国立国語研究所学術情報リポジトリ

ウェブコーパス「梵天」による敬語研究：その活用可能性に関する事例的検討

メタデータ	言語: Japanese 出版者: 公開日: 2018-03-20 キーワード (Ja): キーワード (En): Balanced Corpus of Contemporary Written Japanese (BCCWJ), Corpus of Spontaneous Japanese (CSJ), NINJAL Web Japanese Corpus (NWJC) 作成者: 金, 賢眞, KIM, Hyunjin メールアドレス: 所属:
URL	https://doi.org/10.15084/00001502

ウェブコーパス「梵天」による敬語研究
—その活用可能性に関する事例的検討—

金 賢眞（大阪大学大学院文学研究科）

Reaearch on Honorific using Web Corpus `BonTen'
Case study on its utilization potential

Hyunjin Kim (Graduate School of Letters, Osaka University)

要旨

ウェブコーパスによる敬語研究の可能性について事例的に検討する。具体的には、謙譲表現「ご～する」が尊敬・丁寧表現として誤用される現象に注目し、ウェブコーパス「梵天」がこの問題の検討に有益なデータとなり得るか否かを、現代日本語書き言葉均衡コーパス及び日本語話し言葉コーパスとの比較により評価する。この三種のコーパスについて、「ご～する」の「～」の部分に使用される動詞の異なり（種類）と、それぞれの誤用率を調査すると、「梵天」では多様な動詞が収集でき、他のコーパスの結果をほぼ全てカバーすること、また、他のコーパスより誤用率の大きくなる動詞の多いことが確認できた。このことは、ウェブコーパスが「ご～する」の正用・誤用を含む使用実態の調査に適切な資料であることを示すとともに、とりわけ誤用例の分析を充実させ得るという点で、コーパスを用いた敬語研究を更に発展させるものであることをも示唆している。

１．はじめに

日本語の敬語の分類については諸説あるが、学校教育でも用いられる伝統的かつ一般的な理解としては、「尊敬語」「謙譲語」「丁寧語」という三分類が挙げられる。この三分類はそれぞれ明白に区別される性格を持つものであり、従って、それぞれ異なる敬語形式によって表現される。「謙譲語」は「向かう先」に対する敬語か、それとも「相手」に対する敬語かによって更に「謙譲語Ⅰ」と「謙譲語Ⅱ」に細分化されることがあるが、本発表で扱う「ご～する」という表現形式は、所謂「謙譲語Ⅰ」の一般的な形式の一つである。この「謙譲語Ⅰ」とは「自分側から相手又は第三者に向かう行為・ものごとなどについてその向かう先の人物を立てて述べるもの」（文化審議会 2007）である。つまり、「ご～する」の主語は必ず自分側で、その行為の向かう先は相手側、又は立てるに相応しい第三者でなければならない。しかし、最近はこの「ご～する」を「お客様がご利用しやすい鉄道づくり」「どうぞご安心してご連絡ください」のように自分ではなく相手側の行動について直接用いる使用例が目立つ。このような誤用[1]は日常的に見られ、それなりに定着してきているように思われる。

本発表では、上述したような敬語の所謂「誤用」を含む使用実態に関する研究に用いる資料としてのウェブコーパスの有用性について事例的に検討する。具体的には、謙譲表現

[1] 本研究において「誤用」という表現は、一般に謙譲語とされる「ご～する」の形式が自分側ではなく、相手側の行為に直接用いられた謙譲ではない使用のことを便宜的に指すものであり、その使用が正しい、あるいは間違っているといった価値判断を含むものではない。

「ご～する」が尊敬・丁寧表現として誤用される現象に注目し、国語研日本語ウェブコーパス NWJC がこの問題の検討に有益な資料となり得るか否かについて、現代日本語書き言葉均衡コーパス及び日本語話し言葉コーパスとの比較により評価していきたい。

２．先行研究

謙譲表現「ご～する」に関しては、今までその表現の成立の背景、使用条件、使用など、様々な観点において研究されてきた。中でも「ご～する」の誤用の原因に関しては、「ご～する」が成立当時から謙譲としてのみ用いられていたわけではなかったという歴史的な事実を指摘した小松（1967）、「ご～する」の「～」には名詞が入ることが多いため、「ご～する」が一つの塊として認識されるのではなく、「ご～」＋「する」のように、区別される二つの概念で捉えられやすいと形態論的観点から指摘した蒲谷（1992）、「ご～する」が一般的な尊敬表現形式である「ご～になる」より一拍短いという形態論的観点と敬語の大衆化の問題について指摘した菊地（1994）などがある。

