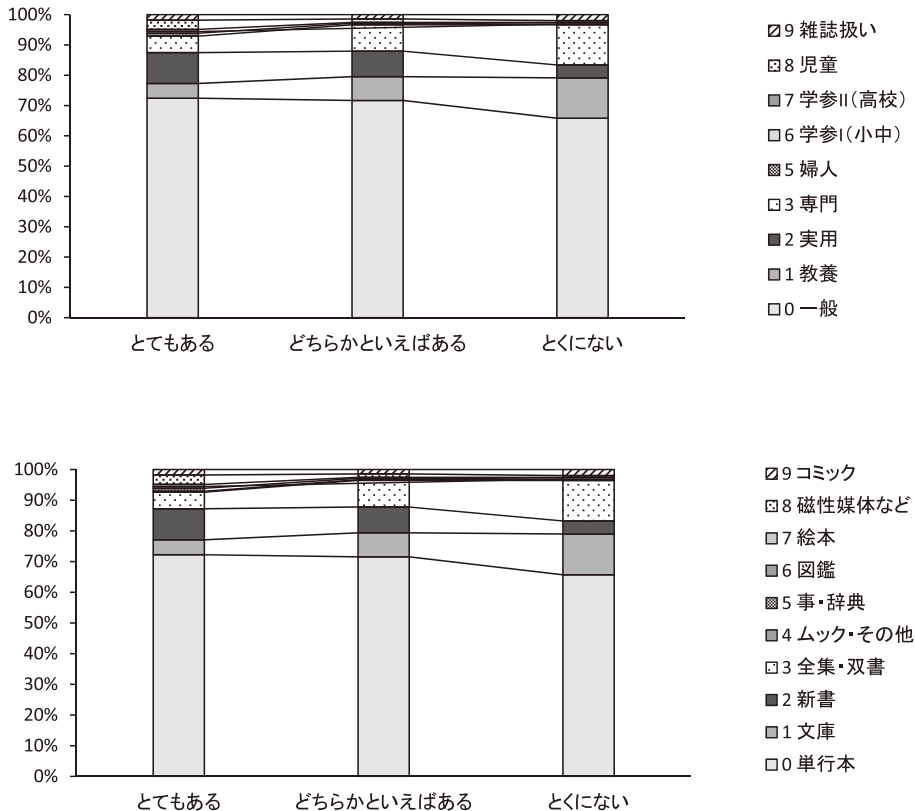


図5 分類群別Cコード分布（上：1桁目（販売対象），下：2桁目（発行形態））

## 6.2 語りかけると推測される書籍タイトル

書籍のタイトルは、その内容を代表して示すものと考えられる。たとえばハウツー本であれば、読み手が予め目的意識を持ってテキストを読むことが推測されることから、テキスト内容の代表たるタイトルには、読み手の知りたい回答の要求に応える旨を明示してあることが期待されよう。本研究の用いたアノテーション結果においても、アノテーターのテキスト閲覧・作業時に書籍タイトルが示されているため、アノテーターは書籍タイトルの影響を受けている可能性がある。

そこで、書籍タイトルから「語りかける」テキストであるかという推測が可能か調査を行った。調査ではまず、2名の作業員<sup>15</sup>が、本稿の調査で用いたサンプルに含まれる書籍タイトルから本文テキストが語りかけているか否かを推測した。

まず、書籍のタイトルと「語りかける」テキストとの関係を見たい。タイトルから「語りかける」テキストであると推測した結果と、実際の「語りかけ性」有無群（実際にテキストを読んで「語りかける」か否か判断された結果）との対照を行った。タイトルからの推測を行った作業員2名の判定の一致率は77.3%で、概ね一致する傾向にあった。また、本文から「語りかけ性」が「ある」と判断されたテキストのうち、タイトルから2名の作業員が一致して「語りかける」と推測したものは、50.5%であった。タイトルから「語りかける」テキストであると推測されたとしても、実際に「語りかける」文体が用いられているとは限らないが、半数程度はタイトルから「語りかける」テキストであることの予測が可能ということになる。

