コーパスに基づいた
語彙情報つき文法ハンドブックの構想

中　俣　尚　己

1. はじめに

筆者は現在、文法形式と実質語の共起関係に関心を持ち、コーパスを用いた研究を進めており、すでにいくつかの成果を発表している（中俣 2011a, 中俣 2011b）。この研究プロジェクトは、特定の文法項目、例えば「てある」の前にどのような実質語が多くくるのか、ということを調査するものである。この研究プロジェクトは日本語教育への応用が目的であるが、個々の形式について論文や学会発表の形で散発的に発表したとしても、日本語教育の現場で使えるリソースにはならない。

そこで、初級で扱う全ての文法形式について、網羅的にどのような実質語と共起するのかを網羅的に調べ上げ、一冊のハンドブックとしてまとめれば、現場でも活用しやすいのではないかと考えた。

すでにハンドブック作成に向けて、文法形式と語彙の調査を始めているが、それと平行して、ここに紙幅を借りてこのハンドブック作成の目的と動機、そしてどのような情報を載せるべきかという理念を記しておきたい。

以下、2. はハンドブックの趣意書にあたる内容であり、実質語の共起情報に注目する意義と、どのようにしてハンドブック制作の計画が生まれたかについて述べる。3. はハンドブックの設計図にあたる内容であり、どのような情報を載せるのかということについて述べる。4. はまとめである。また、付録として、掲載予定の項目のリストとハンドブックの実際のページの案を掲載する。

2. 語彙情報つき文法ハンドブックの趣意書―なぜ語彙情報か？―

ここでは、筆者が語彙情報つき文法ハンドブックという構想に至った経緯をいくつかに分けて述べる。まず、2. 1では筆者が日本語教員養成に関わった中で直面した「例文を作ることの困難さ」について述べる。2. 2では研究という側面からの問題意識を述べる。2. 3ではハンドブック制作の直接の起因ではないが、制作を強力に後押ししたコーパス『BCCWJ』と検索システム『中納言』の完成について述べる。2. 4では語彙情報つきハンドブックのメリットについてまとめる。

2. 1　日本語教員養成からの問題意識

筆者は2009年4月から2年間、日本語教員養成に携わった。日本語教員を目指す学生たちは模擬授業を行っていたが、その準備室に筆者は勤務し、多くの学生が筆者に質問に来るという毎日であった。その質問で最も多かったのは文法事項の確認であり、これについては庵・高梨・中西・山田（2000）などの文法ハンドブックを見ながら解説すれば事足りた。しかし、次に多い質問は「授業で使用すべき例文が思いつかない」というものであった。

言うまでもなく、授業において例文は重要なファクターである。文法項目の導入においては何よりも典型的で馴染みやすい例文が望ましい。また、その後のドリル練習や宿題の問題もすべて「例文」であるということができる。この質問に対しては文法ハンドブックでは例文が少なすぎて対応できない。一方で、教科書の例文をそのまま使うのでも意味がない。学習者の生活に根ざした、活き活きとした例文でなければならない。そのような例文を考えるのは、経験の浅い学生にとって難事であるということに筆者は気付かされた。

例えば、進行の「ている」を結果残存の「ている」と区別して例文を作るという課題でさえも、スムーズにいかないことが多かった。のみならず、初級の例文作りには様々な制約が存在する。例えば、「まだ習っていない語は扱わない」「動詞のグループは1グループ・2グループ・3グループすべて扱う」「1グループの動詞はカ行のもの、サ行のもの…ワ行のものすべて扱う」といった制約である。このような制約を全て満たして例文を作るのは学生にとっては至難であった。

良い例文を作る際の指針が欲しいという声は学生からも、筆者とともに学生の指導にあたっていたTAからも聞かれた。「良い例文」の条件は色々考えられるだろうが、その中の一つには「よく使うもの」が考えられるだろう。実際に、

検索エンジンを使う学生もいたが、その結果は必ずしも初級での導入に適切なものではないことも多いし、頻度が多いものを一覧するには不適である。そこで、コーパスを使って、ある文法項目と共起する述語にどんなものが多いかをまとめれば、後は身近な名詞を格のスロットに入れることで、効果的な例文が作れるのではないかと考えた。