更に、 井上（1999）は「ご～する」に形の上で謙譲を表す要素がないこと、人称が紛れやすいこと、そして、「ご～する」の使える動詞と文脈が限られていることが「ご～する」の誤用の原因であると指摘している。井上（1999）の指摘通り、謙譲表現「ご～する」は制限の多い表現であり、「～」に入ることができる動詞の条件に関しては、北條（1978）、菊地（1994）、川上（1996）などに詳しく論じられている。

また、蒲谷（1992）は、「ご～させて頂く」が一般に使われる表現であるのに対し、「お～させて頂く」が成立しにくいことから、「お・ご～する」絡みの誤用も一律ではなく、「ご～する」と「お～する」では違いがあることを指摘している。蒲谷（1992）の指摘は「ご～する」と「お～する」の誤用の差に関しての指摘であるが、「ご～する」の内部でも同じことが言える。本来謙譲表現である「ご～する」を謙譲以外に用いることが単なる母語話者の敬語能力の低下を原因とする誤用の問題に過ぎないものであれば、「ご～する」の誤用に関しては全ての動詞に同様の現象が見られてもおかしくないはずである。しかし、この誤用の成立しやすさには動詞による差が認められるため、「ご～する」の誤用に関しても正用同様、ある種の制限が働いているように思われる。そこで、発表者は「ご～する」の誤用は単なる誤用ではなく、敬語の用法シフトの過程を見せる一つの現象であるという立場から、どのような特徴を持つ動詞において所謂誤用が生じやすく、どのような過程で用法シフトが進んでいるかを分析している。以下本発表では、そのための資料としてのウェブコーパス NWJC の有用性について論じていきたい。

３．敬語の誤用の調査の資料として必要な条件

本研究のように、誤用を含む使用の全体像を調査する際に必要とするデータは、正用のみならず、自然な誤用も現れるもので、かつ、誤用率が安定し、一般化できるほどの用例数が必要となる。そして、使用される動詞の種類の数も多い方が望ましい。数が少なくとも分析はできるが、それでもグループ化して傾向を導き出すほどの数は必要となる。そして、話者に偏りのない、自然な用例が必要とされ、誰がいつ調査しても同様の結果が得られるような、安定したデータでなければならない。

このような研究には、従来使われてきたデータは十分ではない。従来の敬語研究は、人間を直接対象とするものが主流であった。その典型がアンケート調査や対面調査などである。従来用いられてきた方法は主に次の 4 つである。

第一に、対面調査である。国立国語研究所（2016）と井上（2017）は国立国語研究所で行っている「岡崎敬語調査」という経年調査の結果をまとめたものである。「『岡崎敬語調査』は敬語行動や敬語に関する知識・意見・内省などを聞く面接調査、調査協力者や職業といった社会生活調査（留置式）を中心とし、自然な談話調査を収集する『対話の実験的調査』等、数多くの調査群から成る」（国立国語研究所 2016）ものであり、昭和 28 年から経年的に行われているものである。

このような対面調査は、母語話者の生の言語使用を観察できる点、調査の途中でも対象者の答えによって柔軟に質問を追加できる点、そして、国立国語研究所（2016）や井上（2017)のように、経年的な調査ができるという点において重要なデータを提供するものである。しかし、他の調査方法に比べ莫大な費用がかかり、調査協力者の確保が難しいため、個人レベルでの大規模調査は難しい。

次に、アンケート調査は敬語研究において最も多く見られる調査方法の一つであり、伊藤（2009）はアンケート調査によって「ご～する」の誤用を分析している。アンケート調査は母語話者の内省を問うものであり、言語使用意識の効果的な調査方法である。これは、前述した対面調査に比べ、比較的短時間に大勢の人からの回答を収集でき、対面調査よりは個人でもある程度規模を拡大して調査することができる。しかし、これも大規模で行うには研究目的に沿った調査協力者を探すことは簡単ではなく、費用と時間が要求される。そして、専門会社に依頼しない限り、個人では特定地域内での調査しかできず、結果に地域性と方言の影響が生じてしまいがちである。また、一番の問題は、母語話者の言語使用意識は必ずしも実態とは一致しないということである。この意識と実態のずれに関しては、文化庁で毎年行っている「国語に関する世論調査」の結果からも確認できる。

