

B-3

日本語における目的語位置での主語の記憶の想起：視線計測からの考察

江村 玲^{1,3}、磯野 真之介^{2,3}、小泉 政利¹

¹東北大学文学研究科、²東京大学総合文化研究科、³日本学術振興会特別研究員 DC

要旨

本研究の目的は、日本語母語話者が日本語他動詞文を読む際にいつ主語の意味役割を付与するのかを明らかにすることである。先行研究では、意味役割付与のために動詞で主語が想起されると考えられ、動詞位置での処理を分析対象としていた。しかし本研究は、主語と動詞の間に置かれる目的語の時点で、主語が想起され、主語の意味役割が付与されているという仮説を検証した。日本語の SOV 構文を用いた視線計測実験において、目的語位置と動詞位置での主語への読み戻り率とその読み時間を計測した。結果、目的語位置で 60% を超える主語への読み戻りが見られた。さらに、目的語に基づいた動詞の予測可能性が低い条件では、予測可能性が高い条件に比べ、動詞から主語へ読み戻る率が上がった。これらの結果は、目的語位置で具体的な動詞を予測しており、目的語位置で主語の意味役割が与えられるために主語が想起されることを示唆する。

1. はじめに

人間は、文脈に基づく予測処理と語の入力に基づく処理の両方に基づいて文を理解している。本研究は、日本語における主語－動詞依存関係に焦点を当て、これらの二つの処理のインターフェースを議論する。具体的には、目的語および動詞からの主語への読み戻りのデータから、日本語母語話者は予測処理によって目的語位置で主語をワーキングメモリから想起し、主語の意味役割を付与するという仮説を支持する。

読み手が主語－動詞依存関係をどのように結んでいるのか、という問題は文処理研究の大きなテーマの一つである。文処理モデルの一つであるキューに基づく想起モデルのもとでは、動詞位置で意味役割付与のために主語の情報がワーキングメモリから想起される (Lewis & Vasishth, 2005; Van Dyke & McElree, 2006)。動詞位置での主語の想起があるという主張は、主語－動詞間の距離が伸びた際の動詞の読み時間の増大 (局所性効果) や、主語に似た名詞が文中にある場合の動詞の読み時間の変化 (類似性に基づく干渉効果) を証拠として検証されてきた (Grodner & Gibson, 2005; Van Dyke & Johns, 2012; Van Dyke & McElree, 2006 等)。

しかし、日本語では、動詞位置での主語の想起の証拠となる局所性効果が見られにくい (Nakatani & Gibson, 2010; Yano, Minemi, & Tsumura, 2016)。たとえば、Nakatani and Gibson では、(1) のように主語－動詞距離が長い条件と短い条件で、動詞「報告した」の読み時間に有意な差が見られなかった。

- (1) a. **書記**が [代議士が [首相がうたた寝したと] 抗議したと] 報告した。
b. [代議士が [首相がうたた寝したと] 抗議したと] **書記**が報告した。

(Nakatani & Gibson, 2010)

他の SOV 言語でも局所性効果が見られにくいとされている (ドイツ語: Konieczny, 2000; ヒンディー語: Vasishth, 2003; Vasishth & Lewis, 2006)。これに対し、Isono and Hirose (in press) は、日本語では動詞より前で局所性効果が見られることを実証し、動詞が知覚される前にも主語が想起されているため、動詞位置では主

語の減衰による局所性効果が生じないと考察した。そこで、Isono and Hirose は場所句での主語の想起をテストしたのに対し、本研究は目的語位置での主語の想起に関する検証を行なった。というのは、日本語では、目的語の時点で動詞の意味や統語に関して予測をしていることが先行研究で示されている (Kamide, Scheepers & Altmann, 2003; Momma, 2016)。このことから、具体的な動詞を目的語位置で予測しており、主語の意味役割付与が目的語位置で行われていると仮説を立てた。本研究では、視線計測を用いてこの仮説を検討した。Isono and Hirose は、局所性効果を主語の想起の証拠として用いたが、本研究は視線計測で読み戻りを計測することで、より直接的に主語の想起の有無を確認した。

