

NWJCにおける敬語使用とレジスターとの関係

著者	金 賢眞
雑誌名	言語資源活用ワークショップ発表論文集
巻	3
ページ	68-83
発行年	2018
URL	http://doi.org/10.15084/00001639

NWJCにおける敬語使用とレジスターとの関係

金 賢眞 (大阪大学大学院文学研究科)

Relationship between Use of Honorific and Register in NWJC

Hyunjin Kim (Graduate School of Letters, Osaka University)

要旨

発表者は、現在、NWJC を用いた敬語研究を志向し、その一環として、謙讓表現「ご～する」を尊敬用法に用いた「誤用例」の調査・分析を進めている。WEB コーパスは、この種の「誤用例」を収集するのに好適な言語資料だが、残念ながら、NWJC にはレジスターの情報が付与されていない。敬語の使用および誤用にはレジスターによる差があると考えられ、たとえば、ブログなどでは敬語の使用は多くないが、誤用が現れやすい傾向があり、それに対して、企業や公的機関のホームページなどでは敬語が多く使用されるものの、誤用は少ないと予想される。しかし、こうした予想をNWJCによって直ちに確認することはできない。そこで、本発表では、NWJC から得た「ご～する」の使用例を仮設したレジスターに分類し、動詞別の使用頻度や動詞ごとの誤用率がレジスターによってどのように異なるかを統計的に分析して、NWJC における敬語使用とレジスターとの関係を具体的に検討する。

1. はじめに

発表者は現在、国語研日本語ウェブコーパス（以下、NWJC）を使用し、謙讓表現「ご～する」の動詞別の使用頻度、そして、「ご～する」という表現が謙讓語Iであるという性質から考えたとき、その使用が「正用」であるか、それともいわゆる「誤用」であるか、その使用状況を調査・分析している。ウェブコーパスは、この種の「誤用」も含まれるデータを収集するに好適な言語資料であるが、残念ながら、NWJC にはレジスターの情報は付与されていない。

敬語の使用およびいわゆる「誤用」にはレジスターによる差があると推測される。たとえば、サービス業を含む企業や公的機関のホームページなどでは敬語も多く使用され、その性質上、誤用は少ないと予想される。一方、日記や個人同士のやり取りが多いと思われるブログなどでは、そのくだけた雰囲気から、そもそも敬語の使用自体が少なく、使用された場合には「誤用」が現れやすいと考えられる。しかし、NWJC で確認できる情報は URL のみで、こうした予想をNWJCによって直ちに確認することはできない。

しかし、動詞別の「敬語の誤用」を調査している以上、資料のレジスターを把握することは必要である。もし、レジスターによる差が明確に存在し、NWJC が偶然その特徴的なレジスターの資料を多く含んでいるだけであれば、それは調査全体の結果の信用を損なうためである。また、仮にそのような差が存在するとしても、その事実を知らずにデータを利用しているか、それともそれを理解した上で利用しているかでは結果の解釈が大きく異なる。

そこで、本発表では、NWJC から得た「ご～する」の使用例を仮設したレジスターに分類し、動詞別の使用頻度や動詞ごとの誤用率がレジスターによってどのように異なるかを統計的に分析し、NWJC における敬語使用とレジスターとの関係を検討する。具体的には、NWJC におい

て確認できる唯一の情報である URL を利用し、その「ドメイン」を分析することによって、NWJC というデータのレジスターの特徴を把握することを試みる。

2. 研究の枠組み

2.1 謙譲表現「ご～する」の正用と誤用

現代の「ご～する」の規範的用法として規定されるのは、「謙譲語I」の用法のみである。「謙譲語I」用法は大前提として行為の主体は必ず自分自身、もしくは、話し手がウチの人間として捉えているもので、「客体」との何らかの関係をもつ使用でなければならない。この何らかの関係としては二つがあげられる。一つは(1)のように、その動詞がヲ格、ニ格などに客体をとる、「客体」に直接関わる行為である。

(1) ↓今月コチラを申し込みされた方をご招待します

もう一つの関係は、前述したものとは異なり、文そのものには客体が形式として明示的には現れないが、「客体のために」行われる行為で、(2)のように、その行為の結果が客体に恩恵を与える行為である。

(2) その人がもっともお似合いになる「白」をご提案します

蒲谷(1992)は前者の使用をI類、後者の使用をII類と説明している。以下の記述では蒲谷(1992)のI類とII類という用語を援用したい。

I類の中には、一般的に動詞そのものはI類に分類されない行為であっても、(3)と(4)のように特定の使用場面においては客体に直接関わりをもつ用例もある。

(3) 色々とご注文しましたが、納車までスムーズにご対応頂けて有難うございます

(4) 今後もご利用したいと思いますので宜しくお願いします

(3)と(4)は文には客体が形式として明示的には記されず省略されているが、共通して「(相手の)商品・サービス」をヲ格の補語によって取る表現である。これらは客体と直接、関係し、補語となる対象を高めている使用であるため、「謙譲語I」として解釈できるものである。

一方、実際使用されている「ご～する」の用例には上述のような「謙譲語I」としての用法ではなく、以下のように、相手・相手側、もしくは、謙って表現すべきでない相手がする行為について使用されている例が目立つ。

(5) 『カード決済』をご利用する事はできません

(6) 中国産の低価たばこは取り扱っておりませんのでご安心してください

(7) ロコミで広まって 100 万人がご利用しています

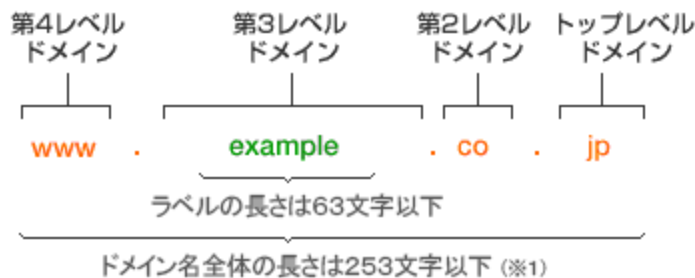
これらは厳密に言うとそれぞれレベルは異なるが、全て「謙譲語I」としての使用ではないという点においては共通している。