そして、特に資料を用いない、研究者個人の内省による研究もある。これは、最も費用も時間も掛らない調査方法である。しかし、他の資料を用いない内省だけの調査は客観的な根拠に乏しく、また、非母語話者には向かない方法である。

最後に、既存の書籍を使う方法がある。その資料としては例えば、現代日本語研究会編（2011）の『合本　女性のことば・男性のことば（職場編）』などがあり、ほかにも、金（2009）のように、「手紙文」「スピーチ」の文例集を資料として使っているものがある。古い敬語の研究の資料としては古典もよく使われる。これは、資料とする本によって、長所と短所が異なる。前述した現代日本語研究会編（2011）は職場における自然談話のテキストデータで母語話者の実際の使用をそのまま資料として活用できる反面、データ量に限りがある。一方、「文例集」は規範に基づく用例が収集できるという点においては長けているが、実際に用いられている規範に沿わない表現は収集されないため、実態の調査には向かない。最後に、小説などの資料は、古典のようなものの場合は、他のデータのない、現代には使われない敬語を調査するときには大事な手掛かりとなる。しかし、一般的な小説の場合、言語に対する規範意識が一般人より高いと思われる「作家」という職業の人によって作られた作家の頭の中の「正しい敬語」、もしくは、作中の登場人物が使いそうな言い方でしかないため、自然な使用ではなく、現代語の敬語の研究にはあまり適切な資料ではないように思われる。

これらの点を踏まえて考えると、理想的な調査資料は、従来の資料の問題点を解決できる資料、つまり、個人でも大量のデータが収集でき、費用が少なく、母語話者の内省と実際の使用が同時に現れているデータである。そして、その条件を満たすものとしてはコーパスが考えられる。しかし、コーパスでも小説などの書き言葉を主とするものよりは、自

然な「敬語」の使用が見られるものが望ましい。その点において、ウェブコーパスは理想的な資料のように思われる。本発表では実例を持ってこれについて検討したい。

４．調査概要

４．１　調査資料

本研究では現代日本語書き言葉均衡コーパス（以下、BCCWJ）と日本語話し言葉コーパス（以下、CSJ）、国語研日本語ウェブコーパス（以下、NWJC）を調査資料とする。

４．２　調査対象

本来「ご～する」など、「ご」の付く敬語形式は、一般に「お＋和語動詞語幹＋する」のように、「お」の付く形式も含まれる。しかし、本研究では、「お～する」は対象外とし、「ご＋漢語サ変動詞語幹＋する」のみを研究の対象とする。なぜなら、「お＋和語動詞語幹」については「お使い」のように、「お＋和語動詞語幹」の組み合わせで一語として定着しているものや、「お掃除する」のように、美化と謙譲の区別が曖昧になっているものも含まれるため、それらを区別するのは難しいためである。

具体的には、BCCWJ と CSJ、NWJC において「ご～する」形式をとっている漢語サ変動詞を調査対象とするが、NWJC に関しては、対象語の数が多すぎるため、現段階では NWJC において用いられている全ての漢語サ変動詞に関する誤用と正用の分類はまだ完了していない。そこで、本発表の 6 節においては、資料としての有用性をコーパス相互で比較するという観点から、その調査対象を BCCWJ と CSJ において実際の使用例が観察された漢語サ変動詞に限る。また、以下本発表においては「漢語サ変動詞」の「語幹」のみを提示しているものや「ご＋漢語サ変動詞語幹＋する」の形式で提示しているものが混在しているが、これらは便宜上まとめて「動詞」と称する。

４．３　調査方法

まず、各コーパスの「ご～する」をとる動詞の語彙表を作成する。この語彙表は、基本短単位検索により「御」＋「名詞」＋「為る」を検索し、その結果をダウンロードした上で、名詞の部分をピボットテーブルにより分類する形式で作成した。具体的な検索語は下記の通りである。