良い例文を作ることは日本語教師に必要な資質の一つである。しかし、これまでそのテクニックは「教師の経験と勘」に委ねられてきたきらいがある。コーパスからの情報だけでその「経験と勘」をすべてカバーできるべくもないが、できるところから少しずつ可視化していってこそ、日本語教育の発展がある。

また、教科書における語彙の出現順序についても、これまでは科学的な根拠ではなく「経験と勘」によって決められてきたのではないだろうか。さらに、「まだ習っていない語は扱わない」「動詞のグループは1グループ・2グループ・3グループすべて扱う」「1グループの動詞はカ行のもの、サ行のもの…ワ行のものすべて扱う」は全く教える側の勝手な都合であり、学習者のニーズに沿っているとは言いがたい。文法項目にどのような語彙が共起するかを調べれば、個々の授業における例文だけでなく、新しい教科書を作る際の語彙配列の決定における情報も提供できる。

2. 2　研究からの問題意識

上述した日本語教員養成の仕事とほぼ同時期に、筆者は日本語教育文法の論文集（森・庵 2011）で「てある」と「ておく」の導入に関する考察を行った（中俣 2011b）。しかしながら、筆者はアスペクトの研究においては全く蓄積がなく、そのため、まずはコーパスから例文を集めるところから作業をスタートした。特に何か仮説があったわけではなく、まずは大量の例を集めたのである。

当時はその名称は知らなかったが、仮説を立てずにコーパスを調べる調査法は Corpus-Driven Approach と呼ばれ、仮説をコーパスを使って検証する Corpus-Based Approach とは区別される。そして、Tognini-Boneli（2001）は Corpus Driven Approach をとることを推奨しているのである。

期せずして筆者が行った Corpus Driven Approach の結果は驚異的であった。何と「書く」の 38%～52%が「書いてある」であったのである。これは、『日本語話し言葉コーパス』『名大会話コーパス』『現代日本語書き言葉均衡コーパス』（BCCWJ）のどのコーパスでも非常に高い数値であり、さらに BCCWJ のどのジャンルで切り取っても「書く」が圧倒的多数を占めていたのである。この数字は直接的には言語学の興味にはならないかもしれないが、言語教育では

無視できない情報である。何よりも「書いてある」というコロケーションが重要であるということであるし、更には「状態を表す時は「ている」を使うが、「書く」の時だけは「てある」を使う」という暫定的なルールさえ、一定の効果を持つといえるからである。

また、文法形式と語彙の共起関係が言語研究にとっても興味深い問いを投げることもある。条件表現「と」「ば」「たら」の前に来る語を調べると、従来の条件表現研究では非典型的とされてきた「この表を見ると～」や「簡単に言えば～」のような例が圧倒的多数を占め、「もし宝くじがあたったら、海外旅行に行きたい」のような典型的な仮定条件文は少数であった。仮定条件文が典型例とされてきたのは論理学の伝統からであろうが、こと日本語の実例を虚心に見つめる限り、それを典型例と呼ぶことには一考の余地がある。コーパスを使った研究は日本語における「条件」の定義にも一石を投じるのである。

2. 3　大規模コーパス「BCCWJ」と検索システム「中納言」の公開

ここでは、問題意識とは別に、ハンドブックの制作を強力に後押しした出来事について付言する。日本語教員養成と研究の両側面から「語彙情報つき文法ハンドブック」の制作を決意したわけであるが、当初の計画では、有志を集め、グループで制作を進めるつもりであった。しかし、2011 年 8 月 2 日に状況は一変する。