本発表のデータでは、(1) から (4) までのような用例を「正用」として認め、(5) から (7) までのような用例は「誤用」として分類する。

2.2 ドメインとは

ドメインとは、簡略に説明するとインターネット上で使用されている住所のことである。具体的には、「www.example.co.jp」のようなものであり、「.」で区切られている各部分は「ラベル」と呼ばれる。ドメイン名を構成するラベルの中で、最も右側に位置するラベルを「トップレベルドメイン」と称し、以下左へ順に「第2レベルドメイン」「第3レベルドメイン」などのように続く。これを図式化すると、<図1>のようになる。

ドメイン名の構成



<図1>ドメイン名の構造

(出典：一般社団法人日本ネットワークインフォメーションセンター (JPNIC))

この中で、登録されているウェブページの性質を区別する重要な基準となり得るドメインはトップレベルである。ドメイン名をトップレベルドメイン (TLD) で分類すると、分類別トップレベルドメイン (gTLD : genericTLD) と国コードトップレベルドメイン (ccTLD : country code TLD) に大きく分けることができる。gTLDはTLDだけでそのドメインの用途と登録対象が確認できるが、ccTLDの場合、その用途と登録対象を確認するためには、TLDだけでなく、第2レベルドメインまで確認する必要がある。

そこで、本発表では、基本的にはTLDを基準にドメイン进行分类し、ccTLDであるJPドメイン名に関しては、第2レベルまで进行分类基準とする。各ドメインの詳細な説明に関しては紙面の都合上、割愛するが、必要に応じて随時説明を加えていきたい。なお、その場合、各ドメインの詳細の説明は株式会社レジストリサービス (JPRS) の「JPドメイン名の種類」の説明ページ (<https://jprs.jp/about/jp-dom/spec/>) を参考に記述する。

2.3 コレスポネンス分析

コレスポネンス分析は対応分析とも呼ばれる分析の手法で、1960年代にフランスの Jean-Paul Benzecri によって提唱され、1970年代以降に広まった比較的新しいものである。小林 (2010) では、コレスポネンス分析について、以下のように説明している。

コレスポネンス分析は、データ表の行や列に含まれる情報を少数の成分（コレスポネンス分析では次元 [dimension] と呼ぶ）に圧縮し、それらの関係を散布図上に布置することで、視覚的なデータの俯瞰を可能にする。¹

コレスポネンス分析は、アイテムごとのカテゴリ間に潜む複雑な関係性の分析に有用な手法であり、本発表においても、ドメインと使用される各動詞と誤用率の関係性を明らかにする方法として最も有効な方法であると思われる。従って、本発表では、分析のための統計手法としてこのコレスポネンス分析を利用する。

2.4 先行研究

田野村 (2012) は現代日本語書き言葉均衡コーパス（以下、BCCWJ）に収められた特定目的サブコーパスという言語資料の特性について分析しており、同論考の中で、BCCWJの利用者はBCCWJが複雑な内部構成を有することと、BCCWJが従来の日本語研究にはあまり使われてこなかった新種のデータを含んでいることに注意しなければならないことを指摘した。

本発表で分析するNWJCは従来の日本語研究においてあまり使用されてこなかった新種のデータとしての性質が強い。そのため、NWJCを利用するときには、BCCWJ以上に資料の特性を把握することが必要になると思われる。しかし、それにも関わらず、NWJCはレジスターが分類されていないため、その複雑な内部構成を『梵天』によっては直ちに確認することはできない。そこで、本発表では「ご~する」という限られた言語表現からではあるが、NWJCから収集された結果から誤用率を求め、「ご~する」の誤用率におけるNWJCのレジスターとデータの特性を考察したい。

また、田野村 (2012) は、BCCWJに格納されているYahoo!ブログのデータから、ブログという媒体が言語研究の資料としてどのような性質であるかについて分析し、その結果、ブログは一般に抱かれている「個人が毎日あるいは気の向いたときに近況やエッセイを書いてインターネット上に公開する」という印象とは異なり、現状としてはその少なからぬ部分を各種の宣伝目的の記事と外部からの引用による記事が占めていることを指摘した。

NWJCによる「ご~する」の用例収集の結果からブログのデータを分類しても、その傾向は同様であり、実際、必ずしもブログだからと言って、個人が気の向くままに書いているものとは限らないことが確認できた。自分でホームページを作成する技術を持たない個人経営の小規模店舗が、広報の必要性和ホームページ作成を依頼する経費を勘案し、手軽に作成・管理できるブログをホームページ代わりに活用している例が多い印象であり、本来のブログの用途と言

¹ 小林雄一郎 (2010) 245ページより抜粋

われる、完全な個人の近況報告などはむしろ少ないように思われた。また、実際にはもう一つの用途として、既にホームページを持っている企業や団体であっても客との心的距離を縮める、もしくは、その内容更新の手軽さから、簡単な情報発信のためにブログを活用することもあり、単純に「ブログ」というだけでは、その性質は定義できない。本考察では、その点も踏まえ、NWJCのデータの性質について総合的に考えていきたい。

3. 調査概要

3.1 調査目的

本発表は資料としてのNWJCの性質を把握することをその目的とする。具体的には、「ご～する」という特定の敬語表現の調査結果を基に、NWJCの中のデータにはどのようなレジスターの特徴が見られ、各ドメインをどのように分類することができるかについて考察する。

