

# 日本語語順分析に言語モデルを用いることの妥当性について

栗林樹生<sup>1,3</sup> 伊藤拓海<sup>1,3</sup> 鈴木潤<sup>1,2</sup> 乾健太郎<sup>1,2</sup>

<sup>1</sup> 東北大学 <sup>2</sup> 理化学研究所 <sup>3</sup> Langsmith 株式会社

{kuribayashi, t-ito, jun.suzuki, inui}@ecei.tohoku.ac.jp

## 1 はじめに

語順には、文脈の影響を受け特定の要素が強調された語順（有標語順）と、文脈の影響がない場合に支配的な語順（基本語順）が存在する。本研究では後者の基本語順について、日本語の書き言葉に焦点を当てて分析を行う。日本語は語順が比較的自由な言語であるが、多くの研究によって基本語順の存在が示唆されている。基本語順とテキストの読みやすさは密接に関わっており、特に日本語教育やライティングフィードバックのために、基本語順の仕組みを記述・整理することは重要である。

しかしながら、基本語順を決定する要因については様々な説が存在する。要因の解明に向けて、言語学や自然言語処理の分野において、様々な仮説の検証が行われてきた [1, 5, 8, 16]。既存研究では、主に心理実験に基づく検証やコーパス頻度による仮説の検証が主流であり、それぞれの検証方法には長所・短所が存在する (2章)。例えばコーパス頻度による分析は、コストが低く網羅的な仮説検証に向いているものの、対象の言語現象を数えるためにしばしば述語項解析などの前処理が必要であり、自動的な同定が難しい現象 (例えば、非明示的な格の出現順序など) に関する仮説の検証には向いていない [16]。

本研究では、ニューラル言語モデルを用いた基本語順の仮説検証方法を検討する (図 1)。本検証方法は、従来の検証方法とは異なる長所・短所を持つため、これまで実現できなかった分析が可能になることが示唆される (3章)。本検証方法による結果の妥当性を (i) 言語モデルの生成確率が高い語順と人間が基本語順であると考える語順の相関 (4章)、(ii) 様々な仮説における、既存の検証方法による結論と言語モデルを用いた検証による結論の一致 (5章) から示す。(ii) において、本研究では基本語順に関連する 10 個の仮説に焦点を当てる。本論文の実験結果から、言語モデルを用いた方法論の妥当性が示唆された。

6章では、言語モデルを用いた方法論の長所を生かした分析の例として、格助詞が明示的に出現しない場合の語順について、日本語のとりたてに関する仮説に焦点を当てた分析を行う。検証結果から、「とりたて助詞を伴う格の位置は、とりたて助詞や、とりたてられる格の種類に依存して傾向が異なる」「基本語順で前方に位置する格ほど主題になりやすい」「目的語の主題になりやすさは動詞のをに率の影響を受ける」ことを支持する経験的な根拠を提示する。

本論文の貢献は以下の通りである。

- 言語モデルを基本語順の分析に用いる方法論の長所について整理した。
- 二種類の実験結果から本分析方法の妥当性を示した。
- これまで大規模な実験が困難であったとりたてと語順の関係について、言語モデルを用いて大規模に分析を行った。

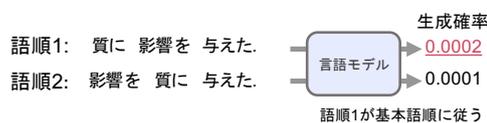


図1: 言語モデルを用いた基本語順の推定

## 2 既存の仮説検証方法

基本語順に関する仮説の検証には大きく三つの方法がある。

**人間の反応に基づく検証:** 語順に関する仮説を検証する典型的な方法として、読み時間や脳波などの人間の反応を観察する方法論があげられる [17]。これらのアプローチは、実験コストの高さから大規模な検証には向いていない。また、用いられた用例や被験者によるバイアス、実験の再現性の低さなどの懸念もあげられる。