同日、国立国語研究所のコーパス整備計画「KOTONOHA 計画」の主要な位置を占める『日本語書き言葉均衡コーパス』（BCCWJ）が一般公開された[(1)]。そしてほどなく、形態素解析を利用した検索システム『中納言』も公開された[(2)]。『中納言』は形態素解析済みの情報を使っており、例えば「動詞＋かもしれない」といった組み合わせも検索し、一覧で表示できる。ここで、動詞の方をキーにして検索すると、その動詞の読み方や語彙素（いわゆる代表形）の情報も付与された表として出力される。ブラウザの画面に表示されるのは 500 件であるが、実際には全てのデータをダウンロードできる[(3)]。これをエクセルで読み込み、ピボットテーブルに変換、キーの項目数で並び替えを行えば、容易に文法形式に前接する動詞の上位リストを作成できるのである。比較的頻度の少ない項目で、検索がスムーズにいけば、ここまでの作業は 10 分とかからない。

筆者は公開後、早速申し込みを行い、『中納言』の威力を実感した。そして、仮に筆者一人が作業を行なっても、初級の文法項目を網羅する調査を行うことは十分に可能であると判断し、調査計画を立て、実行に移したのである。

とはいえ、100 の文法項目について、ハンドブック用に限られた情報を提

示するだけであれば筆者一人でも可能であるが、個々の項目の精査、言語学的な研究を行うことは、個人の力ではとうてい不可能である。『中納言』を利用した文法形式と実質語の共起関係の研究が今後広がっていくことを期待している。

2. 4　語彙情報つき文法ハンドブックの有効性

ここでは、語彙情報つき文法ハンドブックの必要性についてまとめる。筆者は、日本語教員養成に携わる中で例文作りの難しさに直面し、また、文法研究から語彙の重要性を感じ、さらにBCCWJの完成を受けて、語彙情報つき文法ハンドブックの制作に着手した。このハンドブックは、個々の文法形式について、前接する実質語のリストとそれを使った例文をあげ、さらに文末形式やよく共起する他の機能語についての情報も載せたものである。このハンドブックは以下のような有効性を備えている。

(1) 教師にとっては、毎回の授業で準備する例文作りに有効である。特に、日本語の内省がきかないノンネイティブ教師にとっては、教科書にとらわれず、母語話者がよく使う例文を提示することが容易になる。
(2) 学習者にとっては、日本語でよく使われるコロケーションを学ぶことができる。
(3) 教科書作成において、導入すべき語彙の選定に必要な情報を提供できる。
(4) 言語研究において、新たな角度からの知見を提供できる。

3. 語彙情報つき文法ハンドブックの設計図 ―何を載せるか？―

ここでは、語彙情報つき文法ハンドブックに載せるべき情報について吟味する。

多くの項目を搭載するハンドブックという形態上、一項目につきあまり多くの情報を載せることはできないが、例えば、以下のような情報は有益であろう。

(5) 文法項目の頻度
(6) 文法項目に前接する頻度の多い述語
(7) 文法項目に前節する頻度の多い述語を使った例文
(8) 文法項目に前節する頻度の多い述語の格配列
(9) 文法項目の出現位置（主節か、従属節か）
(10) 文法項目と共起する（共起しない）テンス・アスペクト・モダリティー標識やその他機能語

(5) から (7) は全ての項目について掲載する情報であり、(8) から (10) は項目によって掲載する情報である。

(6) に述語とあるように、基本的には述語と結びつく文法項目のみを扱う。格助詞「の」は名詞句と名詞句を結ぶ情報であるが、名詞句は傾向が見えにくいこと、また学習者にとって身近な名詞句を使えばよいことから、名詞句と結びつく文法項目はハンドブックには収録しない [(4)]。付録 1 に、掲載予定の項目を列挙する [(5)]。

以下、3. 1 から 3. 6 では (7) から (10) の各項目について具体例を挙げながら述べる。

3. 1　文法項目の頻度

基本となる情報である。初級文法項目の頻度調査にはすでに森 (2011) があるが、本ハンドブックにも載せるべき情報であろう。1 億語における頻度の絶対的な数値を載せることで、それぞれの語の重要度を相対的に比較することができるようになる。例えば、「ている」は 817,689、「てある」は 14,603、「てしまう」は 82,575 である。