3.2 調査資料

本発表では国語研日本語ウェブコーパス(以下、NWJC)を調査資料とする。NWJCは国立国語研究所が2017年に公開したウェブを母集団とする100億語規模のコーパスで、2014年10月から12月までの間に収集したデータを格納している。

3.3 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」など、「お」の付く形式も含まれるが、「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、「ご～する」の誤用とはその性質が異なる可能性が考えられる。従って、本発表では、「お～する」は対象外とし、「ご+漢語サ変動詞語幹+する」のみを分析の対象としている。

その中でも、NWJCにおいて「ご～する」形式を多く取っている動詞を頻度数順に上位30語まで選定した。その30語を以下に示す。動詞の前の数字はNWJCにおける「ご～する」形式での頻度数の順位を表しているものである。

- 「01. 紹介」 「02. 用意」 「03. 案内」 「04. 提供」 「05. 報告」 「06. 説明」
- 「07. 挨拶」 「08. 提案」 「09. 連絡」 「10. 利用」 「11. 相談」 「12. 購入」
- 「13. 招待」 「14. 来店」 「15. 披露」 「16. 安心」 「17. 協力」 「18. 奉仕」
- 「19. 対面」 「20. 確認」 「21. 使用」 「22. 参加」 「23. 注文」 「24. 訪問」
- 「25. 指導」 「26. 理解」 「27. 注意」 「28. 心配」 「29. 満足」 「30. 予約」

次に、30語の動詞をそれぞれ「ご～する」形式に入れてNWJCから得られた用例をダウンロードし、その中からランダムで各400例ずつを選定した。そして、用例の文脈を見ながら、そのデータを正用と誤用に分類した。その過程の中で、30語の内、使用の傾向が偏っている、コーパスから得られる結果のみでは前後の文脈や人物同士の関係の把握ができないなどの理由によって分析対象として不適切であると判断した「18. 奉仕」「19. 対面」の2語は分析対象

から除外し、最終的には総計 28 語が対象となった。そして、分類の結果から、各動詞の「ご～する」の誤用率を求め、その結果を今回の分析の対象とする。

3.4 調査方法

まず、調査対象をまとめ、各動詞の正用と誤用の実数をドメイン別に再集計した上で、各動詞と各ドメインの誤用率のクロス表を作成する。そして、その結果を基に統計解析ソフトである IBM SPSS Statistics によりコレスポンデンス分析を行い、その結果について考察する。

ただし、クロス表内に欠損値が多くては正確な分析の妨げになるため、次のような補正処理を行った。第一に、補正として欠損地が多い動詞を削除した。その基準としては、他のドメインと明白に性質が異なるドメインに現れる動詞を残すことにした。特に、公的機関のドメインであり、「公的なテキスト」としての性質が強いという性質から一番規範意識が高いと思われるドメイン `go.jp` の結果を最大限残すべく、`go.jp` で使用されていない動詞を削除した。その結果、「01. 紹介」「02. 用意」「03. 案内」「05. 報告」「06. 説明」「08. 提案」「09. 連絡」「10. 利用」「11. 相談」「15. 披露」「17. 協力」「20. 確認」「25. 指導」「26. 理解」の 14 語が残った。そこから、更に教育機関という明確な性質をもつドメイン `ac.jp` において 1 例も使用されていない「05. 報告」「08. 提案」「15. 披露」「20. 確認」の 4 語を削除した。残りの 10 語の内、「01. 紹介」「02. 用意」「03. 案内」の 3 語においてはほとんど全てのドメインが誤用率 0%、「10. 利用」「26. 理解」においてはほとんど全てのドメインが誤用率 100%という偏った傾向をみせていた。コレスポンデンス分析では、このようにドメインによって動詞の誤用率に差がみられないものを分析の対象として含めてしまうと、動詞とドメインの誤用率の関係を説明する際に信頼性が大きく損なわれる。従って、上記の 5 語を削除して、ドメインによって誤用率の傾向が異なっている 5 語、「06. 説明」「09. 連絡」「11. 相談」「17. 協力」「25. 指導」のみを分析対象として残した。

そして、ドメインの中でも、性質が全く同じドメインであるため、まとめずにそのまま分析する意味のないものは一つのカテゴリにまとめた。たとえば、以下の<表 1>において、「local」という項目は、`.local` という純粋なドメインを表しているわけではなく、`tokyo.jp`、`osaka.jp` のような地域型 JP ドメイン名をまとめたものである。地域型 JP ドメイン名は詳細な地域によって第 2 レベルドメインこそ異なるが、その性質は全て同様であるため、本研究においてはまとめても特に問題のないものである。地域型 JP ドメイン名は地域を表すドメイン名として、地方公共団体・特別区およびその機関、他の属性型 JP ドメイン名の登録資格を満たす組織、日本に在住する個人、病院が登録できるドメイン²である。

また、<表 1>には「他の国」という項目も存在するが、これも理屈としては「local」と同様である。まとめる前のデータには `ac`, `at`, `bz`, `ca`, `cc`, `fm`, `in`, `is`, `md`, `me`, `ms`, `nu`, `sc`, `sg`, `th`, `to`, `tv`, `us`, `vc`, `ws` のようなドメインが収集されたが、これらのドメインは、全て `ccTLD` で、日本の `jp` ドメイン名のようなものである。詳細なドメイン名は異なっても、基本的にはその中に有意な差は

² 株式会社レジストリサービス (JPRS) 「JP ドメイン名の種類」の説明ページ (<https://jprs.jp/about/jp-dom/spec/>) を修正・加筆している。(2018年07月13日最終確認)