**頻度に基づく検証:** 検証したい言語現象が実際の文章 (コーパス) 中で出現する頻度を数える、頻度に基づく検証も典型的に用いられる方法論である。多くの研究が各語順のコーパス上での頻度と基本語順には関係があるという仮定をおいている [7, 16]。この方法論の長所として、網羅的な仮説の検証や多くの用例を用いた実験を行うことが可能であることが挙げられる。しかしながら頻度に基づく方法論では、検証の対象となる言語現象をコーパス中で数えるために、その言語現象の存在を同定する必要がある。この同定にはしばしば述語項解析器などの解析器が用いられ [16]、対象とする現象を解析器が正しく同定できるという仮定を担保する必要がある。時に興味の対象となる現象の同定は困難であり、この制約は研究の可能性を狭めることとなる。例えば笹野ら [16] は、述語項解析器のエラーを防ぐために、助詞が明示的に出現する用例のみを収集し、格要素の主辞のみを分析に用いている。したがって、彼らの研究では上記の条件を満たさない用例・言語現象に関しては分析ができていない。このような問題は、解析器が十分に整備されていないリソースの乏しい言語で特に重大な問題となる。

Bloemら [2] は、3 グラム言語モデルを用いて頻度を数えた。この研究は我々の研究に近いと考えられるが、頻度ベースの言語モデルを用い、ドイツ語の two-verb cluster 現象のみに焦点を当てて分析を行っている点で本研究とは異なる。

**数理モデルの内部を分析する検証:** 対象となる言語現象の発生を予測する解釈が容易なモデル (例えば、ロジスティック回帰モデル) を学習し、モデルの内部 (例えば、素性に対する重み) を分析する検証方法も挙げられる [1, 3]。この方法では、各要因の影響度の強さといった高度な統計情報を得ることができ、モデルを学習するための大規模なデータの作成や素性の設計などには、しばしばコストがかかる。

### 3 言語モデルを用いた仮説検証方法

#### 3.1 概要

言語モデルを用いた仮説検証方法では、語順の異なる用例同士を言語モデルの生成確率で比べる。最も言語モデルの生成確率の高い語順が基本語順に従うという仮定をおく(図 1)。この仮定の妥当性を4, 5章で検証する。言語処理の分野では、言語モデルの生成確率はテキストの容認度を推定するために用いられ、英語語順に関する特定の言語現象では、言語モデルの生成確率と人間の語順選好に相関があることも示されている [4]。

#### 3.2 言語モデルを用いた検証方法の長所

頻度に基づく方法論では、2章で述べたとおり、対象となる言語現象の出現をコーパス中で高精度に同定可能な解析器を必要とする。対照的に言語モデルを用いた方法論では、ある仮説に従った語順と従わない語順のペアといった、仮説検証のための評価データを必要とする。解析器の準備と評価用データの準備のどちらが容易かは状況に依存する。

例えば、格助詞が省略された際の語順を分析するという状況を想定する。ヲ格が省略された際に語順 (1) -a と語順 (1) -b のどちらが自然かという問いである。

- (1) a. 生徒に 本を あげた。
- b. 本を 生徒に あげた。

この検証では、評価用データの準備の方が高精度な解析器の準備よりも現実的である。省略された格助詞を述語項解析器などで高精度に同定することは実現が難しいが、格の順序を検証する既存のデータを編集して、格助詞が省略された場合の格の順序の自然さを検証するデータを作成することは容易である。

基本語順は人間の語順選好に基づくため、人間の反応に基づく検証の結果は信頼性が高いと考えられる。しかしながら、この検証方法は時にコストが高く、大量の用例を用いたり、数多くの仮説を検証していくことには向いていない。一方で言語モデルを用いた検証方法では、人間の代わりに言語モデルを用いるため、言語モデルを用いた検証方法は比較的成本が低く、網羅的な仮説の検証に向いている。さらに、言語モデルを用いた検証は人間の反応に基づく検証に比べて、実験の再現性が高いと考えられる。

適切な検証方法は状況に応じて異なり、検証方法の選択の幅を増やすことは言語学分野において有益であると考えられる。また近年は、様々な言語モデル実装や言語資源、言語モデルの学習に利用可能なクラウドコンピューティングサービスが利用可能であり、本分析方法の実装可能性を後押ししている。また、言語モデルと人間の語順選好が必ず一致する保証はないため、言語モデルを用いた検証は、あくまでコストの高い検証の前段階の検証として用いられるべきであると考えられる。