この仮説から、以下の二つの予測を立てられる (図 1 参照)。まず、主語の意味役割が目的語位置で付与されているならば、目的語位置で主語が想起される。目的語位置で主語が想起されるならば、目的語位置から主語位置への読み戻りが起こる可能性が高くなると予測される。第二に、目的語位置での意味役割付与の動機が、目的語位置ですでに具体的な動詞の予測をしていることであれば、予測された動詞と実際の動詞が一致しているか否かによって、動詞位置での処理が異なるはずである。具体的には、目的語位置で予測された動詞と入力された動詞が一致した場合は、動詞位置で主語が想起される可能性が低くなる。つまり、動詞位置での主語への読み戻りが少なくなると予測する。反対に、動詞の予測が外れた場合は、意味役割の修正を行うために主語が想起される可能性が高くなると予測する。よって、本研究では、動詞の予測可能性も操作し、動詞位置での主語の読み戻りを測定した。

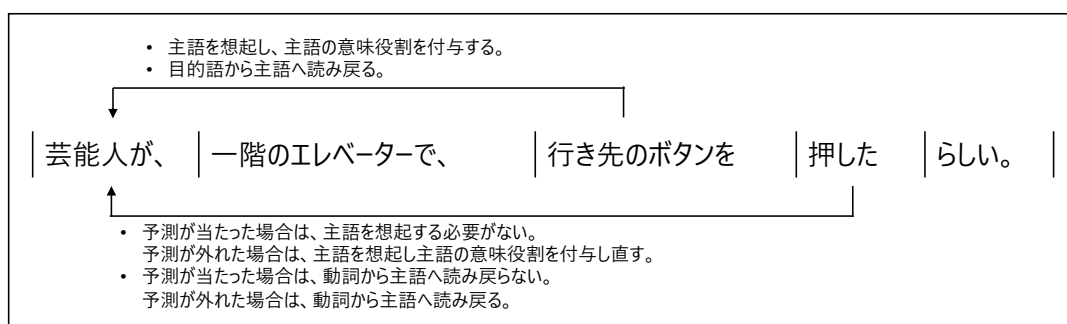


図 1 仮説と予測の概要図

2. 方法

2.1. 被験者

日本語語母語話者 15 名が視線計測実験に参加し、内容理解問題の正答率が 70%以上の 13 名 (男性 9 名; 19-24 歳; 平均年齢 20.6 歳) を分析対象とした。被験者の全員は、利き手が右であり、視力に問題がなく、言語障害の診断歴もない。被験者の利き目は、11 名が右であり 2 名が左である。本研究は、事前に東北大学文学研究科倫理委員会の承認を受けた。また、実験前に被験者全員が同意書に署名した。

2.2. 刺激

各アイテムは、(2) のような文脈文と (3) のような SOV 語順の他動詞文のターゲット文からなる。文脈では、ターゲット文の主語となる人物の状況について述べられている。ターゲット文は目的語に基づいた動詞の予測可能性の高さを操作した。予測可能性の高さは、別途実施した、動詞を空所とした空所補充課題によって決定した。回答の 7 割以上を占める動詞を高予測条件、回答に無い動詞を低予測条件とした。

(2) 文脈

ある芸能人は、テレビ局に撮影に来ていた。

(3) 高予測条件／低予測条件

領域 1 (主語)	領域 2 (場所句)	領域 3 (目的語)	領域 4 (動詞)
芸能人が、	一階のエレベーターで、	行き先のボタンを	押した／壊した
領域 5 (助動詞)			
らしい。			

刺激文を 72 アイテム作り、ラテン方格法に基づいて 2 リストに分けた。つまり、それぞれのリストは 36 アイテムを含む。また、意味的にも文法的にも自然である 36 アイテムをフィラーとして加えた。フィラー文は他の実験のターゲット文となるため、全てのリストが同じフィラー文を含んでいるわけではない。

