

日本語文章の可読性評価のための 項省略判断モデル

久保田智天¹ 石月由紀子² 松林優一郎^{1,3}

¹ 東北大学 ² 無所属 ³ 理化学研究所

tomotaka.kubota.s4@dc.tohoku.ac.jp yukiko.ishizuki.p7@gmail.com

y.m@tohoku.ac.jp

概要

本研究では、日本語の書き手が行う述語の項省略の適切さを、読み手側の視点に立った際に適切に指摘できるかという新しいタスクの設計を検討する。この設定は、項省略の可否を第三者の指摘に基づいて学ぶ学習支援システムの開発を見据えたものである。この問題設定を精緻に検討するために、我々は、読者側が得られる情報のみを用いて項省略の不自然さを指摘できるかについて、判断の要因になりうる複数の観点に基づく質問の回答をアノテーションするデータを作成し、結果を詳細に分析する。

分析の結果、我々の設定した省略判断の観点の組み合わせによって、読み手の実際の省略判断を高い割合で説明できることが示された。

1 はじめに

日本語では、主格や目的格といった述語の項がしばしば省略される。この項の省略は、文意を的確に伝達するために、省略内容を読み手が正確に補完しながら読めるよう適切にコントロールされる必要がある。しかし、この観点で項省略の可否を説明する試みには、これまでに十分な研究が見られない。また、この能力は一般に言語コミュニケーションの中で経験的に体得されており、体系的、効率的な学習は困難である。したがって、省略判断の基準となる要因の分析や、文章内の項省略の可否を指摘できる計算モデルを構築することで、人間の項省略判断能力の向上に貢献できる可能性がある。

同様の動機を持つ先行研究として、Ishizukiら [1]の研究がある。Ishizukiらは、人間の実際の項省略判断データを収集し、判断の要因となる要素を分析した。この研究では、アノテーターは書き手の意図（省略しようとする項の内容 = 項の filler）を予め

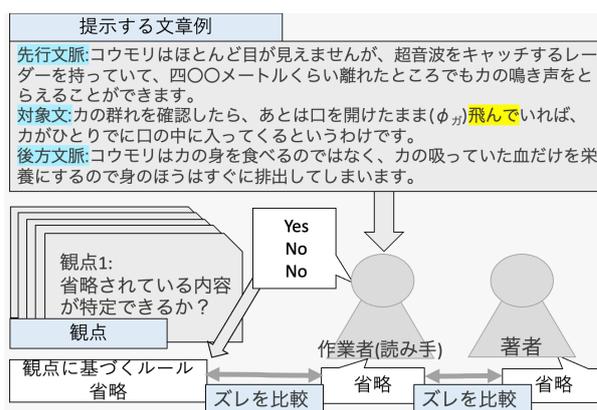


図1 項省略判断タスクの概略図。文脈情報を元にして、省略されている項の省略の可否を判断する。

知ったうえで、与えられた文脈の中でその項が省略できるかという判断を行っていた。しかしながら、読み手視点では省略された項を補いにくい不適切な省略を書き手が行った場合、項省略後の文章のみからは書き手の意図する元の項の内容を知ることができない。したがって、実際の執筆支援システムの応用場面では、省略された項の filler が何であったかという書き手のみが知る情報が秘匿された状況下で、項の不足を指摘する必要がある。この状況は、人間同士の対話で、聞き手が文脈理解のための情報に不足を感じたり、誤解等、互いの認識に齟齬がある疑いを感じた時点で話し手側に確認や補足を求める行為に類似性がある。そこで、本研究では、Ishizukiらの項省略判断タスクを、(1) 省略された項の filler を知ることができない状況下における、(2) 情報の不足を感じる状況、文脈に違和感を感じる状況の検出、という観点で再整理し、その下で、(3) 省略判断のアノテーションデータの作成、(4) 判断の基準となる要因の分析、および、(5) 読み手と熟達した著者による省略判断のずれについての分析を行う。

BCCWJ [2] から抽出した 100 事例の分析の結果、我々の設定した判断の観点に基づく省略判断結果と

読み手の直感的な省略判断は、93%という比較的高い割合で一致することが分かった。また、読み手と著者の省略判断の一致率は76%であり、読み手と著者は一定程度異なる基準を持つことが示された。本稿の後半では、結果を詳細に分析する。

2 関連研究

省略解析 省略現象に関連して、英語では動詞句省略に関連した研究や日本語では省略された述語の項を同定するゼロ照応解析が行われている[3, 4, 5, 6, 7]。関連研究では、省略された要素に補完する要素の同定が目的であり、省略された要素に対する省略可否判断に焦点を当てた研究は少ない[1]。