短単位検索
前方共起（キーから―1 語）：語彙素＝「御」
キー：品詞＝名詞―普通名詞―一般
後方共起（キーから 1 語）：語彙素＝「為る」

まず、各コーパスの「ご～する」をとる動詞の語彙表を作成する。検索語から確認できるように、語彙素の「御」を前方共起と設定しているため、加工前のデータには「ご」のところに「お」や「み」、「する」のところに「される」など、本研究の目的に沿わないものも多数含まれていた。それらの対象外のものは全て削除し、分類後、ピボットテーブルを用いて、語彙表を作成している。この語彙表を用いて、実際使用されている動詞と頻度数を対象に分析を進める。

次に、それぞれのコーパスに BCCWJ と CSJ の語彙表の各動詞を入れた「ご～する」の形式を検索し、実際の用例をダウンロードする。BCCWJ は「現代日本語書き言葉均衡コ

ーパス」という名前からもわかるように、国会会議録、出版物、ブログなど、多様なレジスターからのデータが混在している。本研究では、コーパス別に扱っているデータによる特徴をより明確に比較すべく、BCCWJ のデータに関しては、「国会会議録（以下、HK）」「知恵袋・ブログ（以下、WK）」「その他（以下、ET）」と分類している。そして、それぞれの分類ごとに用例を収集した。

そして、それぞれの用例を確認し、それが誤用であるか正用であるかを判断する。その判断を基に、以下の式により誤用率[2]を計算する。

$$\text{誤用率} = \frac{\textit{誤用の用例数}}{\textit{全体用例数}} \times 100$$

また、NWJC に関しては、その用例数が何十万以上となる動詞も多く、全ての動詞を分析することは非効率的である。そのため、NWJC に関しては、結果の中でサンプルを求め、そのサンプルによる誤用率を求めていきたい。以下のような式により母比率のサンプルのサイズを求めている。

$$\mathrm{n} \geq \left(z_{\frac{a}{2}}\frac{\sigma}{\delta}\right)^2 = \left(\frac{z_{\frac{a}{2}}}{\delta}\right)^2 P(1-P)$$

ここで、n=サンプルサイズ、z_x=標準正規分布の上側 100x%点、σ＝母比率の標準偏差、δ=誤差、 p=母比率である。本研究の場合、母比率はわからないため、正確度が最も高くなる 50%と仮定し、上記式により信頼度 95%、許容誤差 5%に設定して計算した結果、必要となるサンプルサイズは 385 件である。本発表ではそれより少し高く設定し、最大 400 件までを調査対象としている。データの中でのサンプルの求め方としては、Excel のランダム機能を用い、ランダムに混ぜた用例の中で上から 400 件ずつ取り、分析した。本研究において、このように 400 件ずつとって分析した NWJC の誤用率は、基本、誤差±5%で、95%信頼できる結果となる。一方、今回は扱っていないが、中には反対に NWJC においても用例数が少なく、400 件に満たないものもある。それらに関しては得られるデータ全てを分析しているが、そもそも NWJC のデータにはウェブ上のデータ全てが入っているわけではなく、NWJC そのものがサンプリングされたデータであるため、この場合にも全数調査ではなく、サンプルによる調査になってしまい、誤差は更に増える。また、後に見るが、NWJC 以外のデータに関しては、BCCWJ の「ご紹介する」以外は用例数 400 を超えるものはなかった。しかし、この BCCWJ の「ご紹介する」に関しても、HK、WK、ET にわけると 400 件を超えているのは 446 件の ET のみであったため、これはサンプリングはせずに、全てのデータをそのまま対象としている。

最後に、それぞれのコーパス別、動詞別の誤用率の結果をまとめ、比較する。

---

[2] 後にみるが、コーパス別の用例数を見ると、その数が何百のものもあれば、10 に満たないものもある。例えば、用例数が 1 つしかない「a」という動詞がたまたま誤用で表れ、それを用例数 100 で誤用が 90 件あった「b」という動詞と同線上で比べ、「『a』のほうが『b』より誤用されやすい」というふうに判断していいのかなどの問題も問われるが、この発表はあくまでもコーパス別の特徴をみるためのものであり、それぞれの用例数も提示しているため、誤用率はあくまでも参考として示すものである。

５．コーパス別使用動詞の種類

５．１　BCCWJ

BCCWJ における「ご～する」形式を取る動詞の数は異なり 125 語である。その詳細を頻度数別に分類したい。頻度数 10 以上のものに関しては、頻度数がまとまらず分散しているため、頻度 10 以上というグループにして、括弧の中にそれぞれの頻度数を示す。