存在しないと考えられるため、これらも「他の国」として一つの項目にまとめている。この「他の国」項目に含まれる ccTLD は基本的には登録資格に特に制限はなく、誰でも自由に登録できるドメインである。

ほかのドメインに関しては、「local」と「他の国」ほどドメイン間の性質が完全に一致するものは見られず、それぞれ程度の差はあれど、異なる特徴を有するため、ドメインそのものを特定基準によって任意にまとめたものはこの二つだけである。

最後に、調査対象の動詞 5 語のうち、2 つ以下の動詞でしか使用されていないドメインを削除した。この過程で 1 語の動詞においてのみ使用されている ed.jp、travel、2 語の動詞においてのみ使用されている lg.jp が除外され、最終的にドメインは ac.jp、biz、co.jp、com、go.jp、info、jp、「local」、ne.jp、net、or.jp、org、「他の国」の 13 項目が分析対象として含まれた。

4. 調査結果

まず、分類の結果を以下<表 1>に示す。<表 1>では「local」とorgにおいて「指導」が空欄になっているが、これは総計 1 例も使用されなかったものである。

<表 1>動詞・ドメインの使用頻度・誤用・正用・誤用率のクロス表

		ac.jp	biz	co.jp	com	go.jp	info	jp	local	ne.jp	net	or.jp	org	他の国	総計
06 ・ 説 明	総数	4	11	41	150	19	4	90	8	9	50	3	4	5	398
	誤用	0	0	1	6	2	0	8	0	1	0	0	0	0	18
	正用	4	11	40	144	17	4	82	8	8	50	3	4	5	380
	誤用率%	0	0	2	4	11	0	9	0	11	0	0	0	0	5
09 ・ 連 絡	総数	3	1	38	210	3	7	81	1	15	29	4	3	2	397
	誤用	1	0	7	16	0	1	11	1	2	5	0	1	1	46
	正用	2	1	31	194	3	6	70	0	13	24	4	2	1	351
	誤用率%	33	0	18	8	0.0	14	14	100	13	17	0	33	50	12
11 ・ 相 談	総数	1	1	31	160	2	9	121	1	35	27	9	1	2	400
	誤用	1	1	12	73	0	3	55	1	10	12	4	0	1	173
	正用	0	0	19	87	2	6	66	0	25	15	5	1	1	227
	誤用率%	100	100	39	46	0	33	46	100	29	44	44	0	50	43
17 ・ 協 力	総数	4	6	10	222	2	9	89	2	5	43	5	1	2	400
	誤用	4	5	6	196	0	6	78	1	5	32	5	1	2	341
	正用	0	1	4	26	2	3	11	1	0	11	0	0	0	59
	誤用率%	100	83	60	88	0	67	88	50	100	74	100	100	100	85
25 ・ 指 導	総数	5	7	28	157	1	23	99		27	38	9		3	397
	誤用	5	6	11	105	1	15	49		11	14	4		1	222
	正用	0	1	17	52	0	8	50		16	24	5		2	175
	誤用率%	100	86	40	67	100	65	50		41	37	44		33	56
総計		30	60	362	1775	37	83	925	21	182	388	61	14	33	3971

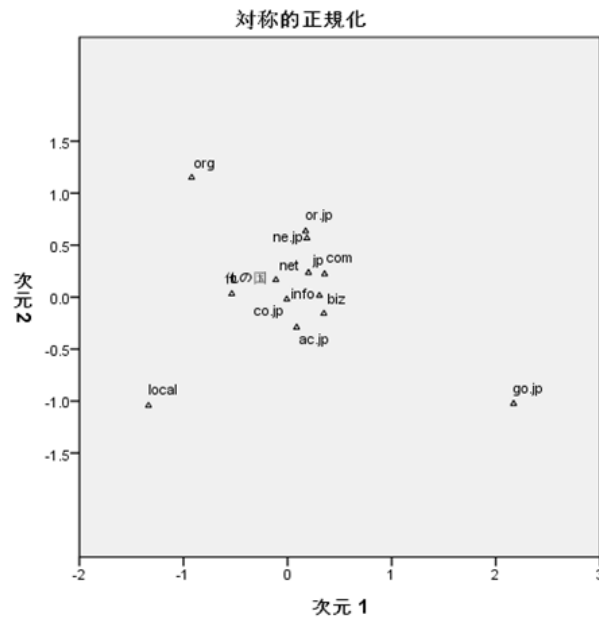
次に、<表 1>の結果により求めたコレスポネンス分析の結果の要約は以下<表 2>の通りである。

<表 2> コレスポネンス分析結果の要約

次元	特異値	要約イナーシャ	カイ 2 乗	有意確率	イナーシャの寄与率		信頼特異値	
					説明	累積	標準偏差	相関
								2
1	.473	.224			.586	.586	.013	.214
2	.290	.084			.220	.805	.018	
3	.241	.058			.152	.957		
4	.127	.016			.043	1.000		
要約合計		.382	1008.266	.000 ^a	1.000	1.000		

a. 自由度 48

コレスポネンス分析の結果、4 の次元が抽出され、次元 2 までの累積寄与率は 80.5%であった。以下では、この結果の散布図を基に分析を進める。



<図 2> ドメインの行ポイント散布図

<図 2>の結果をみると、10 個のドメインが原点に近い中央の方に集中しており、残りの 3 個のドメインは中央に位置するドメインを囲み、次元 1 を底辺とする直角三角形を形作るよう

に位置していることが確認できる。具体的には、次元 1 の正方向の特徴は **go.jp** によって、負方向の特徴は「**local**」と **org** によって特徴づけられ、次元 2 の正方向の特徴は **org** によって、負方向の特徴は「**local**」と **go.jp** によってそれぞれ特徴づけられている。