2.3. 手順

視線計測には EyeLink 1000 Plus デスクトップマウント (SR Research 社) を用いた。参加者は両目で画面を見て、右目の視線が 1000 Hz のサンプリング・レートで記録された。刺激提示には Experiment Builder (SR Research 社) および、24.5 インチの液晶ディスプレイ (解像度 1920×1080; リフレッシュ・レート 165 Hz) を用いた。参加者はディスプレイから 92 cm の位置に置かれたあご台にあごを置き、刺激を読んだ。文字は 33 ポイントの MS 明朝で提示された。刺激は 1 列で、上下中央に提示された。

参加者は、7 アイテムの練習を行ったあと、本番に入った。1 試行の手続きは以下である。画面の左端に固視点が現れ、ドリフトチェックを行う。その後、固視点到代わり文脈文が 3000 ミリ秒、続いて固視点が 2000 ミリ秒提示される。そして、ターゲット文が 1 文で提示され、自己ペースでターゲット文を読み、Cedrus Response Pad RB-740 7 番キーを右手で押す。最後に、内容理解問題が現れ、内容理解問題が正しい場合は左手で 2 番キー、間違っている場合は左手で 1 番キーを押す。内容理解問題は幅広い内容を出した。

2.4. 分析方法

本実験は、目的語および動詞からの主語への読み戻りの割合と、目的語および動詞から主語へ読み戻った際の主語の読み時間に焦点を当てた。どちらの分析においても、内容理解問題に正解した試行のみを分析対象とした。また、どちらも R (R Core Team, 2022) の *lme4* パッケージ (Bates, Mächler, Bolker, & Walker, 2015) を用いて、線形混合モデルを用いて統計的に分析した (Baayen, 2008)。最終モデルは最尤推定法が適合されたもので、*p* 値は *lmerTest* パッケージ (Kuznetsova, Brockhoff, & Christensen, 2017) を用いて計算した。

はじめに、読み戻り率の統計分析には、二項分布を確率モデルとする一般化線形混合効果モデルを用いた。モデルは従属変数として、主語への読み戻りの有無を含み、固定要因として予測可能性の高さ (高予測／低予測)、主語への読み戻りの起点 (目的語／動詞)、それらの交互作用を含む。また、ランダム要因としてアイテムごとのランダム切片が含まれる。なお、被験者ごとのランダム切片を含めたモデルは収束しなかった。その後、事後検定として、*emmeans* パッケージ (Lenth, 2023) を用いて、読み戻りの起点の条件ごとに、主語への読み戻り確率の推定周辺平均を求めた。読み戻りの起点の条件それぞれにおいて、予測可能性の高さにより主語への読み戻り確率が変化するのかを *z* 検定により確かめた。

次に、主語の再読時間を分析について、主語と目的語を注視した後かつ動詞を注視する前の主語の注視時間（以降、目的語を起点とする主語の再読時間と呼ぶ）と、主語と目的語と動詞を注視して以降の主語の注視時間（以降、動詞を起点とする主語の再読時間と呼ぶ）を比較した。まず、主語の再読時間がゼロの場合は分析から除外した。統計分析には、線形混合効果モデルを用いた。モデルの従属変数には、主語の再読時間を設定し、固定効果として、予測可能性の高さ（高予測／低予測）と再読の起点（目的語／動詞）、そしてそれらの交互作用をモデルに組み込んだ。また、ランダム切片として被験者とアイテムを設定した。その後、事後検定として、再読の起点の条件ごとに、主語の再読時間の推定周辺平均を求めた。読み戻りの起点の条件それぞれにおいて、主語の読み時間が予測可能性の高さによって変化するのかが t 検定により確かめた。

3. 結果

まず、主語への読み戻り率について、各条件における主語への読み戻り率は、図2と表1の通りである。表1より、主語への読み戻りの起点の主効果が有意であり、動詞からの読み戻り率が目的語からの読み戻り率よりも高くなった ($\beta=0.98, p<0.001$)。また、交互作用が有意であった ($\beta=0.79, p=0.020$)。目的語からの読み戻りでは予測可能性の有意な差は見られなかったが ($z=0.45, p=0.656$)、動詞からの読み戻りでは低予測条件は高予測条件よりも高い値を示した ($z=-2.55, p=0.011$)。