頻度 1：ひいき、鑑賞、激励、結婚、契約、供養、帰宅、祈念、記入、納得、丹精、同乗、同情、動座、来訪、来院、留意、満足、勉強、命令、没落、返答、発言、配慮、負担、批判、賞味、想像、賞讃、説得、守衛、守護、承知、乗車、試食、信頼、神幸、愛読、養育、研究、要求、容赦、誘導、融資、融通、移祭、一任、自慢、自愛、調達、卒業、注文、準備、尽力、賛成、参考、請求、推進、推薦、出産、出席、忠告、褒美、表記、下命、協議、回答、回向

頻度 2：期待、祈禱、来場、返金、予約、応募、伝授、静養、提示、指名、支援、進講、参加

頻度 3：購入、祈祷、返事、奉仕、使用、辞退、注意、創設

頻度 4：無礼、安心、提出、判断

頻度 5：対面、送付、要望、議論、認識、指導

頻度 6：苦労、奉公

頻度 7：心配、指摘、確認

頻度 9：検討、提供

頻度 10 以上：訪問(10)、遠慮(12)、利用(13)、答弁(14)、提案(14)、理解(16)、協力(18)、招待(27)、披露(28)、質問(34)、用意(37)、挨拶(44)、相談(58)、報告(68)、連絡(70)、一緒(125)、説明(166)、案内(169)、紹介(769)

上記の分類から確認できるように、BCCWJ における「ご～する」の形式をとる動詞は異なり 125 語であったが、その半分以上である 68 語の動詞が 1 回しか用いられていないものであった。また、頻度の低いものは殆どが、本来謙譲としての用法を持たないものであるように思われる。これらは先行研究において「ご～する」の形式を取ることができないとされている動詞が殆どである。

また、上記動詞の中には「ご一緒する」が含まれているが、これは「一緒する」という形式で用いられることはなく、「ご一緒する」でひとまとまりの表現である。そのため、「一緒」も本調査においては除外する。

５．２　CSJ

CSJ における「ご～する」形式を取る動詞の数は異なり 16 語である。その詳細を頻度数別に分類すると以下のようになる。CSJ に関しては、頻度 2 以上の語の頻度数がまとまらず分散しているため、CSJ に関しては、頻度 2 以上の結果をまとめて提示し、BCCWJ 同様括弧の中にそれぞれの頻度数を示す。

頻度１：対処、連絡、逝去、相談、入場、提示、注意、指摘、質問、披露

頻度２以上：挨拶(2)、提案(3)、案内(4)、報告(38)、紹介(113)、説明(144)

上記リストの中で BCCWJ において使用されなかった語は「対処、入場、逝去」の 3 語であり、どれも頻度 1 の語である。上記の内訳をみると、16 語の内、半数以上である 10 語が頻度 1 であった。そして、残りの 6 語の中でも、3 語が頻度 5 未満で、5 回以上使用されている動詞は「報告、紹介、説明」の 3 語のみであった。このことから、CSJ においては「ご～する」形式は非常に限られた動詞と場面においてのみ使用されることが確認できる。

５．３　NWJC

分類未完了のため本発表では扱わないが、参考までに載せると、NWJC のデータの中から漢字 2 文字からなる動詞は異なりで 1,591 語あった。まだ作業途中の数であるため、分類作業が終わっておらず「お勉強する」の「勉強」ように習慣的に「お」を取る動詞や「取引」などのように和語動詞のもの、異表記のものなどが含まれている可能性がある。これらを合算すると結果が多少変わる可能性はあるが、この 1,591 語の内、現段階で頻度 1 のものは 615 語あり、これらは誤字や誤認識による結果である可能性も高いと考えられる。そのため、分類が完了すると今後 NWJC の「ご～する」をとる動詞の異なりの数は現在の 1,591 語よりはかなり減るだろうと予想される。しかし、それを念頭において考えるとしても、BCCWJ や CSJ、そして、3 節で述べたような他の方法により収集できる動詞の種類に比べると、NWJC で使用される動詞の異なりは圧倒的に多いと言えよう。また、その内訳から比較すると、NWJC で抽出された動詞には、CSJ・BCCWJ から抽出されたほとんど全ての動詞が含まれる。そして、これらの動詞の中には本来謙譲表現としての「ご～する」の用法を持たず、誤用としてのみ現れると思われる動詞も多数含まれている。