図6にタイトルから「語りかける」テキストであると予測した結果（タイトル判定結果）と実際に内容を読んで「語りかけ性」が「ある」とアノテーションした結果を、NDC別に対照した結果を示す。縦軸は、それぞれのNDC別分類における全サンプルのうち、「語りかける」（「語りかけ性」が「ある」）と判断された割合を示している。

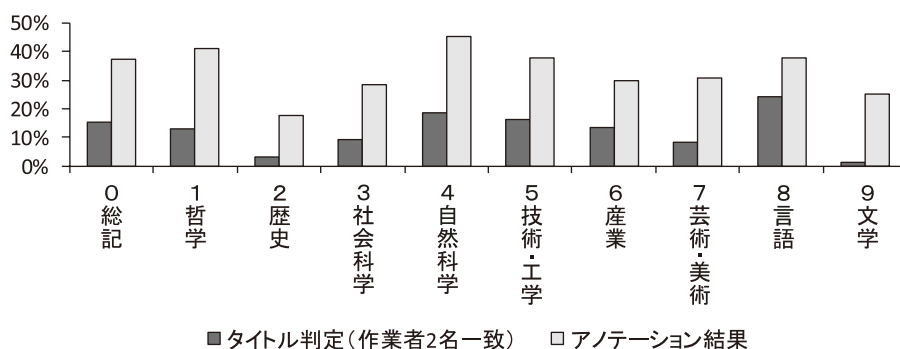


図6 タイトルからの判定結果と本文からの判定（アノテーション）結果（NDC別）

図6を見ると、「5（技術・工学）」と「6（産業）」「8（言語）」では、タイトルから「語りかけ」と予測した結果が、本文からの判定結果（アノテーション結果）に近づいている傾向がある。

<sup>15</sup> 本稿の第一著者と第三著者が作業を行った。



一方、「7 (芸術・美術)」「9 (文学)」などではタイトルからの判定結果の割合が低い。とくに「9 (文学)」の随筆やエッセイ<sup>16</sup>、「7 (芸術・美術)」に含まれる芸能人やスポーツ選手のエッセイは、タイトルからの予測が困難であったと考えられる。タイトルからの予測が困難であっても、本文を見ると、先行研究で指摘されてきた「ブログや軽い文体のエッセイの文章 (野田 2012)」には、「語りかける」テキストが多く含まれるためであろう。

また、「語りかける」と作業者がタイトルから判定した場合、タイトルに含まれるどのような表現によって推測したかを分析した。

実際の書籍タイトルの例には、『みんなでつくるビオトープ入門』『リクガメが100% 喜ぶ飼いや遊び方』『商標登録の実務がよくわかる本』『強制執行の仕方と活用法』『部屋と押入れの整理術』『はじめてのリバー・トレッキング』『大倒産時代の生活防衛マニュアル』『不登校母親にできること』『自分でできる住まいの修理と演出』『人に好かれることばレッスン』『髪に悩んでいるあなたへ』などがあつた。

タイトルから「語りかける」テキストと推測した根拠として、作業者の挙げていた表現を表6に示す。

先の表3に見た表現がタイトルに含まれている場合も観察され、ハウツー本であることがタイトルによって明示されている傾向を見て取ることができる。書籍タイトルから本文が「語りかける」と推測可能なテキストは、ハウツー本らしいタイトルの場合が多いといえる。これまでに挙げた例でも、(1)は『魅女ってみませんか』、(2)は『はじめてつくるプリアンプ』など、ハウツー本らしいタイトルが見られている。書籍のタイトルからハウツー本であることが予測され、読み手が教えられる目的で教示的な内容を期待するために、「語りかける」という印象が生じやすい可能性が考えられる。

表6 本文が「語りかける」と推測された書籍タイトルに含まれる特徴的表現

- |  |
|--|
| <ul style="list-style-type: none"> <li>・「語りかける」テキストと判断される根拠となる表現 (4節と5.2参照)             <ul style="list-style-type: none"> <li>- 呼びかけ (「あなた」など) を含む</li> <li>- 可能・勧誘・命令・評価・希望などを含む</li> </ul> </li> <li>・対象読者を示す表現             <ul style="list-style-type: none"> <li>- 「～のための」</li> <li>- 初心者向けであると示す「入門」「はじめて」「やさしい」など</li> <li>- 方法解説であると示す「テクニック」「レッスン」「手引き」「教える」など</li> </ul> </li> </ul> |
|--|