中央のグループには商用ドメインである **co.jp** と **biz** や、教育機関である **ac.jp** などのような特徴的なものをはじめ、特徴を限定しにくい **com**、**net** などのようなドメインまで、多様なドメインが含まれている。これら、中央のグループは、これまでのように TLD と第 2 レベルドメインだけをみると共通する特徴がまとまらず、解釈が難しいが、一つの共通点がある。それは、このグループに含まれる結果は全体的な URL をみると、**or.jp** を除いては、全てブログという特殊なものが含まれているということである。具体的には、今回分析の対象となった用例総計 1992 件の中で、URL だけでブログであることが判明するものが 938 件と、ほぼ半数近くを占めており、**ac.jp**、**biz**、**co.jp**、**com**、**info**、**jp**、「**local**」、**ne.jp**、**net**、「他の国」においてブログでの使用が見られた。しかし、上記に見られるように、中央ではないところに布置されている「**local**」において 1 例、また、URL そのものからは判定しないため、上記の数字には含まれていないが、分析の中で元のリンクに辿って確認した結果、**org** にもブログでの使用が 1 例確認された。一方、中央に布置されている **or.jp** においては、ブログの使用は見られない。以上の結果をまとめると、NWJC において収集されるデータの中には、全体的にブログの用例が多く、このブログの用例が何か影響を与えている可能性もあると思われるが、中には例外も存在し、それが具体的にどのように影響しているかまでは今回のデータだけでは解釈できない。

また、ブログとは言え、その用途と性質が一律ではない点、**org** と「**local**」においても 1 例ずつではあるがブログの用例が含まれている点などを考えると、単純にブログが特徴的な性質を示すと解釈するのは、多少説得力が弱いように思われる。全体の用例数の割合から考えても、ほとんど全ての用例が含まれている中央グループのドメインがこの原点に近いところに布置されているということに鑑みると、NWJC において収集されるデータは、「ご~する」の誤用においては特徴的な偏りは少ない可能性が高い。

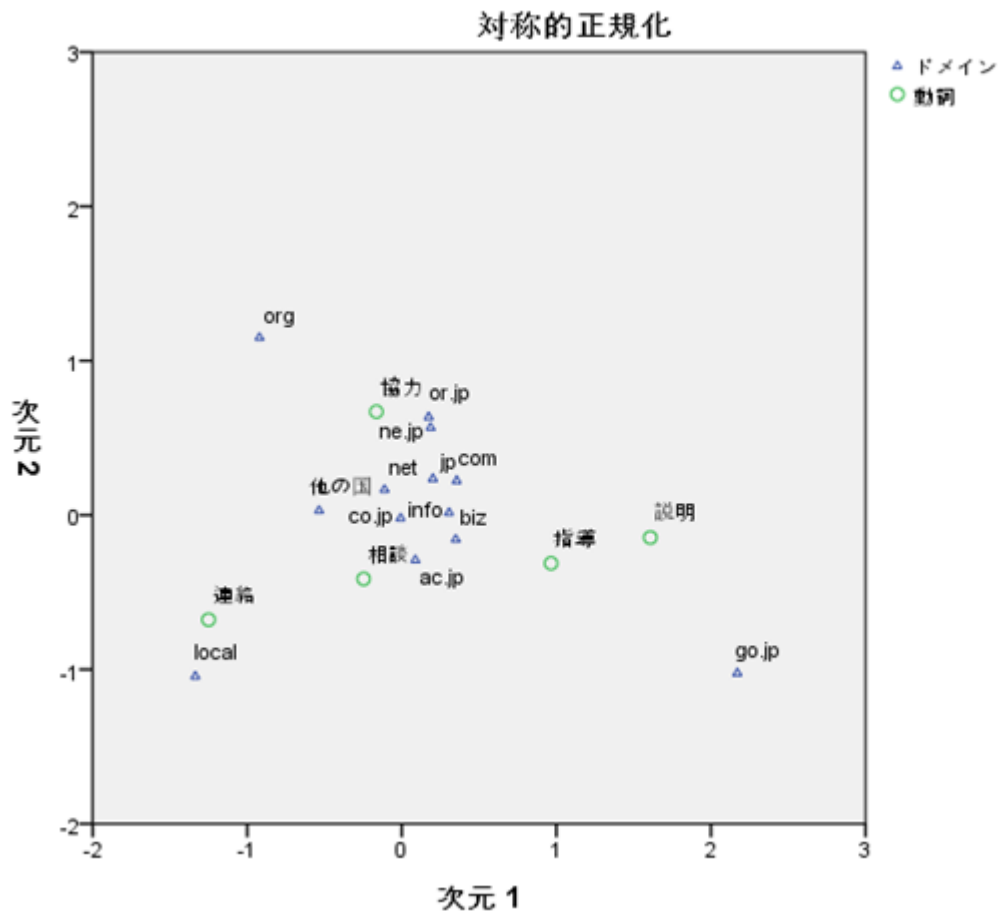
一方、特徴的に現れた三つのドメインは以下のような性質をもつドメインである。以下の記述において、**go.jp** と「**local**」のドメインは株式会社レジストリサービス (JPRS) の「JP ドメイン名の種類」のページの説明を参考にしている。1 つ目の **go.jp** は日本の政府機関や各省庁所管の研究所、特殊法人、独立行政法人が登録できるドメインであり、中でも政府機関は、一つの組織で複数の **go.jp** ドメイン名を登録できる。2 つ目の「**local**」(地域型ドメイン) は地域を表すドメイン名として、地方公共団体・特別区およびその機関、他の属性型 JP ドメイン名の登録資格を満たす組織、日本に在住する個人、病院が登録できるドメインである。最後の **org** は用途としては非営利組織用のドメインであるが、その登録条件には制限がないものである。