６．　動詞別・コーパス別誤用率

本節では、BCCWJ と CSJ の結果をまとめ、頻度の高い順に並べた時、上位 10 位までの計 10 語の動詞に関して、それぞれのコーパス別に誤用率を求め、比較した。その 10 語の結果を以下の表 1 に示す。

表 1　頻度数上位 10 語の動詞別・コーパス別誤用率

		HK	WK	ET	CSJ	NWJC
ご紹介する	全体用例数	21	272	446	113	400 (824,755)
	誤用の数	2	7	3	1	8
	正用の数	19	265	443	112	392
	誤用率	9.5%	2.6%	0.7%	0.9%	2%
ご案内する	全体用例数	1	38	128	4	400 (98,534)
	誤用の数	0	1	6	1	13
	正用の数	1	37	122	3	387
	誤用率	0%	2.6%	4.7%	25%	3.3%

		HK	WK	ET	CSJ	NWJC
ご説明する	全体用例数	81	12	73	144	400 (38,463)
	誤用の数	15	3	3	1	18
	正用の数	66	9	70	143	382
	誤用率	18.5%	25%	4.1%	0.7%	4.5%
ご連絡する	全体用例数	3	39	28	0	400 (26,556)
	誤用の数	0	9	1	0	47
	正用の数	3	30	27	0	353
	誤用率	0%	23.1%	3.6%	・	11.8%
ご報告する	全体用例数	12	17	39	28	400 (39,865)
	誤用の数	1	2	2	0	15
	正用の数	11	15	37	28	385
	誤用率	8.33%	11.76%	5.13%	0%	3.8%
ご相談する	全体用例数	18	14	26	1	400 (14,008)
	誤用の数	0	8	4	0	173
	正用の数	18	6	22	1	227
	誤用率	0%	57.1%	15.4%	0%	43.3%
ご挨拶する	全体用例数	1	23	21	2	400 (27,615)
	誤用の数	0	3	5	0	53
	正用の数	1	20	16	2	347
	誤用率	0%	13.04%	23.81%	0%	13.3%
ご用意する	全体用例数	0	19	18	0	400 (199,039)
	誤用の数	0	2	0	0	14
	正用の数	0	17	18	0	386
	誤用率	・	10.53%	0%	・	3.5%
ご質問する	全体用例数	24	9	0	1	400 (2,389)
	誤用の数	2	2	0	1	121
	正用の数	22	7	0	0	279
	誤用率	8.33%	22.2%	・	100%	30.3%
ご披露する	全体用例数	5	3	20	0	400 (7,722)
	誤用の数	2	1	2	0	91
	正用の数	3	2	18	0	309
	誤用率	40%	33.3%	10%	・	22.8%

HK=国会議事録、WK＝知恵袋、ブログ、ET=その他（BCCWJ における分類略語）

上記表 1 の動詞は BCCWJ の頻度の高い順に並べられている。これら BCCWJ と CSJ において高頻度の結果を得た動詞は本来「ご～する」の形式で謙譲の意味を表す用法を有するものが多い。そのため、中には当然正用の用例が多く含まれており、その分、誤用率は低くなることが多い。BCCWJ と CSJ はその資料の性質上、そもそも「ご～する」形式はそれほど多く使われておらず、CSJ に比べるとまだその数の多い BCCWJ においても HK、WK、ET に分けているため、殆どの語に関して用例数が少なくなっている。そのため、誤用が 1、2 例増えるだけでも誤用率が大きく増えてしまう。従って、BCCWJ と CSJ のデータのみを用いて誤用率を求め、その傾向を一般化することは難しいように思われる。

一方、NWJC の結果をみると、用例数の最も少ない「ご質問する」に関しても 2,389 件確認され、このリストの中からは全ての動詞において 400 以上のサンプルを取ることが可能であり、比較的に安定した結果が期待できる。

結果の詳細をみると、CSJ を除く残りの 2 資料で最も頻度の高かった「紹介」は「ご紹介する」の形式で謙譲表現として一般的に使われているため、NWJC の中でも最も低い誤用率を記録した。本動詞は BCCWJ の ET のグループにおいても 446 回も使われ、NWJC 以外の結果の中では圧倒的な高頻度を示したが、その中での誤用はわずか 2 件のみであった。