#### 3.3 検証方法の妥当性をどのように担保するか

本論文では、二種類の実験結果を示すことで言語モデルを用いた検証方法の妥当性を示す。一つ目に、言語モデルの語順選好が人間が基本語順であると考えられる語順に高い相関があることを示す(4章)。二つ目に、様々な仮説において、言語モデルを用いた検証の結果が、既存の検証方法により得られた結論と一致することを示す(5章)。

もし、言語モデルを用いた検証による結論が、既存の検証方法による結論と一致するならば、本検証方法は既存の検証方法

と同等の仮説検証力をもつことが支持される。本検証方法の妥当性が示されれば、これまで実現できなかった検証を可能にすることが期待できる。

#### 3.4 設定

次の単語を予測する Transformer ベースの言語モデル<sup>\*1</sup>を用いた。文字レベルの言語モデル (CLM) と、サブワードレベルの言語モデル (SLM) を用いた。SLM への入力は、UniDic を用いた MeCab で形態素に分割したのち、サブワード<sup>\*2</sup>に分割した。30 億 Web ページから無作為に抽出した 1.6 億文を言語モデルの学習に用いた。4, 5, 6章で使用するデータは、言語モデルの学習で用いたデータと重複しない。各用例に対するスコアを、左から右に生成する言語モデルによる生成確率と、右から左に生成する言語モデルによる生成確率の積で求めた。

### 4 実験 1: 人間との相関

本章では、語順の自然さを比較するタスクにおいて、言語モデルの選好と人間の選好が高い相関をもつことを示す。

**アノテーション:** 30 億 Web ページから無作為に収集した 1 万文それぞれに対して、語順を変えたかき混ぜ文を一種類作成した。かき混ぜ文の作成方法は栗林ら [18] と同様であり、依存構造で兄弟関係にある要素を入れ替えている。ヤフークラウドソーシングを利用し、元の語順の文とかき混ぜ文のペアに対してラベルをつけた。ラベルは、(a) 文 1 が自然である、(b) 文 2 が自然である、(c) 意味的に崩壊した文が存在する、のうちから一つ選ばれる。ワーカーには文ペアのみを見せ、特定の文脈を想像しないように指示している。

チェック設問を利用して明らかに不自然なラベルを付与しているワーカーを除き、チェック設問を通過した合計 756 人のワーカーによるラベルを用いた。それぞれのペアに対して 10 人のワーカーがラベルをつけた。(i) 意味的に崩壊した文が存在するというラベルが付与されていない、かつ (ii) 10 人中 9 人以上が同じラベルを付与した、約 2,600 ペアを採用した。

**結果:** 文 1 と文 2 のどちらが自然であるかについて、言語モデルの選好と人間の選好の間にはピアソンの相関係数で 0.89 (CLM)、0.90 (SLM) の非常に高い相関があった。

### 5 実験 2: 既存の検証方法による結論との一貫性

本章および次章の概要を表 1 に示す。本章では、様々な仮説において、言語モデルを用いた検証による結論が、既存の検証方法による結論と一致したことを示す。本研究では、新情報・旧情報などの文脈に依存する要因は分析の対象外とした。既存研究によって検証されてきた 10 個の仮説全てにおいて、言語モデルを用いた検証による結論が既存研究と一致した。

#### 5.1 二重目的語の語順・副詞の位置

二重目的語の語順と副詞の位置に関する実験設定は栗林ら [18] と基本的に同様であり、全ての仮説で笹野らと一致する結論<sup>\*3</sup>が得られた。栗林ら [18] の実験では、二重目的語の語順

<sup>\*1</sup>fairseq の実装を用い、デフォルトのハイパーパラメータを使用している。

<sup>\*2</sup>sentencepiece を用い、語彙数を 100,000、character coverage を 0.9995、分割方法を bpe に設定した。

<sup>\*3</sup>栗林らとの差分は以下の点である。(i) 言語モデルのアーキテクチャを CNN から Transformer ベースに変更した。(ii) 文字ベース言語モデルとサブワードレベル言語モデルの二種類の言語モデルを用い、両者で同等の結論が示された。(iii) 有生性と語順の関係について、笹野ら [16] に従い仮説を支持する(しない)動詞の数の比率で検定したところ、笹野ら [16] と同様の結