まず、次元 1 として、正方向に特徴づけられた **go.jp** は、公的機関としての性質が最も強いドメインであり、一般的に考えると最も誤用が現れにくい、強い規範意識が働いているドメインである。一方、負方向に特徴づけられた「**local**」は、**go.jp** のように政府機関である「特別区およびその機関」のページも含まれるが、**go.jp** とは異なり、ほかの属性型 JP ドメイン名の登録資格を満たす組織か日本に在住する個人でも登録できるドメインであるため、公的機関としての性質と、その他の性質が混在しているドメインである。また、**org** は非営利組織で、通常の用語の定義から考えると、非営利組織の中に政府組織は含まない。このことから、次元 1

では純粋な公的機関としての性質が強いドメインであるか、それとも、公的機関を全く含まない、もしくは、他の性質が混在し、純粋な公的機関でないドメインであるかが特徴として現れているように思われる。一方、次元 1 においては、org と「local」は近いところに布置されているが、非営利組織の中に一部公共的な活動を行う団体が含まれるため、両者は公共的な団体としての性質が共通している。

次に、次元 2 では org と go.jp、「local」の間に特徴的な差が見られる。次元 2 において、「local」と go.jp に共通する性質で、org とは異なる最も大きな性質は、公的機関、中でも政府機関のページが含まれているか否かであると解釈される。

つまり、次元 1 においても、次元 2 においても、共通して公的機関のもつ性質が大きく作用していると推測される。



<図 3> ドメインの行ポイント・動詞の列ポイントの散布図

<図 3>は行ポイントである「ドメイン」と、列ポイントである「動詞」の結果を合わせた散布図である。レジスターを分析するためには、本来であれば影響を与え得る要因を全て総合的にみていかなければならない。しかし、今回は「ご～する」形式での動詞の誤用の性質とい

う極限られた現象に注目して分析を行っている上に、13 のドメインに対し、その特徴を解釈するための動詞の数は5語だけになっている。従って、本来のコレスポネンス分析のように最も近くに特徴づけられているものだけを考慮して分析してしまうと、その傾向性が解釈しにくいと想定される。そこで、本来のコレスポネンス分析の手法とは若干異なるが、動詞に関しては単純に近いところに布置されているものを基準に分析するのではなく、各次元において正方向で現れているか、負方向で現れているかを基準にし、<図 2>のドメインだけの結果において特徴的に現れたドメインと関連付けて解釈していきたい。

動詞に関しては、次元1では「指導」「説明」が正方向に布置され、「相談」「協力」「連絡」が負方向に布置されている。しかし、この中で「協力」と「相談」は、負方向に布置されてはいるが、原点に近いものであるため、次元1においてはそこまでの特徴は示されていない可能性が高いと予想される。一方、次元2に関しては、「協力」だけが正方向に布置され、「連絡」「相談」「指導」「説明」は負方向に布置されている。ここでも「説明」と「指導」は0に近いところに布置されているため、この2動詞に関しては、次元2においては有意味な結論は得られない可能性が高い。

ドメインと動詞を合わせて考えると、まず、次元1からは、go.jp が正方向に特徴づけられ、同じく正方向に布置されている動詞は「説明」と「指導」の2語であった。一方、「local」とorg は負方向に布置されており、動詞の中では「協力」「相談」「連絡」が同じく負方向に布置されていた。<表 1>のそれぞれの「誤用率」からも確認できるよう、go.jp においては、第一次元において同じく正方向に布置された動詞である「説明」と「指導」においてのみ誤用が現れ、負方向に布置された「協力」「相談」「連絡」においては、1例の誤用も現れていない。これに対し、負方向に布置された「local」とorg は正方向に布置された動詞「説明」「指導」に関しては1例の誤用も現れておらず、同じく負方向に布置された動詞「協力」と「連絡」において誤用が現れた。

もう一つの負方向の動詞である「相談」はorg に関しては誤用が現れておらず、「local」でのみ誤用が現れる。このように対照的な結果を示すorg と go.jp であるが、<表 1>の誤用率に照らし合わせて考えると、「相談」に関しては、誤用率が0%であるという点において両者に共通点が見られ、org とドメインとしては共通する性質を含めていた「local」に関してのみ差がみられる。動詞の解釈で既に述べている通り、<図 3>において、「相談」は原点に近い動詞であるため、今回の結果においては、特徴が弱く、解釈が難しい。

「協力」も「相談」同様、第1次元においては原点付近に布置されているため、次元1に関してはあまり有意味な特徴は見られないと予想される。しかし、次元2に関しては、正方向においてorg と「協力」の間に関連性が見られ、政府機関を含まない非営利組織で用い誤用が現れやすいという「協力」の誤用率の特徴が説明できると解釈される。

以上をまとめると、次元1において、純粋な公的機関のドメインにおいて「説明」と「指導」は誤用が現れやすく、「連絡」は誤用されにくいと思われる。一方、次元2においては、「協力」は政府機関を含まない非営利組織において誤用されやすいと解釈される。以下では、これらの解釈について、事例をもって検証する。

まず、動詞の中で最も正方向に布置されている「説明」は go.jp において総計 19 例確認され、その中で 2 例が誤用、17 例が正用で、誤用率は 10.5%であった。以下 (8) と (9) は go.jp におけるその誤用例であり、(10) から (12) まではその正用例である。

(8) その上で、この後ちょっと御説明していただくほうで議論したほうがいいのかもかもしれませんけれども、(後略) (go.jp)

(9) 今、実態につきましては課長からご説明したとおりでございます (go.jp)

(10) あなたは、こういう事情を私が御説明しても、そうではないというふうに言い張られますか (go.jp)