また、NWJC と BCCWJ の WK は両者ともウェブ上の書き込みに関する資料であり、この両者はそれぞれの結果の中で、いくつかの動詞において共通する使用傾向を見せている。両者とも「紹介」の頻度が最も高く、誤用率は最も低い。一方、「質問」と「披露」に関しては、それぞれ頻度順位からいうと、9 位・10 位、10 位・9 位となり、使用頻度が低い動詞に関してもその傾向が共通していることが分かる。しかし、同じく頻度の低い動詞であっても、それはあくまでもそれぞれのコーパス内での傾向の話であり、NWJC と WK は用例の絶対数は大幅に異なるため、低頻度語の場合、絶対的な用例数の少ない WK に関しては、その誤用率は多少信頼性に欠ける面がある。

そして、誤用率が最も高かったのは両者とも「相談」であった。「相談」は、ET においてもかなり高い誤用率を示していた。「相談」という動詞は他の動詞に比べると、やや異なる性格を持つものであり、「相談」という行為は話者単独で完結する行為ではなく、もう一人の相手を必要とする。そのもう一人は会話の相手とは全く無関係な第三者にもなり得るが、直接会話の相手になる場合も多く、用例の中には「あなたとご相談して決める」のように、「と」という格助詞を以て表される場合も多い。このような用例は統語論的に考えた場合、自分が相談するという事実には変わりがないため、正用として認めた。一方、「あなたがご相談する」のように、相談の主体が専ら相手になる用例に関しては誤用に分類している。しかし、「あなたとご相談する」のような用例を全て正用に分類しても、「相談」は高い誤用率を示している。実際の用例の中には、「いつでもメールでご相談してくださいね」のように、「ご相談してください」の用例が多数含まれており、今回対象としているほかの語よりその誤用はかなり定着しているように思われる。その原因の一つとして、前述したような「相談」という行為の性質が考えられる。「あなたとご相談する」のように、相手に直接言及する用例が一般に認められ、用いられているため、他の動詞に比べ、「ご～する」形式を相手に直接用いることについてこれは失礼だという心理的拒否感が弱くなり、その結果「主治医にご相談してみて下さい」のように、自分とは関係のない相手の「相談」行為に関しても用いられやすくなっているように思われる。このように、

誤用として定着している用例に関しては、コーパスの種類に関係なく、似たような傾向が確認された。

７． 考察

最後に、調査した結果をもとに、総合的に NWJC の有用性について評価したい。前述した通り、敬語の誤用の全体像を把握することを目的とする資料の条件には用例の数、使用される動詞の種類の数、話者の多様性、用例の自然さ、データの安定性の 5 点が考えられる。NWJC がこれらの条件に適した資料であるか否かを順に追って考えて行きたい。

第一に、用例の数に関しては、表 1 からも確認できるよう、NWJC は他のデータに比べ、用例を圧倒的に多く収集でき、統計的にも安定的に誤用率を求めることができる。これは、他のコーパスとの比較により、更に際立つ長所である。

第二に、動詞の種類の数に関しても、まだ分類の途中ではあるが、125 語の BCCWJ、16 語の CSJ に比べると圧倒的に多い 1,591 語が確認できた。今後分類が完了したとき、NWJC のデータが多少減るとしても、他の資料に比べると依然として多様な動詞と用例による分析ができる。

第三に、話者の多様性という点に関しては、使用される場面が限られる他の資料に比べ、ウェブ上の書き込みは誰でもアクセスでき、投稿できるという特徴を持つ資料である。例えば、国会議事録はその中の話者が国会会議に出席する国会議員や官僚など、限られた人のデータしか取れない。また、BCCWJ のその他のデータに関しても、小説などの書き言葉が多く、文脈も限られる。これは、CSJ に関しても同様のことが言える。一方、ウェブ上の書き込みを扱う NWJC は、老若男女誰でも自由に書き込むことができ、その全てのデータがランダムに収集されているため、一般に敬語の用いられやすいと想定される企業の書き込みから、日本国の政府機関、各省庁所轄研究所、特殊会社を除く特殊法人などからの掲載、そして、個人のブログに至るまで幅広い日本語母語話者による多様な使用が収集できる。これは、用例の自然さとも関係する問題である。小説のような作為的な文とは異なり、ウェブ上に記載されている企業の広告などの用例はウェブに限らず我々の日常生活の中でも見られる自然な使用である。