**表1:** 各仮説に対する既存の検証方法による結論と、言語モデルを用いた検証方法による結論。太字が本論文の実験および分析で得られた結論であり、いずれも統計的な検定から得られた結論である。

対象	仮説	既存の検証方法	言語モデル
目的語 (5.1節)	動詞によらず基本語順は「にを」である [5]	棄却 [16]	<b>棄却</b> (一致)
	基本語順は動詞のタイプによって異なる [10]	棄却 [8, 11, 16]	<b>棄却</b> (一致)
	省略されにくい格は基本語順において動詞の近くに位置する [16]	支持 [16]	<b>支持</b> (一致)
	基本語順は二格名詞の意味役割や有生性によって異なる [6, 10]	支持 [16]	<b>支持</b> (一致)
副詞 (5.1節)	対象の動詞と高頻度に共起するヲ格、二格名詞は動詞の近くに位置する [16]	支持 [16]	<b>支持</b> (一致)
	副詞のタイプにより基本生起位置は異なる [9]	支持 [9]	<b>支持</b> (一致)
主語 (5.2節)	主語は時間・場所を表す格よりも後ろにくる [14, 15]	<b>支持</b>	<b>支持</b> (一致)
場所の情報 (5.2節)	場所を表す格は時を表す格よりも後ろ、主語よりも前に位置する [14, 15]	<b>支持</b>	<b>支持</b> (一致)
時の情報 (5.2節)	時間を表す格は場所、主語よりも前に位置する [14, 15]	<b>支持</b>	<b>支持</b> (一致)
一般 (5.3節)	モーラ数の大きい句が小さい句よりも前に来る	支持 [1, 13]	<b>支持</b> (一致)
主題・とりたて (6章)	とりたてられた格の位置はとりたて助詞に応じて異なる	-	<b>支持</b>
	とりたてられた格の位置はとりたてられた格に応じて異なる	-	<b>支持</b>
	基本語順で文頭側にある格ほど主題になりやすい [12]	-	<b>支持</b>
	目的語の主題になりやすさは、動詞の「をに」率に依存する	-	<b>支持</b>

**表2:** 時を表す格、場所を表す格、主語の位置関係の傾向

	$o_S$ (時, 場所)	$o_S$ (場所, 主語)	$o_S$ (時, 主語)
CLM	.757	.604	.642
SLM	.708	.615	.632
頻度	.686	.682	.666

と二格の有生性の関係に関して笹野らと異なる結論が出ていたが、検定方法を笹野ら [16] と揃えたところ、一致する結論が得られた。また、副詞の位置に関する実験結果についても、栗林ら [18] 同様、小泉ら [9] の結果と強い相関<sup>\*4</sup>が見られた。

## 5.2 時の情報、場所の情報、主語の提示される順序

佐伯 [14] は、時の情報 < 場所の情報 < 主語、の順で情報が提示されることを主張した。この仮説について、頻度ベースの検証による結論と言語モデルを用いた検証による結論が一致するか調べた。信頼できる規模でこの仮説を検証した既存研究は存在しないため<sup>\*5</sup>、本節では大規模なデータを用いた頻度ベースの検証と言語モデルによる検証の二つを行う。

30 億 Web ページからランダムに収集した 26 万用例  $S_{orig}$ <sup>\*6</sup> と、それぞれの用例について言語モデルが支持する語順に並べ替えた用例の集合  $S^{LM}$ <sup>\*7</sup> を用いた。  $S_{orig}$  を用いた検証は頻度ベースの検証、  $S^{LM}$  を用いた検証は言語モデルを

論が得られた ( $p < 0.05$ )。 (iv) 言語モデル学習データと重複する用例は除いた。

<sup>\*4</sup> 栗林らと同様、各副詞タイプについて副詞が置かれる三つの候補箇所の選好を人間の選好と比較した。3 箇所に対する選好順位の相関を各副詞タイプで求め、それらのマクロ平均は 0.88 (CLM)、0.75 (SLM) であった。

<sup>\*5</sup> 佐伯 [14] は、非常に少数の用例の観察と内省から本仮説を導いた。

<sup>\*6</sup> KNP を用いて文末の動詞が格を三つ以上もつ文を収集し、文末の動詞とその格 (統語関係で格要素に係る部分も含む) 以外の部分は削除した。