可読性評価 読み手にとっての文の読みやすさを文章の評価に取り入れた例については、Flesch Kincaid Reading Ease や、日本語リーダビリティといった可読性指標が提案されている[8, 9, 10]。これらの指標は、読み手のレベルに合わせた語彙や表現の易しさの側面から文章単位の評価を行う。本研究では、述語の項の省略が読み手にもたらす読解の困難さをアノテーションし、文章の可読性に影響のある項省略判断を分析する。

3 読み手視点での省略判断タスク

書き手によって執筆され、述語の特定の格が省略済みの文章に対し、その文章の表層から得られる情報を頼りに、対象の格の省略の可否を指摘するタスクを定式化する。タスクの概要を図1に示す。このタスクでは、(a) 著者が実際に項を省略している事例と、(b) 元の文章では項が述語との係り受け関係を持って出現しているが、その項のフレーズを意図的に削除して擬似的に不完全な文章にした事例の2種類を作成し、これらのデータ中に生じる情報の不足や、解釈の齟齬の可能性を検出するタスクを行う。具体的な(a), (b) 事例の作成方法は付録Aに記載した。

タスクの作業者は、対象文の前方 k 文の先作文脈と、対象の述語を含む文、対象の述語、対象の格助詞、後方 k 文の後作文脈が与えられ、省略の可否判断を尋ねられる。本研究では、読み手が省略の可否を判断する観点として、大きく分けて2つの観点を仮定をする。1つ目は、その文章の読解に際して、対象の格の filler が分からない事による情報の不足を感じるかという観点であり、2つ目は文脈を読み進めるうちに解釈の齟齬の可能性を認識できるかと

いう観点である。省略判断タスクは、この仮定に基づいて設定された5つの質問に答える形で行われる。作業者は、前方文脈から対象の文までを読んだ時点で、指定された格について、

- (1): 省略されている内容が、読解に支障がない程度に特定できる
- (2): その項の内容を特定することが、読解に必ず必要である
- (3): 項の内容の特定に、2~3回の読み返しが必要な程度の負荷がかかり、困難を感じる

かどうかをそれぞれ Yes/No の2値で判定し、さらにそこから後作文脈を読み、

- (1b): (観点1) が No の場合に、次の1文を読めば特定可能である
- (4): 後方の文脈を読んだ際に、自分のそれまでの解釈に疑いを持つ

かをそれぞれ Yes/No の2値で判定する。この観点別の回答の組み合わせに対応させる形で、想定される省略判断をあらかじめ定める事ができる(表1のルールによる判断)。しかし、実際にはここで想定していないより詳細な判断基準もあると予想できるため、作業者は、これら5つの項目に回答することに加えて、最終的に「省略されていても読解に強いが負荷かからず自然に読めるか」という観点で「省略されていても良い/表出してほしい」という判断を OK/NG の2値で行う。

分析では、我々があらかじめ想定した回答パターンに対する判定ルールと、最終的な作業者の判断とのずれを分析することで、より詳細な判断要因を洗い出す。また、このタスクを通して検出する、情報不足や解釈の違和感に起因する「読み手の省略不可能判断」が、書き手の判断とどのように一致するか、ずれが生じた部分は実際に理解の齟齬を生む可能性があるかどうかの検討を行う。

4 アノテーションデータの作成

対象データ 現代日本語書き言葉均衡コーパス(BCCWJ)[2]に対して係り受け、並列構造、述語項構造、共参照情報が付与された BCCWJ-DepParaPAS[11]に含まれる書籍ドメイン 83 文書の中から事例を抽出する。3節冒頭で説明した(a), (b)の2種類の事例を作成する。この際、作業コストの観点から、対象文の前方文脈と後作文脈の長さは $k=5$ 文と定めた。(a), (b)それぞれランダムに50事例、計100事

表 1 観点別の判断結果に基づく省略判断ルール，および作業者の判断結果. Fは概ね表出と判断されると想定しているが，一部は省略と判断される可能性も想定する.