また、芸能人やスポーツ選手のエッセイであっても、ハウツー本と同様に、読み手が前提として教示を期待することによって、積極的に「語りかける」と受け取る場合もあり得よう。すなわち、先行研究で指摘されてきた書き手による「読み手めあて (岸本 2005)」の表現のみならず、読み手の前提的態度が関わっている可能性も考えられるということである。

<sup>16</sup> 本稿では小説外のテキストに限定した分析を行っているため (3節参照)、「9 (文学)」に含まれているのは、主に随筆・エッセイ・文芸評論・研究書となっている。

## 7. まとめ

本稿は、書き手が読み手に対して直接語りかけているような印象を与えるテキストの文体特徴を、BCCWJの図書館サブコーパス（書籍サンプル）に付与された「語りかけ性」のアノテーション結果を用いて調査した。分析の結果、「語りかける」文体には、以下のような要素が寄与していることがわかった。

- ・直接的に語りかけていると受け取られる表現・頻出する表現（2節，4節参照）がある。

エッセイやブログなどの文末に多い「読み手めあて」と考えられてきた表現（岸本 2005, 野田 2012）などはこの種であり、「です・ます（助動詞）」「よ・ね（終助詞）」や、呼びかけ表現や問いかけや同意を求める表現、「あなた（代名詞）」「ね（終助詞）」など（柏野 2010）もある。

- ・一般的に「語りかける」文体に特徴的な表現とは呼び難く、語の出現頻度としても高くないものの、文脈上語りかけていると受け取られる語や表現（5.2 参照）がある。

読み手が当該テキストにおいて「語りかける」という印象を受けるもので、読み手を想定していると読み取れる表現の「私たち」「いただく」「思うかもしれません」などや、「からだ」・「わけだ」群のような文末をはじめ、教示的な表現との印象に関わる「～てほしい」（希望）・「～はいけない」（注意・禁止）・「（～すると）よい」（評価）などもこれにあたる。また、書き手との関係（「教えられる」関係）が設定されることによって、読み手が語りかけていると感じる表現として、「～できる」「～と思う」「～だろう」などもある。

また、頻度の低い中にも強く語りかける印象を与える表現があり、典型的には、「筆者」あるいは「ベテラン（教師的位置づけ）」のように、直接的な情報伝達であることが示されている例が見られる。また、この場合の「ベテラン」のような表現は、「目指すべき姿であるからあなたも倣いましょう」のように教示的な態度を前提とするものであり、「語りかける」文体をつくる特徴的な表現になり得ると考えられる。

- ・語レベルにとどまらず文脈として特徴的な表現（5.3 参照）もある。

たとえば、教示的な態度として、読み手の求める条件とそれへの回答を述べる表現が用いられるものである。語レベルでは抽出し難い場合が多いが、読み手へ回答を与えることで「語りかける」印象に関係している可能性が考えられる。「～するなら+～しましょう」「～する場合+～なります」「～すると+よいでしょう」などのパターンは、「語りかける」と判断されるテキストに多く含まれている。

- ・「語りかける」文体はハウツー本に多く見られ（6節参照）、教示的な態度を強調するために用いられている可能性が考えられる。読み手の態度が「語りかける」という印象に関わっている可能性がある。

NDCやCコードなどの分類から、回答の提示がある・書き手が読み手に教示を与えることを前提とするハウツー本に、「語りかける」文体が多く観察されることがわかった。テキスト内容が代表されていると考えられる書籍のタイトルが、そもそも「語りかける」表現類を含む

か、あるいは「入門」「わかる」のようにハウツー本であることを明示している場合も見られる。

以上のように、書籍においては、いわゆるハウツー本をはじめとするような、教示的な態度を示すテキストに「語りかける」文体が出現しやすい傾向があり、それらは必ずしも直感的に「語りかける」ととらえられる表現ばかりで形成されているのではないということが明らかになった。これらの要素・要因が多様な形で出現することによって、テキストが「語りかける」という読み手の印象がつけられているということがわかった。