<sup>\*7</sup> 各用例について、かき混ぜ操作を適用して可能な語順を全て列挙した後、言語モデルの生成確率が最も高い語順を収集した。

**表3:** 言語モデルが基本的な格の並びから逸脱した語順を支持した際の、モーラ数の大きい句の位置の変化

モデル	前に移動した用例数	後ろに移動した用例数
CLM	7,196	5,401
SLM	7,277	5,750

用いた検証に対応する。それぞれのコーパス  $S$  を用いて、時を表す格要素<sup>\*8</sup>、場所を表す格要素<sup>\*9</sup>、主語のそれぞれのペアについて、  $o_S(a, b)$  を求めた。

$$o_S(a, b) = \frac{N_{a < b}^S}{N_{a < b}^S + N_{b < a}^S},$$

ここで、  $a, b$  は格要素、  $N_{a < b}^S$  は  $S$  において格  $a$  が格  $b$  より文頭側に位置する用例の数を表す。  $o_S(a, b)$  が  $o_S(b, a)$  よりも優位に大きければ、格  $a$  が格  $b$  より文頭側に位置する傾向がある。

いずれの実験設定においても、  $o_S$ (時, 場所)、  $o_S$ (場所, 主語)、  $o_S$ (時, 主語) が優位に大きいことが示され<sup>\*10</sup>、言語モデルによる検証と頻度による検証の結果は一致した (表2)。また、佐伯 [14] の仮説は支持された。

## 5.3 句のモーラ数と語順

長い句が短い句よりも先に現れるという傾向の存在が、既存の検証から示唆されている [1, 13]。この既存の検証結果に対する言語モデルの傾向を、言語モデルが基本語順から逸脱した語順を好んだ用例の傾向から分析する。5.2節で用いた用例のうち、言語モデルが支持する語順における各用例で最もモーラ数の大きい格要素<sup>\*11</sup>の位置が、基本的な格の並び<sup>\*12</sup>に従った語順における位置と異なった、約 12,500 用例を対象とした。

CLM、SLM が支持した語順では、もっとも長い格要素の位置が基本的な格の順序と比較してより前方に移動しているケースが優位に<sup>\*13</sup>多かった (表3)。この結果は既存の分析手法による結果と一致し、長い句が短い句よりも先に現れるという傾向が支持された。

## 6 分析: とりたて・主題と基本語順

4, 5章の結果から、言語モデルを日本語基本語順の分析に用いることの妥当性が支持された。本章では、言語モデルを用いた方法論の長所を生かした語順分析の例として、これまで大規模に分析されることがなかった、格助詞を伴わない格が存在する文における語順の自然さについて分析を行う。本分析では特に日本語のとりたてに焦点を当てる。例えば、とりたて助詞「は」を伴った要素は主題として文頭に出現するなど (例2-b)、とりたてと語順には密接な関係がある。

- (2) a. 先生が 本 を あげた。  
b. 本 は 先生が あげた。

本分析では、とりたてが生じた文における支配的な語順について、文脈非依存な要因に焦点を当てて分析をする

格要素がとりたてられると、一部の格助詞は消えてしまい、解析器を用いて元の格を高精度に同定することが困難になる。したがって、とりたてられた格の元の格を、解析器を用いて大規模コーパス中で同定したり、頻度を数えることは難しい。一方で、とりたてが生じていない文から、とりたてが生じた文を

<sup>\*8</sup> 格要素に時間と関係するカテゴリの形態素が含まれる二格。

<sup>\*9</sup> 格要素に場所と関係するカテゴリの形態素が含まれるデ格。

<sup>\*10</sup> 符号検定 ( $p < 0.05$ ) によって検証した。

<sup>\*11</sup> MeCab の-Oyomi オプションを用いた。

<sup>\*12</sup> 時 < 場所 < 主語 < 間接目的語 < 直接目的語とする。

<sup>\*13</sup> 符号検定 ( $p < 0.05$ ) によって検証した。

表4: 格要素(列)をとりたて助詞(行)でとりたてた時に, とりたてられた要素が文頭に現れた方が自然と見なされた割合

モデル	とりたて助詞	時	場所	主語	間接目的語	直接目的語	平均
CLM	は	.715	.777	.675	.624	.623	.683
	こそ	.492	.423	.521	.313	.486	.447
	も	.560	.557	.458	.343	.271	.438
	だけ	.385	.340	.312	.227	.184	.331
	平均	.538	.525	.544	.377	.391	-
SLM	は	.667	.751	.635	.565	.580	.640
	こそ	.567	.596	.574	.398	.462	.519
	も	.511	.531	.457	.292	.259	.410
	だけ	.334	.309	.285	.172	.126	.303
	平均	.520	.547	.560	.357	.357	-