パターン	観点					ルールによる判断	事例数	作業者の最終判断とルールとのずれ	作業者 vs. 著者	
	1	1n	2	3	4				表出 vs. 省略	省略 vs. 表出
A:	Y	-	Y	Y	N	表出	3	0	1	0
B:	Y	-	N	Y	N	省略	2	0	0	0
C:	-	-	-	-	Y	表出	1	0	0	0
D:	Y	-	-	N	N	省略	52	3	0	14
E:	N	Y	Y	Y	-	表出	3	0	0	0
F:	N	Y	Y	N	-	表出?	3	2	0	0
G:	N	Y	N	-	-	省略	1	1	0	0
H:	N	N	Y	-	-	表出	25	1	4	0
I:	N	N	N	-	N	省略	10	0	0	5
Total							100	7	5	19

例を収集した。各事例は、(b)の事例に関し、削除された項の内容が露呈することを極力防ぐため、判断対象の文が重複しないよう収集した。

抽出事例の絞り込み Ishizukiら [1]が収集した省略判断データによれば、BCCWJ-DepParaPASのアノテーションにもとづいて項の省略事例を無作為抽出すると、並列構造や機能動詞構文等における構文構造上の項共有など、複数人のアノテーター間で容易に省略判断が一致する事例がサンプルに入ることが分かっている。本研究では、このような自明な判断となる事例を極力避けるため、(a)、(b)事例の抽出時に、同一文内に項の filler が現れてしまうケースは収集しないという戦略を取った。具体的には(a)の事例は、文間ゼロ照応のみを対象とし、(b)の事例は、対象の文から対象の直接係り受けの項を削除したあとに、同一文内に削除した項と共参照関係にある表現が現れていない文のみを抽出対象とした。

作業者間の揺れ 著者のうち2名が3節の省略判断タスクを全件並行してアノテートした。作業者の最終的な省略判断には、一致率で25%の揺れが見られた。2名の作業結果を著者らで比較検討した結果、これらの揺れは、作業者の読解能力や、厳密な解釈を求める傾向によって本質的に起こり得るものであり、一つの判断結果に集約することは困難であるという結論に至った。したがって、今後はより多くの作業者による判断の分散を加味しつつ分析を行う必要がある。本稿の次節以降の分析では、作業者1名分の結果に絞って判断ルールや著者とのずれの要因を分析する。

5 アノテーション結果の分析

表1に観点別の判断結果の事例数、ルールと作業者の最終判断にずれが生じた件数、著者の省略判断と読み手の最終判断にずれが生じた件数を示した。元文章の事例数は表出50事例、省略50事例であるのに対し、ルールでの判断は省略が65事例、作業者の最終判断も省略が64事例と、やや省略可能な側に傾いている。このずれの分析については後の段落で行う。観点別の判断に基づくルールによる省略判断と、読み手の最終判断との一致率は93%であり、本研究で設定した省略判断の観点を読み手の省略判断を一定程度説明できることが示唆された。

観点別の判断結果のパターンとしては、Dの省略52件、Hの表出25件、Iの省略10件で、これら上位の3パターンが全体の87%をカバーした。

判断ルールと作業者判断のずれ パターンルールに基づく判断と作業者の最終判断には7件のずれが生じており、これらの事例には予め列挙されていない別の何らかの判断要因が内在すると考えられる。そこで、これらの事例に含まれる要因を分析した結果、「読み手が厳密な意味理解を求める場合」、「理解できるが文脈上初出の場合」、「物語文における動作主の明示」等の要因が見られた(詳細は付録Cを参照)。このうち、最も多い要因は「読み手が厳密な意味理解を求める場合」であった(3/7件)。これは、表2の(事例1)のように、「500万」という対象の項は容易に類推できる一方で、文脈上、金額の表記が厳密な理解のために重要であると読み手が考えたために「表出が望ましい」と判断した事例である。このように、各文脈における個別の情報要求に応じて、ここで取り上げたものの他にも更に複数

表2 ずれの生じる事例。

事例	文章	読み手の判断	ずれの対象
事例1	・・・中島さんは、知り合いの田中さんから、貸したお金500万円を返せという裁判を起こされてしまいました。・・・(中略)・・・この裁判で中島さんは、確かに(φ _ヲ) [借り pred] たが、そのうちの100万円は、田中さんではなく田中さんの親から借りたものだ。	表出	判断ルールでは省略
事例2	・・・ぼくは何か悪さをした子どもみたいに答える。「ああ、そう」女はぼんやりと言い、パスタもかじきまぐるも中途半端に残したまま、グラスにワインを満たしている。ちびりと(φ _ヲ) [飲ん pred]] で、舌で唇をなめる	表出	著者は省略
事例3	・・・仕方なく雑誌でも買いに行こうと夜半にホテルを抜け出した。・・・(中略)・・・適当に二、三冊、(φ _ヲ) [買った pred] た。	省略	著者は表出