なお、おそらく「ている」が最も頻度の高い項目となる。これより頻度の高い形式として「た」があるが、「ていた」「ほうがよかった」など他形式との複合で論じるべき場合も多いことから対象から外す。

3. 2　文法項目に前接する頻度の多い述語

本ハンドブックの特徴となる部分である。例として、BCCWJ における「てある」に前接する動詞のベスト 10 を表 1 に示す。左の数字は実際の頻度、() 内は「てある」全体に対する割合である。

表 1 「てある」に前接する動詞ベスト 10

1	書く	4710 (31.8%)	6	掛ける	185 (1.3%)
2	置く	1507 (10.3%)	7	入れる	176 (1.2%)
3	する	728 (5.0%)	8	取る	156 (1.1%)
4	張る	407 (2.8%)	9	止める	150 (1.0%)
5	飾る	217 (1.5%)	10	作る	145 (1.0%)

実用と紙幅を勘案すれば、ベスト 10 を収録するのが妥当であろうか。形容詞と名詞については 5 ずつでも構わないであろう。

上述した「てある」については語彙が非常に偏っているため、実数の頻度でも問題はないが、他の文法形式の場合、単純に出現頻度の多い動詞に左右されることがある。その影響を排除するために石川（2006）ではTスコア、MIスコア、対数尤度比などの統計的指標が紹介されている。

表2は「ている」に前接する動詞のベスト10に対して、「共起率」（その動詞が「ている」と共起する割合）「Tスコア」「MIスコア」「対数尤度比」を示したものである。

表2 「ている」に前接する動詞スコアベスト10

順位	語彙	前接頻度	語彙頻度	共起率	Tスコア	MIスコア	対数尤度比
1	する	45908	625265	7.34%	190.40	3.17	124782
2	成る (6)	30905	475332	6.50%	153.69	2.99	76612
3	思う	17796	218327	8.15%	120.02	3.32	51353
4	持つ	13946	82142	16.98%	112.41	4.38	57675
5	言う	9588	420301	2.28%	62.82	1.48	7447
6	考える	9197	87387	10.52%	88.45	3.69	30979
7	見る	9093	152574	5.96%	82.27	2.87	20922
8	知る	8356	61557	13.57%	85.90	4.05	33448
9	入る	6601	60516	10.91%	75.16	3.74	22680
10	遣る (5)	6134	74262	8.26%	70.57	3.34	17782

このようにして見ると、「言う」のように共起率がわずか2%しかなくとも、頻度が莫大な語は上位にランクインすることがわかる。田野村（2010）の調査によれば「ている」との共起率が高い語はほかにも多くあり、最も高い「似る」は97%に達する。しかし、共起率の高い語が必ずしも多く出現するわけではなく、「持ち合わせる」「秘める」などは初級で導入すべき語とはいえない。

Tスコアは「言う」のように共起率が極端に低いものでは若干数値が下がるが、おおむね共起頻度の実測値と相関する。

また、MIスコアは共起率と高い相関を示し、一般には3以上で有意な結びつきとされる石川（2006）。そうすると、1位の動詞「する」は「ている」と一定の結びつきやすさを持つということになる。筆者の直感では頷けないこともないのだが、MIスコアはそもそも出現頻度の少ない語から有意な組み合わせを見つけるのに適した指標であり、出現数が1万を超えるような例に無制限に適用できるかは不明である。

石川（2006）によれば、対数尤度比が最もバランスのとれた指標であるという。試しに対数尤度比でベスト10をピックアップすると、上記の表中の語のうち、8語までが重なる。網掛けを施した「言う」と「遣る」がランクから外れ、代わりに「残る」と「付く」がランクインする。これらはそれぞれ共起率が26.79％、12.28％と高い語である。対数尤度比には「言う」のような「不当にランクインしている語」を排除する働きがあるといえよう。

しかしながら、ハンドブックに採用する数値としては、対数尤度比は一考の余地がある。頻度であれば5000は500の10倍であるが、対数尤度比は極めて抽象的な値であり、直感的な把握は難しい。また、「言っている」が多く使われているということも厳然たる事実であり、ハンドブックには頻度ベースでの表を載せる。ただし、MIスコアや対数尤度比を用いた分析もまた必要であることは言うまでもない。