作することは簡単である。例えば, ある格要素をとりたて助詞「は」でとりたてた時には, その要素を文頭に移動し, 格助詞を「は」で置き換えれば良い\*14 (例3)。

- (3) a. 本は先生が 本をあげた。  
 b. 先生は先生が 本をあげた。

従って, 4, 5章で用いたデータを編集し, とりたてと語順に関する大規模な検証用データを作成することが可能である。

### 6.1 とりたてられた格の移動しやすさ

とりたて助詞「は」でとりたてられた格要素は, 例4-bのように文頭に移動することが知られている。この傾向が他のとりたてた助詞でもみられるか, またどの格がとりたてられても同様の傾向を示すか分析した。

具体的には例4のように, とりたて助詞によってとりたてられた格が元の格の位置に存在する例と, 文頭に移動した例を作成し, どちらの語順が自然かを言語モデルを用いて分析した。

- (4) a. 先生が 本をあげた。  
 b. 本も先生があげた。

とりたて助詞として, 主題を表す「は」, 累加を表す「も」, 限定を表す「だけ」, 「こそ」の4種類を用いた。結果を表4に示す。各スコアは, とりたてられた格が前に移動した用例が移動しない用例に対して好まれた割合を示す。とりたて助詞ごとの平均スコアから, 「は」以外のとりたて助詞ではとりたてられた格が文頭に移動する傾向は弱いことが示唆された。また, 格ごとの平均スコアから, 基本的な格の並びにおいて文頭に近い側の格ほど, とりたて助詞によって文頭に移動するケースが多いことが示唆された。次節では, 語順に対する影響が強い, 主題のとりたて助詞「は」に注目して分析を行う。

### 6.2 格の主題になりやすさ

本節では, 以下の二つの仮説に焦点を当てる。

1. 基本語順で前方に位置する格ほど主題になりやすい [12]。
2. 目的語の主題になりやすさは動詞の「をに」率に依存する。

**仮説1の検証:** 5.2節と同様, 30億Webページからランダムに収集した26万用例を検証に用いた。それぞれの用例について, 例3のようにとりたて助詞「は」を用いてそれぞれの格要素をとりたてた文を作成し, それらの文を言語モデルの生成確率で比較することで, どの格要素が主題になりやすいかを調べた。言語モデルが支持した, とりたてた文の集合  $S$  を用いて, 格のペア  $(a, b)$  に対して  $t_S(a, b)$  を用いた。

$$t_S(a, b) = \frac{N_{a|b}^S}{N_{a|b}^S + N_{b|a}^S},$$

ここで  $N_{a|b}^S$  は, 用例集合  $S$  中で, 格  $a$  と格  $b$  が存在し, 格  $a$

がとりたてられている用例の数を表す。基本的な格の並びにおいて, 格  $a$  が格  $b$  より前にある全ての格のペア  $(a, b)$  について  $t_S(a, b)$  と  $t_S(b, a)$  を計算し, それらの大小を対応のとれた  $t$  検定で比較した。CLM, SLMの結果ともに,  $t_S(a, b)$  が大きいという傾向が  $p < 0.05$  で優位にみられた。従って, 基本語順で前方に位置する格ほど主題になりやすいという, 佐伯 [14] の仮説は支持された。

**仮説2の検証:** 本仮説は我々の観察に基づく仮説である。仮説1より基本語順で前方にある格要素はとりたてられやすいことが分かった。5.1節より目的語の基本的な順番は動詞ごとに異なるため, 目的語の主題になりやすさも動詞ごとに変わることが考えられる。5.1節と同様の用例を用いて, 仮説1の検証時と同様, それぞれの目的語をとりたてた用例を作成し, 言語モデルによってどちらの用例が自然かを調べた。実験の結果, 動詞ごとの前に位置しやす目的格の種類と動詞ごとのとりたてられやす目的格の種類にはピアソンの相関係数で0.89 (CLM), 0.84 (SLM) の非常に強い相関があった。