最後に、データの安定性が挙げられる。発表者の検索エンジンを用いた同種の研究の経験上、検索エンジンの生の検索結果は結果の件数や、その結果の提示順が検索の度変わるため、安定しない。また、ウェブ上の書き込みは掲載者によっていつでも修正ができるものであるため、毎回調査の度、データを全て自分のパソコンに保存し、コーパス化しておかなければ、後には使えなくなるなどという問題が生じる。これは、動詞をいくつか選定し、その動詞に関してのみ見る研究であれば不可能なことではないが、本研究のように、動詞を幅広く見ることで全体像を把握しようとする研究の場合、検索エンジンのサーバーへの負担やパソコンの負担などから考えると決して容易なことではない。この点に関して、NWJC は既にコーパス化されているデータであるため、安定していると言えよう。

一方、既に安定しているデータであるということは、今後新しい言語使用のデータが追加されにくいことや、検索エンジンのようなリアルタイムの結果を得ることができないという点においては弱点にもなり得る。しかし、データが安定しなければ研究の資料としては使えないということを考えると、NWJC の有用性は否定できない。このような問題点に関しては、検索エンジンのデータを参考にすることで補うことができると考えられる。

８． おわりに

本研究では敬語の正用・誤用を含む使用実態の全体像の調査の資料としての NWJC の可能性について検討した。従来の敬語研究では、コーパスはあまり使われてこなかったが、誤用を含む調査の資料としては、他のデータより求められる用例数と動詞の種類が圧倒的に多く、安定した誤用率が求められる点、多様な位相の話者による自然な用例が得られる点などから、敬語研究にウェブコーパス NWJC が有用なデータとなることが検証された。この点は、本研究だけでなく、他の正用と誤用を含む敬語の使用実態の調査の研究にも活かせる長所であると考えられる。ほかの方法では収集しにくい敬語の誤用の用例を NWJC のようなウェブコーパスを用いて分析することにより、今後の敬語研究の分析の更なる発展が期待される。

謝　辞

本研究の BCCWJ と NWJC の語彙表の作成に際し、国立国語研究所の浅原氏にご協力頂いた。記して感謝申し上げる。

文　献

井上史雄 (1999).『敬語はこわくない—最新用例と基礎知識』講談社
井上史雄編著 (2017).『敬語は変わる—大規模調査からわかる百年の動き—』大修館書店.
伊藤博美 (2009).「承接形謙譲語に関する適切性判断要因と尊敬語転用—『お/ご～する』と『お/ご～される』をめぐって」『日本語学論集』5，pp.162-180(15-33).
蒲谷　宏 (1992).「『お・ご～する』に関する一考察」『辻村敏樹古希記念　日本語史の諸問題』明治書院
川上恭子 (1996).「謙譲語『お～する』『ご～する』の動詞の使用制限」『園田国文』17
菊地康人 (1994).『敬語』講談社学術文庫
北條淳子 (1978).「初級における敬語の問題」『日本語教育』35
金　東奎 (2004).「手紙文」と「スピーチ」から見た敬語接頭辞「お・ご」を用いた「敬語表現」の使用様相『早稲田大学日本語教育研究』4，83:102.
現代日本語研究会編 (2011).『合本女性のことば・男性のことば(職場編)』ひつじ書房.
国立国語研究所 (2016).『敬語表現の成人後採用—岡崎における半世紀の変化—』http://www2.ninjal.ac.jp/longitudinal/files/Late_Adoption_of_Honorific_Expressions_(okazaki_report2016).pdf
小松寿雄 (1967).「『お……する』の成立」『国語と国文学』44:04
文化審議会 (2007).『敬語の指針』http://www.bunka.go.jp/seisaku/bunkashingikai/sokai/sokai_6/pdf/keigo_tousin.pdf
文化庁 (2005).「平成 17 年度『国語に関する世論調査』の結果について」http://www.bunka.go.jp/tokei_hakusho_shuppan/tokeichosa/kokugo_yoronchosa/h17/

関連 URL

コーパス検索アプリケーション『中納言』　https://chunagon.ninjal.ac.jp/
『国語研日本語ウェブコーパス』検索系『梵天』　http://bonten.ninjal.ac.jp/