従来、「語りかける」という文体は、エッセイやブログなどに散見されるくだけたテキストにのみ出現しやすいと考えられてきた。実際、「とても」語りかける（「語りかけ性」が「とてもある」）と受け取られるテキストには、「語りかけ性」が「ある」と判断されたテキスト群に高頻度な特徴的表現が含まれる傾向があり、「話しことば的」表現などの効果も考えられる。しかし、書籍においては、「どちらかといえば」語りかける（「語りかけ性」が「どちらかといえばある」）と受け取られるテキストもあり、本稿で取り上げてきた（2）のような例は説明し難かった。これらは、表現群としてまとめると意識される特徴的な表現が重なった効果によって「語りかける」という文体が形成されているため、高頻度な表現が含まれない場合も、「語りかける」と受け取られ得るのである。

最後に、本稿では扱いきれなかった「語りかける」文体に関わると考えられる特徴的な表現に触れておく。

以下の(18)は、「語りかけ性」が「ある」と判断された例であるが、NDCは「2(歴史)」であり、「語りかけ性」が「ある」と判断されたテキストの多い分類ではない。また、タイトルから「語りかける」とは判定されず、Cコードでも実用書には該当しない例である。

(18) ほんらいなら、これから歴史そのものの話がはじまるわけである。しかし、そのまえに、もうひとつだけ伝えておきたいことがある。それは、この西域、あるいは中央アジアとよばれる土地が、ほんの最近まで、文字どおりのテラ・インコグニター知られざる土地—だったということに関係する。要するに探検の対象となるような未知の地域だったのだ。探検ということが、なにかを目的にして未知の土地に足をふみ入れることだとすれば、あとでその名が出てくるはずの張騫や班超、あるいは玄奘三蔵やマルコ・ポーロ、かれらも古代や中世の大探検家だったといえるだろう。しかし、科学的調査という点からみると、一般に探検時代は一八世紀にはじまり、いま問題の西域の地では、一九世紀、それも後半期になって注目されてくる。そして、それは、まず地理学的調査にはじまった。いまでこそ、大小の河川や山脈の明記された地図を、われわれは手にすることができるけれど、一〇〇年、いや五、六〇年前には空白地帯がいっぱい残されていた。カラコルムと崑崙の山系の区別も明確でなかったし、タクラマカン砂漠を流れる川筋の方向も、想像によるしかなかった。航空写真測量などという新技術のない時代のことだから、こつこつと地上を歩き、ひとつひとつ測量することからはじめねばならなかったのである。

(LBd2\_00055：山田信夫「世界の歴史」)

(18) においては、まず、書きことばとしてはくだけていると考えられる表現（下線を施した）「けれど」「いっぱい」などが散見される点が特徴的である。「話しことば」の典型イメージ（三宅 2005）を加味するようなくだけた表現は、「語りかける」という印象に関係する可能性が考えられる。また、それら以外の手がかりとして、ここまでに見た「わけである」「のだ」（今村（2007）、5.2 参照）「伝えておきたい」「できる」「われわれ」（5.2 参照）などの表現や、「とすれば～いえる」「みると～てくる」（5.3 参照）のような条件と回答の関係が見つかる。しかし、このほかに「ほんらい」「なにか」「あとで」のようにひらがな表記の割合が高いことなどの要因も考えられる。今後は表記の問題をも視野に入れた調査を進める必要がある。

本稿で見たような文体的特徴は、コーパスから高頻度語を抽出するのみでとらえられるものではない。文脈に依存した表現や、テキストに向かう読み手の前提的態度、テキストのジャンルや内容など、多角的な視点によって把握され得るものである。このような、人間が判断してアノテーションしなければ得られない情報をメタデータとしてコーパスに付与していくことにより、コーパスに依拠したテキスト分析・談話研究に対して有益な研究用情報を提供することができると考えられる。