## 7 おわりに

網羅的な検証から, 日本語基本語順の分析にニューラル言語モデルを用いることの妥当性を示した。6章では, 言語モデルを用いて, これまで実現できなかったとりたてと語順に関する大規模な分析を行った。本研究では, 新情報・旧情報などの文脈に依存する要因に関する仮説は分析の対象外としたため, さらなる検証が必要とされる。今後, 本方法論の妥当性を否定する結果が示された場合, どのような条件下では本方法論が妥当に機能するかを解明していくことが, 言語モデルの性質の理解にも繋がり, 重要であると考えられる。

**謝辞** 本研究はJST CREST (課題番号: JPMJCR1513) の支援を受けて行った。日本語二重目的語構文の基本語順分析用データを提供していただいた, 笹野遼平氏に感謝の意を表す。

## 参考文献

- [1] Masayuki Asahara, Satoshi Nambu, and Shin-Ichiro Sano. "Predicting Japanese Word Order in Double Object Constructions". In: *Proceedings of the Eight Workshop on Cognitive Aspects of Computational Language Learning and Processing*. Association for Computational Linguistics, July 2018, pp. 36–40.
- [2] Jelke Bloem. "Testing the Processing Hypothesis of word order variation using a probabilistic language model". In: *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC)*. The COLING 2016 Organizing Committee, Dec. 2016, pp. 174–185.
- [3] Joan Bresnan et al. "Predicting the dative alternation". In: *Cognitive foundations of interpretation*. KNAW, 2007, pp. 69–94.
- [4] Richard Futrell and Roger P Levy. "Do RNNs learn human-like abstract word order preferences?". In: *Proceedings of the Society for Computation in Linguistics (SCL) 2019*, 2019, pp. 50–59.
- [5] Hajime Hoji. "Logical form constraints and configurational structures in Japanese." In: *PhD Thesis, University of Washington* (1985).
- [6] Atsushi Ito. "Argument structure of Japanese ditransitives". In: *Nanzan Linguistics Special 3* (2007), pp. 127–150.
- [7] Gerard Kempen and Karin Harbusch. "A corpus study into word order variation in German subordinate clauses: Animacy affects". In: *Multidisciplinary approaches to language production* (2004), pp. 173–181.
- [8] Masatoshi Koizumi and Katsuo Tamaoka. "Cognitive processing of Japanese sentences with ditransitive verbs". In: *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 2004.125 (2004), pp. 173–190.
- [9] Masatoshi Koizumi and Katsuo Tamaoka. "The Canonical Positions of Adjuncts in the Processing of Japanese Sentence". In: *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society* 13.3 (2006), pp. 392–403.
- [10] Mikinari Matsuoka. "Two Types of Ditransitive Constructions in Japanese". In: *Journal of East Asian Linguistics* 12.2 (2003), pp. 171–203.
- [11] Edson T Miyamoto. "Sources of difficulty in the processing of scrambling in Japanese". In: *Sentence processing in East Asian languages* (2002), pp. 167–188.
- [12] Hisashi Noda. *Wa to ga*. Kuroshio Publishers, 1996.
- [13] Naoh Orita. "Predicting Japanese scrambling in the wild". In: *Proceedings of the 7th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2017)*. 2017, pp. 41–45.
- [14] Tetsuo Saeki. "Gendaigo ni okeru gojun no keikō - iwayuru hogo no baai". In: *Gengo seikatsu [Language life]* 111 (Dec. 1960), pp. 56–63.
- [15] Tetsuo Saeki. *Yōsetsu Nihongo no Gajun*. Kuroshio Publishers, 1998.
- [16] Ryohei Sasano and Manabu Okumura. "A Corpus-Based Analysis of Canonical Word Order of Japanese Double Object Constructions". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Aug. 2016, pp. 2236–2244.
- [17] Yasumasa Shigenaga. "Canonical Word Order of Japanese Ditransitive Sentences: A Preliminary Investigation through a Grammaticality Judgment Survey". In: *Advances in Language and Literary Studies* 5.2 (2014), pp. 35–45.
- [18] 栗林 樹生 et al. "言語モデルを用いた日本語の語順評価と基本語順の分析". In: *言語処理学会第25回年次大会発表論文集*. 2019, pp. 1053–1056.

\*14格助詞「に」, 「で」などではとりたて助詞は格助詞の後ろに追加される。