3. 3　文法項目に前節する頻度の多い述語を使った例文

コーパスは有益な統計的情報をもたらしてくれるが、コーパスの用例自体は、日本語教育でそのまま使うには不適である。現場で活用するためには質の高い例文は必須である。ただし、ここでもコーパスで動詞＋文法項目の形で検索を行い、よく使われるパタンの文を元に、作例を行うことが重要である。以下に、「てある」の上位3語の例文を挙げる。

(11)このクーポンには、「11月末まで」と**書いてある**。

(12)私は机の上に**置いてあった**本を手に取った。

(13)この部屋は冷房を弱く**してあります**。

しかしながら、一人の人間が例文を作っていくと、どうしても偏りが生まれることは否めない。この項目は実践女子大学の学生達にも手伝っていただき、質の高い例文を掲載することを目指したい。

3. 4　文法項目に前節する頻度の多い述語の格配列

例文に使用する述語はランキングに従い、名詞は身近なものを使用するとして、どのような格配列が多いかというのは直感ではなかなかわからない情報である。

試みに「書いてある」の格配列を表3に示す。

表3 「書いてある」の格配列

～に～が書いてある	314	～が～に書いてある	134
～が～と書いてある	221	～と～が書いてある	129
～に～と書いてある	542	～と～に書いてある	160

これは「書いてある」の1つ前の語の格助詞と、さらにそこから3語以内に出現する格助詞を指定して検索した結果である。「書いてある」の直前に格助詞が来るとは限らないので、正確な数値とは言えない。しかしながら、相対的にどの配列が多いのかを知る目安にはなろう。「～に～と書いてある」が最も多いことがわかる。3. 3における（11）の例文はこの情報を元に作成した。

しかしながら、このような各配列は動詞を指定して一つ一つ調べると、多くの情報を載せることはできないという問題がある。

3. 5 文法項目の出現位置（主節か、従属節か）

その文法項目が言い切りの位置で終わるのか、それとも従属節で終わるのかといった情報である。「てある」の場合は「言い切り」が多いが、「てあった」になると「置いてあった本」のように連体用法が多くなる。

3. 6 文法項目と共起する（共起しない）テンス・アスペクト・モダリティー標識やその他機能語

例えば、「てある」「ておく」はほとんど疑問文では使われないことがわかっている（中俣 2011b）。このような情報は必要であろう。また、文法項目によっては丁寧体と共起するか普通体と共起するかということも問題になりうる。さらには、「てある」のように、「字が書いてある」（状態の描写）と「切符は買ってある」（準備完了）の2つの意味をもつ形式もある。どちらの意味が多く使われるのか、は重要な情報であろう。このような多義に関する問題は、機械による自動判定はできない。500例ないし1000例をランダムサンプリングし、人の目で判定するほかない。

4. おわりに

この論文では、語彙情報つき文法ハンドブックの制作に至った理念と、そこに必要となる情報について記した。現在、「中納言」を使ってデータの収集を進めており、まずは初級で扱う約100項目を収録することを目標にしている。

この論文が掲載される『実践国文学』81 号が刊行されるころには、調査はあらかた終了している予定である。

また、今回は端緒として初級項目を選定したが、語彙との結びつきがより重要なのは「～ざるをえない」「～つつある」のような中・上級項目であるといえよう。語彙との結びつきを見ることで、未だ混迷の感がある中上級シラバスにも新たな道筋を見出すことができるかもしれない。今後なすべきことは山積している。