## 参考文献

- 安藤宏（2012）『近代小説の表現機構』東京：岩波書店。
- 池上嘉彦（1986）「日本語の語りのテキストにおける時制の転換について」『記号学研究』6(25): 61-74。
- 今村和宏（2007）「『のだ』の発話態度の本質を探る：「語りかけ度」と「語りかけタイプ」」『一橋大学留学生センター紀要』10: 37-48。
- 柏野和佳子（2010）「『直接的な語り』という表現スタイルをもつ書籍テキストの人手抽出の試み」『ことば工学研究会』35: 63-72。
- 柏野和佳子（2013）「書籍サンプルの文体を分類する」『国語研プロジェクトレビュー』4(1): 43-53。
- 岸本千秋（2005）「ネット日記における読み手を意識した表現」三宅和子・佐藤彰・岡本能里子（編）、204-231。
- 小磯花絵・田中弥生・小木曾智信・近藤明日子（2011）「評定実験に基づくテキスト分類尺度の体系化の試み」『現代日本語書き言葉均衡コーパス』完成記念講演会予稿集』47-52。
- 松本和也（2008）「語りかけるテキスト—太宰治「カチカチ山」」『國文學：解釈と教材の研究』53(4): 23-31。東京：學燈社。
- 三宅和子（2005）「携帯メールの話しことばと書きことば」三宅和子・佐藤彰・岡本能里子（編）、234-261。
- 三宅和子・佐藤彰・岡本能里子（編）（2005）『メディアとことば2「特集」組み込まれるオーディエンス』東京：ひつじ書房。
- 野田春美（2012）「エッセイ末における読み手を意識した表現」『人文学部紀要』32: 39-54。
- Schiffirin, Deborah (1981) Tense variation in narrative. *Language* 57(1): 45-62.
- Silva-Corvalán, Carmen (1983) Tense and aspect in oral Spanish narrative: Context and meaning. *Language* 59(4): 760-780.
- 保田祥・柏野和佳子・立花幸子・丸山岳彦（2012a）「『語り性』を有する書きことばの典型例の分析」『第1回コーパス日本語学ワークショップ予稿集』139-146。
- 保田祥・柏野和佳子・立花幸子・丸山岳彦（2012b）「『語りかけ性』を有すると判断される書きことばの表現」『第2回コーパス日本語学ワークショップ予稿集』43-50。
- 保田祥・柏野和佳子・立花幸子・丸山岳彦（2013a）「書きことばにおける「語りかけ」は何のために用いられるのか」『第3回コーパス日本語学ワークショップ予稿集』143-152。
- 保田祥・立花幸子・柏野和佳子・丸山岳彦（2013b）「『ベテランが足を保護する』が語りかけるとき」『第4回コーパス日本語学ワークショップ予稿集』345-354。

**関連 Web サイト**

現代日本語書き言葉均衡コーパス (国立国語研究所) [http://www.ninjal.ac.jp/corpus\\_center/bccwj/](http://www.ninjal.ac.jp/corpus_center/bccwj/)

## Addressive Expressions in Written Japanese

KATO Sachi<sup>a</sup>      KASHINO Wakako<sup>b</sup>

TACHIBANA Sachiko<sup>c</sup>      MARUYAMA Takehiko<sup>b</sup>

<sup>a</sup>Postdoctoral Research Fellow, Center for Corpus Development, NINJAL

<sup>b</sup>Department of Corpus Studies / Center for Corpus Development, NINJAL

<sup>c</sup>Technical Assistant, Center for Corpus Development, NINJAL

### Abstract

In this paper, we report the characteristics of the writing style called “addressing the reader” found in books. For this study, we used the Library Sub-corpus contained in the Balanced Corpus of Contemporary Written Japanese (BCCWJ). In performing a stylistic analysis of the corpus, we used the information annotated in the corpus and the comments of the annotators (namely, readers), as well as the word frequencies and contextual connections between words. We found that the “addressive” style is more common in texts written in a casual style, such as essays and blogs, and contains phrases used to address readers directly or ask them questions (as if the writer and the readers are having a dialogue). In books, however, this writing style can be found in instructive texts, mainly in so-called how-to or enlightening books. Therefore, expressions that readers intuitively consider “addressive” in nature are not always encountered in these texts. In this paper, we will show that when readers find that the text is “addressing” them, they are influenced by context-dependent expressions and their own presuppositions as participants.

**Key words:** BCCWJ, style, addressive expression, annotation, document classification