表1 主語への読み戻り率のモデルにおける固定効果の推定

	β	標準誤差	z	p	
切片	0.58	0.18	3.29	0.001	***
予測可能性の高さ	-0.09	0.20	-0.45	0.656	
主語への読み戻りの起点	0.98	0.22	4.41	<0.001	***
予測可能性の高さ×主語への読み戻りの起点	0.79	0.34	2.33	0.020	*

注: * $p<0.05$. ** $p<0.01$. *** $p<0.001$.

次に、主語の再読時間について、結果は図3と表2の通りである。表2にある通り、再読の起点の主効果が有意であり、予測可能性の高さにかかわらず、目的語を起点とする主語の再読時間が、動詞を起点とする再読時間に比べて長くなった ($\beta=-1518.37, p<0.001$)。交互作用に有意な結果は見られなかった ($\beta=363.80, p=0.153$)。また、再読の起点が目的語である場合も動詞である場合も、予測可能性の高さに再読時間の有意な差は見られなかった（目的語: $t=1.38, p=0.171$; 動詞: $t=-0.48, p=0.631$ ）。

表2 主語の再読時間のモデルにおける固定効果の推定

	β	標準誤差	自由度	t	p	
切片	2245.00	149.44	48.18	15.02	<0.001	***
予測可能性の高さ	-268.46	185.36	61.77	-1.45	0.153	
再読の起点	-1518.37	200.30	53.21	-7.58	<0.001	***
予測可能性の高さ×再読の起点	363.80	250.79	53.08	1.45	0.153	

注: * $p<0.05$. ** $p<0.01$. *** $p<0.001$.

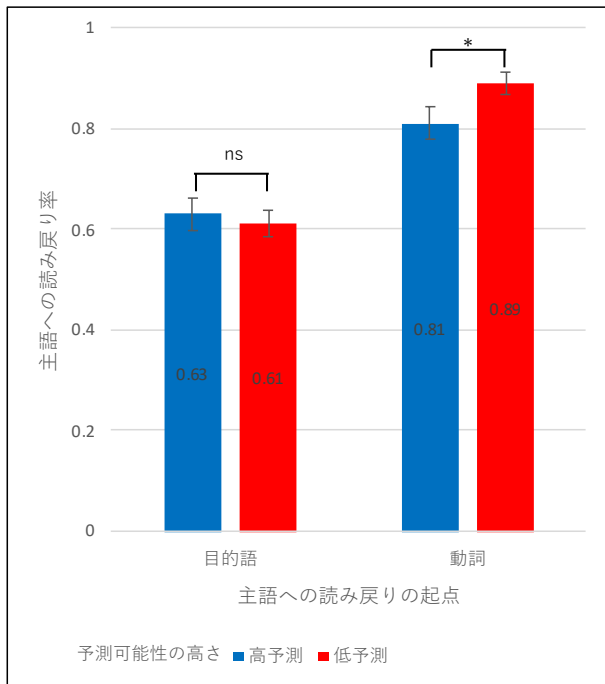


図 2 各条件の主語への読み戻り率の平均値
 注: ns = not significant. * $p < 0.05$. ラベルは平均値を示す。エラーバーは標準誤差を示す。