注

(1) 『現代日本語書き言葉均衡コーパス』の設計に関しては前川（2011）、山崎（2011）を参照。
(2) 『中納言』の設計に関しては小木曽ほか（2011）を参照。
(3) ただし、2011 年 11 月現在ではダウンロードは 10 万件という上限がある。また、共起条件を増やすと、一度にダウンロードできる件数はさらに少なくなる。
(4) 山内（2012 予定）は実質語の難易度について議論を行っている。学習者の発話における単語の出現順は、具体物については単語親密度が関係しているとしている。すなわち、身近なものほど覚えやすいということである。
(5) 選定基準は庵・高梨・中西・山田（2000）で項目として取り上げている文法項目のうち、述語に接続する項目である。ただし、3. 1 で述べるように、除外する項目もある。
(6) 「成る」「遣る」の表記は BCCWJ に使われている UniDic という辞書の表記による。UniDic には「語彙素」「語形」「書字形」「発音形」という 4 つの表記情報が搭載されているが、異なる文字で書かれる同じことば（「大きい」と「おおきい」など）をまとめてカウントすることが本プロジェクトでは必要となるため、「語彙素」を使用している。この語彙素の表記法が「成る」「遣る」であり、実際のコーパスの文例では「なる」「やる」となっているものが大半である。

参考文献

庵功雄・高梨信乃・中西久実子・山田敏弘（2000）『初級を教える人のための日本語文法ハンドブック』スリーエーネットワーク

庵功雄・森篤嗣（編）（2011）『日本語教育文法のための多様なアプローチ』ひつじ書房

石川慎一郎（2006）「言語コーパスからのコロケーション検出の手法―基礎的統計値について―」『統計数理研究所共同研究レポート』190 pp.1-28 統計数理研究所

小木曽智信・中村壮範・鈴木泰山・八木豊・山崎誠・前川喜久雄（2011）「コーパス

検索システム「中納言」デモンストレーション」『『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』pp.43-46　文部科学省研究費特定領域研究「日本語コーパス」総括班
田野村忠温（2010）「日本語コーパスとコロケーション ――辞書記述への応用の可能性――」『言語研究』138　pp.1-23　日本言語学会
中俣尚己（2011a）「コーパス・ドライブン・アプローチによる日本語教育文法研究―「てある」と「ておく」を例として―」『日本語教育文法のための多様なアプローチ』pp.215-233 ひつじ書房
中俣尚己（2011b）「条件を表す「なら」に前接する語彙―例文作りのために―」『2011年度日本語教育学会秋季大会予稿集』日本語教育学会
前川喜久雄（2011）「特定領域研究「日本語コーパス」と『現代日本語書き言葉均衡コーパス』」『『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』pp.1-10　文部科学省研究費特定領域研究「日本語コーパス」総括班
森篤嗣（2011）「『現代日本語書き言葉均衡コーパス』コアデータにおける初級文法項目の出現頻度」『日本語教育文法のための多様なアプローチ』pp.57-78 ひつじ書房
山内博之（2012 予定）「非母語話者の日本語コミュニケーション能力」野田尚史（編）『日本語教育のためのコミュニケーション研究』くろしお出版
山崎誠（2011）「『現代日本語書き言葉均衡コーパス』の構築と活用」『『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』pp.11-20　文部科学省研究費特定領域研究「日本語コーパス」総括班
Tognini-Boneli, E. (2001) Corpus Linguistics at Work. Amsterdam:John Benajamins.

謝辞

この論文では国立国語研究所作成の『現代日本語書き言葉均衡コーパス』ならびに検索システム『中納言』を利用して、データを収集しました。また、対数尤度比の計算には神戸大学の石川慎一郎氏がウェブサイトで公開しているエクセルシート（http://language.sakura.ne.jp/s/stat.html）を使用しました。それぞれの作成者の皆様に感謝申し上げます。