の細分化された観点が存在すると考えられるため、今後は分析事例を増やし、更に精緻な分析を行う必要がある。

読み手判断と著者判断のずれ 読み手と著者の省略判断にずれがあった事例は全体で24件あった。読み手が「表出してほしい」と判断した46件に対し著者が省略をした事例(表出 vs 省略)は5件であり、11%のずれが生じた。一方で、読み手が「省略されていても良い」と判断した64件に対し著者が表出させた事例(省略 vs 表出)は19件あり、30%程度のずれが生じた。

読み手は表出と判断し、著者は項を省略した事例 読み手が表出を求めている項を省略することは、読解の妨げを生む可能性があり、注意深く観察すべき事例であるが、各事例のずれの要因を分類した結果、(a)「読解能力による揺れ」(2/5件)、(b)「参照する前方文脈を制限した事による影響」(2/5件)、(c)「物語文における動作主の明示」(1/5件)に分類された。(a)については次の段落で述べる。(b)は今回のデータ設計上の制約であり、より前方の文脈にアクセスできれば、ずれは生じないことが確認された。(c)の例を表2の事例2に示す。これは、物語文のような複数の登場人物の発話・行動・視点が頻繁に入れ替わる状況においては、各動作の動作主が容易に推論できる場合であっても、言動の流れを明示化することを目的として著者が動作主体を積極的に表出させるために発生するずれである。

読み手は省略と判断し、著者は項を表出した事例 このタイプのずれの要因には、「読解能力による揺れ」と「物語文における動作主の明示」が顕著に見られた(詳細は付録Dを参照)。「読解能力による揺れ」は、例えば表2の事例3のように、「雑誌」という対象の項は文脈から比較的容易に推定できると想定できるが、省略された項の推論負荷は読み手の読解能力に応じて異なるため、幅広い読解能力の読み

手に対して正確に情報伝達することを目的として、著者が表出と判断する事例である。一方、「項の繰り返しによる特定の情報の強調」や「状況の詳説」「文章表現の制約(慣用的表現・対比表現)」といった書き手の執筆戦略や表現上の制約に起因する要因も見られた。これは文章表現の向上の側面から発生する事例であり、項を表出させることで読み手に対して過剰な情報の冗長性を与えるものではないことを確認した。省略判断システムの作成にあたり、読み手の視点に基づく判断に加えて、書き手の執筆戦略に基づく省略判断を考慮する事の必要性が示唆される結果である。

6 議論と今後

本研究では、文章の読み手視点から、項省略の可否を判断するタスクの定式化を試みた。問題設定を精緻に検討するために、読者側が得られる情報のみを用いて項省略の不自然さを指摘するための判断の要因になりうる観点を整理し、アノテートしたデータを作成し、結果を詳細に分析した。分析の結果、我々の設定した読み手の省略判断の可否に関する観点は、読み手の実際の省略判断と高い割合で一致し、その適用可能性が示された。本研究の結果を踏まえ、各観点に関する分類モデルの構築による、読み手視点の項省略判断モデリングが期待される。

加えて、省略判断のずれの分析によって、読み手と書き手間の省略判断に関する戦略の差異や、読解力の異なる読み手間の省略判断の差異を検討する余地があることが明らかになった。今後の方向性として、分析の事例数および、アノテーター数を拡充し、より詳細な省略判断の要因分析を行うとともに、読み手の読解能力と読み手間の省略判断の分散の関係等のより精緻な分析を行うことを検討している。