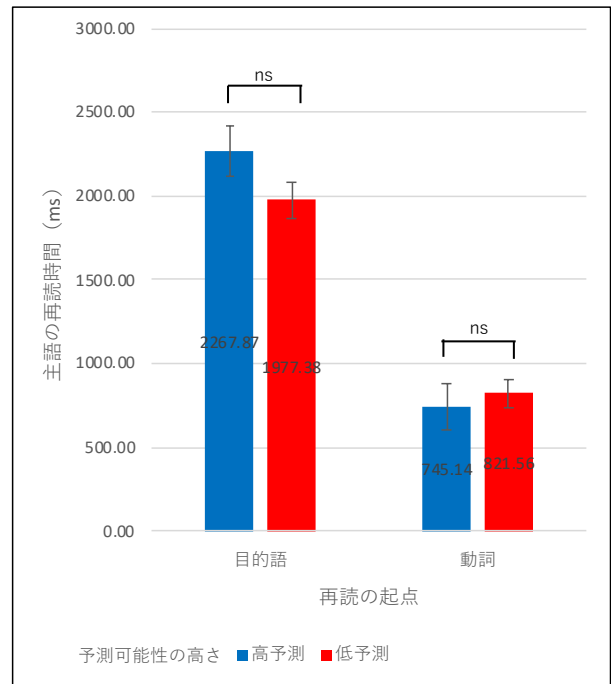


図 3 各条件の主語の再読時間の平均値
 注: ns = not significant. ラベルは平均値を示す。エラーバーは標準誤差を示す。

4. 考察

4.1. 結果の解釈

本研究は、日本語のような SOV 言語では、予測処理により主語の意味役割が目的語位置で付与されるという仮説を、視線計測から検証した。まず、目的語位置での 60% を超える読み戻り率が確認された。これは、目的語位置で主語を想起する処理を行っていることを示唆する。さらに、主語への読み戻りが起こった時、目的語から読み戻った時の主語の再読時間の方が、動詞から読み戻った時の主語の再読時間に比べて有意に長かった。このことは、目的語位置で主語に関するなんらかの処理を行なっていることを示唆し、意味役割付与はそのような処理の一つとして考えられる。

また、目的語に基づいた動詞の予測可能性が低い条件では、予測可能性が高い条件に比べ、動詞から主語への読み戻り率が上がった。この結果は、目的語位置で主語を想起する動機が、目的語位置での動詞の予測可能性の高まりであるという仮説を支持する。つまり、目的語位置で具体的な動詞を予測しており、目的語位置で主語の意味役割付与がすでに行われていると考えられる。よって、その予測した動詞が実際の動詞と異なった場合に、主語に意味役割を付与し直す必要があるため、再び主語の想起を行う (図 1 参照)。

ただし、動詞で主語が想起されないとは言えない。というのは、動詞から主語への読み戻り率は、高予測条件においても、目的語からの読み戻り率よりも高かった。よって、動詞位置での主語の意味役割付与を否定することはできない。しかし、動詞で主語の想起が見られることは、目的語位置で主語を想起することと矛盾するわけではないことを補足する。というのは、高予測条件でも予測が外れている可能性があるため、目的語位置と動詞位置の両方で主語が想起されることは十分に考えられる。さらに、動詞から主語への読み戻りは、文末に見られる文全体をまとめる効果 (wrap-up 効果) の可能性もあるため、更なる検討が必要である。

4.2. 日本語における主語－動詞依存関係の構築

キューに基づく想起モデルを始めとする主語－動詞依存関係の理解処理についての研究では、動詞が知覚された時の処理について議論されることが多い (cf. Vasishth & Engelmann, 2021, Chapter 2)。しかし、動詞の処理に注目した実験では、動詞で主語を想起する証拠となる局所性効果が SOV 言語では観察されないことがある。そこで、本研究は、動詞の前、具体的には目的語における主語－動詞依存関係に焦点を当て、日本語における目的語位置での主語の想起を、初めて視線計測から実証した。

SOV 言語で局所性効果が見られにくい理由として、動詞での処理だけではなく、動詞の前の要素による予測処理が影響していることは先行研究で示されている (Levy & Keller, 2013)。彼らは、動詞の前に追加の要素を加えることによって動詞の予測可能性を高め、結果として動詞の処理が促進されたと考察した。しかしあくまで動詞位置の処理を分析するだけであった。一方で、本研究は、動詞の前の目的語での処理を分析対象とし、主語の意味役割をすでに動詞前で付与しているという仮説を支持した。