(11) ○＝年金課長 400 万人は幾つかの統計を組み合わせて推計したもので、ちょっと何もしないで説明するのは大変ですので、次回以降で御説明したいと思います (go.jp)

(12) 文章のほうで引き続き説明させていただきますが、ここでご説明させていただく試験につきましては繁殖能に関する指標として産卵率、孵化率、育成率、それとこれら 3 つの指標をかけた総合繁殖指数というものでご説明したいと思います [ママ] (go.jp)

今回収集された用例の中で go.jp において「説明」が使用されている場面は、全て何かの会議の「議事録」であった。一方、org と「local」に現れている「説明」の正用には、以下のようなものがあった。

(13) 上下水道料金の支払方法についてご説明しています(「local」)

(14) お客様と直に接してご説明し、表情やご反応を見、時にはご質問やご意見をいただけるギャラリー・トークは、私たちにとっても貴重な場なのです(「local」)

(15) ・採用後、1ヶ月間を研修期間とし、教室の理念や指導についてご説明します(org)

(16) ※弊社からのご案内メールが正しく届かなかった時はお電話にてご連絡します(org)

以上のように、org と「local」の「説明」の用例は、一般向けに発信される公的な文書に使用されているが、議事録のものは見られない。会議の中では、自分側から相手に何かを説明することも、逆に、相手から説明されることも多い。また、「会議」という公的な場面では、敬語を使用しなければならないという意識が強く働き、相手の「説明」という行為に関しても無理に「ご～する」という形式を使用しているため、誤用が現れているように思われる。公的機関の中でも政府機関では、「会議」の「議事録」が一般に公開されることが多い。コレスポン

デンス分析の結果だけでは「純粋な公的機関であるか否か」が次元 1 において大きく影響を与えると解釈したが、さらに具体的には、「純粋な公的機関」の中でも、特に go.jp に含まれている「会議」という場面の、相手が存在する話し言葉としての性質が「説明」において特徴的に現れているように思われる。

次に、正方向に布置された「指導」に関して分析する。「指導」は go.jp において 1 例確認され、その 1 例が誤用で、誤用率 100%となった。その用例を以下 (17) に示す。

(17) また、流された人を助けるための「スローバッグ」の投げ方や受取り方などの方法についてご指導していただきました (go.jp)

「指導」は go.jp、org、「local」の 3 ドメインを合わせても 1 例しか使用されておらず、(17) の用例がその唯一の用例であった。このことから、go.jp、org、「local」のような公的機関や公共的団体のドメインは、他の性質をもつドメインに比べて相対的に、何かを「指導する」という行為は想定されにくいように思われる。(17) の用例は go.jp の「徳島河川国道事務所」で使用された用例で、このページは各種講座などについての開催結果を報告しているものである。基本的には直接他人に何かを「指導」するという行為は想定されにくく、指導するとしても、機関の内部の人が講師を務めるのではなく、外部から講師を招聘することになる可能性が高いことが公的機関の特徴の可能性の一つとして考えられる。

次に、次元 1 において最も負方向に布置されている動詞「連絡」と「local」と org、go.jp の間の関係について考察する。まず、go.jp では、「連絡」は 3 例使用されているが、その 3 例は全て正用であり、誤用率は 0%となっていた。正用の用例は以下、(18) から (20) までのようなものであり、このことから、公的機関は自分側から相手に連絡をするだけで、相手からの「連絡」を求めることは基本的にはないと解釈される。

(18) 審査の結果は 2 月 18 日 (水曜) を目処にご連絡します (go.jp)

(19) また、募集締切後に受講の可否をご連絡します (go.jp)

(20) 参加の可否については、返信はがき又は FAX でご連絡します (go.jp)

しかし、実際の用例をみてみると、「local」で現れている「誤用」も公的機関である政府機関の使用であったため、<図 3>の結果から解釈したように、特徴的に現れている要因が単純に「純粋な公的機関であるか否か」と解釈するだけでは不十分であることが確認される。

「local」では「連絡」が 1 例だけ確認され、その 1 例が誤用であったため、誤用率 100%となっている。その用例は以下の (21) である。

(21) 利用日の 7 日前までにご連絡して下さい (「local」)

(21) は小平市のホームページで、他市の保養施設の利用方法を説明しているページにおいて使用されている用例である。同じ政府機関であっても、国全体をまとめる go.jp とは異なり、都道府県と市町村と、比較的狭い範囲に限られるものであるため、基本的にはその地域に居住している居住者のために発信する情報であり、go.jp に比べ、読み手との関係が近いと思われる。このことから、次元 1 は散布図で解釈したように、単純に「純粋な公的機関であるか否か」だけで区別されているわけではなく、国全体をまとめる公的機関全般での使用であるか、それとも、読み手との関係がより近い、狭い範囲での使用であるかによっても特徴づけられる可能性があると思われる。しかし、「local」に現れている「連絡」の用例が正用・誤用に関係なく (21) の 1 例しかなかったため、今後分析データが増えれば、結果が変わる可能性も高い。

最後に、次元 2 において、「協力」と org の間の誤用率の関係性について考察する。「協力」は org において 1 例使用されており、その 1 例が誤用であったため、誤用率 100%となった。org における「協力」の誤用の用例は以下のようなものであった。

(22) 治験が治験実施計画書（治験を安全に行うにあたって、治験を依頼した製薬会社が決めたルール）を遵守（じゅんしゅ）しているかを細かくチェックしながら、治験にご協力していただいている患者さまからの相談（心のケア）や質問に対する対応を行っています
(org)

(22) は 2018 年 7 月現在はリンクが切れており、全文の確認はできないが、URL からして病院の案内文の一部であることが確認できる。

これに対し、次元 1 において「org」と同様に負方向に布置されていた「local」では、「協力」は 2 例使用されており、その中の 1 例が誤用として現れ、誤用率 50%となっていた。「local」の「協力」の誤用を (23) に、正用を (24) に示す。