また、語彙情報つき文法ハンドブックの構成に関しては、実践女子大学の山内博之教授に様々な助言をいただきました。感謝申し上げます。

付録１　語彙情報つき文法ハンドブックに収録予定の語

庵・高梨・中西・山田（2000）より述語に接続する形式を抜き出した。全92項目。

～た、～ていない、～てしまう、～たことがある・～たことがない、～ことがある、～とき、～ている、～ているところだ、～つつある、～続ける、～始める、～だす、～終わる・終える、～やむ、～ところだ、～てある、～ておく、～てみる、～ようになる、～ようにする、～ことになる、～ことにする、可能、～ことができる、～てあげる、～てくれる、～てもらう、～ていく、～てくる、～だろう、～かもしれない、～はずだ、～にちがいない、～そうだ（様態）、～ようだ、～みたいだ、～らしい、～そうだ（伝聞）、～よう、～つもりだ、～たい、～てほしい、～なさい、命令形、～な、～てください、～ましょう、～ましょうか、～ませんか、～ないか、～ようか、～なければいけない・～なければならない、～ほうがいい、～てもいい・～なくてもいい、～ね、～よ、～よね、～て、～ないで・～なくて・～ずに・～ず・～なく、～ながら、～たまま、～たり、～し、～てから、～あとで、～あと、～まえ、～まえに、～までに、～まで、～あいだに、～あいだ、～うちに～、から、～ので、～ために、～のに・～に（目的）、～と、～ば、～たら、～なら、～ても、～のに、～けれども、～が、～のだ、～か、受け身、使役、～やすい、～にくい、～すぎる、尊敬語、謙譲語

付録２　語彙情報つき文法ハンドブックの見本

てある　出現数 14603

アスペクト形式

接続：他動詞 のテ形

① 他動詞に接続し、状態を表す。部屋などの情景を表す時に使われることが多い。

② 他動詞に接続詞、準備ができていることを表す。

前接する語彙

	動詞	形容詞	名詞
1	書く　4710（31.8%）		
2	置く　1507（10.3%）		
3	する　728（5.0%）		
4	張る　407（2.8%）		
5	飾る　217（1.5%）		
6	掛ける　185（1.3%）		
7	入れる　176（1.2%）		
8	取る　156（1.1%）		
9	止める　150（1.0%）		
10	作る　145（1.0%）		

※動詞 10 のカバー率 57%

「掻く」が多く検出されたが、「書く」であると判断されたため，統合した。また、「描く」も含めた。

格助詞のパターン

「書いてある」の場合、直前に来る格は「が」695 件　（「を」48 件）、「と」1379 件、「に」756 件である。組み合わせは下表の通り。

～に～が書いてある	314	～が～に書いてある	134
～が～と書いてある	221	～と～が書いてある	129
～に～と書いてある	542	～と～に書いてある	160

仮に「書いてある」に重きを置いて指導する場合は、「～に～と書いてある」

も多く使われている。引用の「と」をすでに指導しているのであれば、指導しても構わないだろう。

後ろの要素

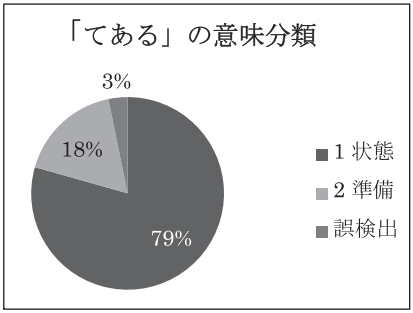
てある。／てあります。（言い切り）	2106/1069 (21.7%)	てあった。／てありました。（言い切り）	1015/409 (9.8%)
てあって	341 (2.3%)	てあり	485 (3.2%)
てある＋名詞	2864 (19.3%)	てあった＋名詞	535 (3.7%)
てある＋終助詞	212 (1.5%)	てある＋接続助詞	565 (3.9%)

「てある」は言い切りの形で使われることが多く、これ単独で指導するのがよいだろう。また、連体修飾節も多いです。また、「てあった＋名詞」の場合は「置いてあった本」のように「置く」の割合が高くなる。

また、疑問文はほとんど出現しない。

用法

1000 例のみを抽出して意味分類を行った結果、以下のグラフのようになった。



最も多い「書いてある」が 1. 状態の意味であるため、当然の結果である。あえて複数の用法を教えなくとも、状態を表す「てある」が典型的と言えることがわかる。すなわち、自動詞の状態を「ている」を使って表すのに対して、他動詞の状態は「てある」を使って表すのである。

なお、2. 準備で多い動詞は「してある」「言ってある」「隠してある」などである。

（なかまた　なおき・実践女子大学助教）