参考文献

- [1] Yukiko Ishizuki, Tatsuki Kuribayashi, Yuichiroh Matsubayashi, Ryohei Sasano, and Kentaro Inui. To drop or not to drop? predicting argument ellipsis judgments: A case study in Japanese. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)**, 2024.
- [2] Kikuo Maekawa, Makoto Yamasaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. Balanced corpus of contemporary written Japanese. **Language resources and evaluation**, Vol. 48, No. 2, pp. 345–371, 2014.
- [3] Marjorie McShane and Petr Babkin. Detection and resolution of verb phrase ellipsis. **Linguistic Issues in Language Technology**, Vol. 13, , 2016.
- [4] Wei-Nan Zhang, Yue Zhang, Yuanxing Liu, Donglin Di, and Ting Liu. A neural network approach to verb phrase ellipsis resolution. **Proceedings of the AAAI Conference on Artificial Intelligence**, Vol. 33, No. 01, pp. 7468–7475, Jul. 2019.
- [5] Ryohei Sasano, Daisuke Kawahara, and Sadao Kurohashi. A fully-lexicalized probabilistic model for Japanese zero anaphora resolution. In Donia Scott and Hans Uszkoreit, editors, **Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)**, 2008.
- [6] Ryohei Sasano and Sadao Kurohashi. A discriminative approach to Japanese zero anaphora resolution with large-scale lexicalized case frames. In Haifeng Wang and David Yarowsky, editors, **Proceedings of 5th International Joint Conference on Natural Language Processing**, 2011.
- [7] Ryuto Konno, Shun Kiyono, Yuichiroh Matsubayashi, Hiroki Ouchi, and Kentaro Inui. Pseudo zero pronoun resolution improves zero anaphora resolution. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, 2021.
- [8] Rudolph Flesch. A new readability yardstick. **Journal of applied psychology**, Vol. 32, No. 3, p. 221, 1948.
- [9] JP Kincaid. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. **Chief of Naval Technical Training**, 1975.
- [10] 李在鎬. 日本語教育のための文章難易度に関する研究. PhD thesis, Waseda University, 2016.
- [11] 浅原正幸, 大村舞. Bccwj-depparapas: 『現代日本語書き言葉均衡コーパス』係り受け・並列構造と述語項構造・共参照アノテーションの重ね合わせと可視化. 言語処理学会第 22 回年次大会発表論文集, pp. 489–492, 2016.

表3 評価データにおける事例の内訳

格助詞	(a)	(b)	合計
が	42	25	67
を	6	15	21
に	2	10	12
合計	50	50	100

表4 判断ルールに含まれない読み手判断の要因.

事例	パターン			
	D	G	F	H
読み手の厳密な意味理解	2	1	-	-
後続文脈を読むまでの負荷	-	-	1	-
物語文における動作主の明示	1	-	-	-
理解できるが文脈上初出	1	-	-	-
会話文	-	-	1	-
参照する前方文脈の制限による影響	-	-	-	1
Total	3	1	2	1

A アノテーション事例の作成方法

タスクの事例は以下のように作られる。まず、ある文章内の特定の1文が選定され、その中の1つの述語が選定される。この述語に対して、文脈上何らかの指示先のエンティティが存在する格が、ガ格、ヲ格、ニ格のうちから一つ選択される。この選択された格の項は、述語と同一文内にその実態を指し示す表現が存在しないか(a事例)、文内で対象述語と直接係り受け関係を持って出現しているか(b事例)、のいずれかであるように選択される。さらに前処理として、(b)の項が述語と係り受け関係を持って出現している事例については、その項のフレーズが句単位で削除される。

B 事例の内訳

作成した評価データの事例の内訳を、格属性ごとに表3に示す。

C 判断ルールと最終判断のずれの要因

判断ルールと読み手の最終判断のずれの要因を、判断ルールのパターンごとに分類したものを表4に示す。

「文意理解の厳密性」・「物語文における動作主の明示」については本文中の説明を参照する。「後続文脈を読むまでの負荷」は、後続文を読んで意味を理解するまでの読解の負荷がかかるかどうかで、項の省略判断が決まる事例である。「理解できるが文

表5 読み手は表出と判断したが、書き手は項を省略していた場合の要因.

事例	パターン	
	A	H
読解能力による揺れ	-	2
参照する前方文脈の制限による影響	1	1
物語文における動作主の明示	-	1
Total	1	4

表6 読み手は省略と判断したが、書き手は項を表出していた場合の要因

事例	パターン	
	D	I
読解能力による揺れ	3	2
物語文における動作主の明示	4	-
文章表現の制約(慣用的表現・対比表現)	3	1
繰り返しによる特定の情報の強調	2	1
状況の詳説	2	1
Total	14	5

脈上初出」は、省略された項を理解はできるが、文脈上初めて出現する情報であるため、表出が望ましい事例である。「会話文」は、項の省略判断という観点からは表出を選択するが、対象文が会話文であるために、実際には省略が望ましい事例である。

D 読み手判断と著者判断のずれの要因

読み手の判断と著者の判断のずれの要因を、判断ルールのパターンごとに分類したものを表5、表6に示す。

「読解能力による揺れ」、「物語文における動作主の明示」については、本文中の説明を参照する。「文章表現の制約」は、慣用的表現や対比表現といった文章表現の制約によって書き手が項を表出する事例である。「繰り返しによる特定の情報の強調」は、書き手が特定のトピックを強調するために、項を繰り返し表出する事例である。「状況の詳説」は、書き手がより詳しく状況を説明するために、特定の情報を伝える項を意図的に表出する事例である。