本研究と同じように、Isono and Hirose (in press) も、主語が入力されてから動詞が入力される前までの処理を分析し、動詞よりも前に主語が想起されることを示唆した。この仮説と本研究のデータは一致する。彼らは移動窓式自己ペース読文課題を用いて、(4) のような主語と、主節の残りの要素間の距離を操作した他動詞文を、日本語母語話者に読ませた。結果、(4a)の方が、(4b)よりも、主節の場所格名詞である「前で」やその後ろの領域で、読み時間が長くなった。

(4) a. **外科医が**、病院が一時的に雇った看護師の前で小柄の患者を優しく撫でたところ…

b. 病院が一時的に雇った**外科医が**、看護師の前で小柄の患者を優しく撫でたところ…

(Isono & Hirose, in press)

ただし本研究と異なり、Isono and Hirose (in press) の研究では、主語の想起の証拠として示した局所性効果は、目的語の前の場所句で見られ、目的語位置では見られなかった。このことから、主語が想起されるのは目的語に限定されないとうかがえる。

本研究の新しい点として、Isono and Hirose (in press) の刺激では、予測可能性の高さを操作していなかった。一方、本研究は動詞の予測可能性の高さを操作することで、目的語位置での動詞に対する予測可能性が高い場合は、動詞で主語に読み戻す率が高くなることを示した。したがって、本研究は、目的語位置で具体的な動詞が予測されることによって、主語が想起されるという新たな仮説を提示し、支持した。

4.3. 研究の限界と今後の展望

本研究は、日本語母語話者は、目的語位置で主語を想起し、主語に関するなんらかの処理を行なっていることを示した。しかし、この目的語位置での主語の想起が意味役割付与によるものかは更なる検討が必要である。

また、本研究は目的語位置で主語への読み戻りが約 60%で起こったことを主語への想起があることの証拠としているが、この 60%という値が高いと言えるかは、読み戻りのうちどれだけが目的語であることに起因するかで評価されなければならない。しかし、今回の実験では直接評価できないため、この点は今後の課題である。

5. 結論

日本語母語話者が日本語の SOV 語順の他動詞文を読んでいるときの視線を計測し、目的語および動詞からの主語への読み戻りについて分析した。目的語位置から 60%以上の割合で主語に読み戻りがあることから、読み手は目的語位置で主語を想起していると示唆された。また、目的語位置での動詞に対する予測可能性が低い条件では、高い条件に比べて、動詞から主語への読み戻り率が高くなった。これらのデータは、日本語母語話者が動詞を知覚する前に主語に意味役割を付与し、主語－動詞依存関係を結ぶという仮説を支持する。

引用文献

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive science*, 29(2), 261-290.
- Isono, S. & Hirose, Y. (in press). Pre-verb reactivation of arguments in sentence processing. *Glossa Psycholinguistics*.
- Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of psycholinguistic research*, 32, 37-55.
- Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research*, 29(6), 627-645.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1-26.
- Lenth, R (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means (R package version 1.8.8).
- Levy, R. P., & Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of memory and language*, 68(2), 199-222.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375-419.
- Momma, S. (2016). *Parsing, generation, and grammar* [Doctoral dissertation, University of Maryland College Park].
- Nakatani, K., & Gibson, E. (2010). An on-line study of Japanese nesting complexity. *Cognitive Science*, 34(1), 94-112.
- R Core Team (2022). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, 6(4), 193-211.
- Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of memory and language*, 55(2), 157-166.
- Vasishth, S. (2003). *Working memory in sentence comprehension: Processing Hindi center embeddings*. London: Routledge.
- Vasishth, S., & Engelmann, F. (2021). *Sentence comprehension as a cognitive process: A computational approach*. Cambridge University Press.
- Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4), 767-794.
- Yano, M., Minemi, I., & Tsumura, S. (2016). Processing benefits from similarity-based partial reactivation: A case study of subject-predicate dependency in Japanese. *Kyushu University Papers in Linguistics*, 36, 309-329.

謝辞

本研究のデータ収集は、矢野雅貴准教授（東京都立大学）に協力いただいた。ここに感謝の意を表す。なお、本研究は JSPS 科研費 19H05589 および特別研究員奨励費 23KJ0199 の助成を受けた。