(23) ご来場いただきました皆さんには、アンケート調査にご協力していただきまして、ありがとうございました（「local」）

(24) （前略）あるいは県営名古屋空港がそれに御協力する機会があるのか、そんなことは当然これからしっかりと検討していかなければなりませんし、（後略）（「local」）

(23) は「ちよだ区議会だより」から抽出された用例で、区民大会の結果を報告したものである。

一方、go.jp では、「協力」は 2 例使用され、その 2 例とも正用の用例であった。具体的には、次のような用例が収集された。

(25) 自主共聴組合さまからのご要望により、以下についてご協力します (go.jp)

(26) したがって、当面の施策としては、先般商工委員会で十一時半ぐらいまでやっていただきまして上げていただいたあの法律の適用ということで私ども大蔵省としての立場か

らも精いっぱい御協力することがあればしていくというのが今日の対応策でございます
(go.jp)

go.jp においては、誤用は現れなかったが、「local」において誤用として現れている (23) の例が政府機関の使用であるため、公的機関としての性質が「協力」の誤用率に影響しているとは解釈できない。しかし、そもそも次元2の寄与率は次元1の寄与率の半分にも満たないことが<表2>より確認できる。このことを総合的に考えると、次元2において現れた特徴は、今回の用例の実数の少なさによるものであり、実際のところは特徴として説明できない可能性が高いと考えられる。

5. おわりに

本発表ではドメインをレジスターとして仮設し、「ご~する」のドメイン別・動詞別の誤用率についてコレスポネンス分析を行うことにより、ドメインをNWJCのデータのレジスター類似のものとして利用する可能性を試みた。その結果、今回のデータに限って言うと、NWJCにより収集されるドメインにおいて、ドメインと動詞との関係の中で特徴的に現れているものは極一部だけであることが確認された。しかも、これら、特徴の見られないドメインがNWJCにより収集されるドメインのほとんど全てを占めており、NWJCはさほど偏りのないデータのように解釈される。「ご~する」の誤用において一部特徴的に現れているものは公的機関での使用のように、全体の日本語の中では極めて限られた部分であるように思われる。つまり、NWJCに含まれているドメイン同士の特徴にそこまで深刻な差は見られず、NWJCを利用して敬語を分析するにあたり、特定ドメインの集中による結果の偏りはあまり現れない可能性が高いように思われる。

今回の結果からは、「ご~する」の誤用においては、go.jp、org、「local」の3ドメインにおいて最も顕著な特徴が見られた。そして、その特徴は純粋な公的機関のドメインであるか否か、さらに、純粋な公的機関のドメインであっても、国全体をまとめる公的機関全般での使用であるか、それとも、読み手との関係がより近い、狭い範囲の公的機関での使用であるか、また、議事録が含まれたデータであるか否かによって説明される可能性があることが明らかになった。

しかし、今回の分析に関しては、以下のようないくつかの問題点があることを指摘したい。まず、そもそも対象としている動詞の数が少ないため、ドメイン別の性質を明確に説明しにくいことが問題として挙げられる。次に、各動詞の用例数も400と限られているため、そこからさらにドメイン別に分類をすると、ドメイン別の用例数が少なくなりすぎるという点である。そのため、1例の変動だけでも誤用率が大きく変わり、結果が変動してしまうため、それが統計の結果に大きな影響を与えている。最後に、「local」のように、複数の性質が含まれているドメインの存在と、ブログのように性質の定義が曖昧なものが各ドメインに含まれている点が結果の解釈に悪影響を及ぼしている可能性も存在する。

このことを踏まえて今回の結果を考えると、NWJCにおいて「ご~する」の動詞別誤用率に関する研究にドメインによる結果の偏りの問題はそこまで大きな影響を与えてはいないように思われるが、go.jpのように、一部明確な性質を持つものについては、他とは区別される明確

な特徴がみられる。しかし、今回の調査は、用例数の少なさ、一部ドメインの性質の定義の難しさなどが統計結果に影響を与え、事実を正確に反映していない可能性も考えられる。これについては、データの数を増やし、更にドメインをなるべくカテゴリ化した上で再度分析を行いたい。これを今後の課題とする。

文 献

- 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, pp.141-157.
- 金 賢眞 (2015) 「謙譲表現『ご~する』の誤用—公的テキストにおける実態とその要因—」大阪大学文学研究科修士論文 (未公開)
- 国立国語研究所コーパス開発センター編 (2017) 『国語研日本語ウェブコーパス』 (2014-4Q データ, 梵天バージョン 1.0.0) <https://bonten.ninjal.ac.jp/> (2018年07月13日最終確認)
- 小林雄一郎 (2010) 「第10章 コレスポネンス分析; データ間の構造を整理する」石川慎一郎・前田忠彦・山崎誠編『言語研究のための統計学入門』くろしお出版, pp.245-264.
- 田野村忠温 (2012) 「BCCWJに収められた新種の言語資料の特性について: データ重複の諸相とコーパス使用上の注意点」『待兼山論叢. 文化動態論篇』46, pp.59-83.

関連 URL

- 一般社団法人日本ネットワークインフォメーションセンター (JPNIC)
(<https://www.nic.ad.jp/ja/>) (2018年07月13日最終確認)
- 株式会社レジストリサービス (JPRS) 「JP ドメイン名の種類」の説明ページ
(<https://jprs.jp/about/jp-dom/spec/>) (2018年07月13日最終確認)
- 『国語研日本語ウェブコーパス』検索系『梵天』 <http://bonten.ninjal.ac.jp/>
(2018年07月13日最終確認)