



梶川康平<sup>1,2</sup> 磯野真之介<sup>1,2</sup> 窪田悠介<sup>2</sup> 大関洋平<sup>1</sup>



1東京大学 2国立国語研究所

kohei.kajikawa1223@gmail.com



🤔 言語横断的な「規則」はなぜ存在するのか？ (←言語学・認知科学のBig Question)

💡 言語は「使いやすい」ように形作られているから！？

😊 語順が自由ならば省略できがち (逆も然り) という言語横断的特徴の存在を、「使いやすさ」の観点から説明

## 背景：自由語順と項省略の相関関係

- 自由語順と項省略の存在には相関関係がある [Hale, 1980; Oku, 1998; Saito, 2002; Takahashi, 2008]
- こうした相関はなぜ存在するのか？
  - 領域固有のパラメータ？
  - それとも、領域一般の認知能力から説明可能？

### 自由語順 (かき混ぜ)

花子が 太郎を ほめた (主語—目的語—動詞)  
 太郎を 花子が ほめた (目的語—主語—動詞)

### 項省略

~~花子が~~ 太郎を ほめた (主語の省略)  
 花子が ~~太郎を~~ ほめた (目的語の省略)

## 実験：文法相関がある/ない日本語を作ってみる

	項省略あり	項省略なし
自由語順あり	Id Swap	Id <sub>no-omission</sub> Swap <sub>no-omission</sub>
自由語順なし	Subj>Obj Obj>Subj	Subj>Obj <sub>no-omission</sub> Obj>Subj <sub>no-omission</sub>

- UD\_Japanese-BCCWJ (v2.10, 長単位分割) [Asahara et al., 2018] と JAOJ [Ishizuki et al., 2024] を用いて、
  - (i) 依存構造と、(ii) どこでどのような項が省略されているかの情報を獲得
- 3,428文、67,088単語+2,895単語分の項が省略
- Id: 語順をそのまま、Subj>Obj: 語順をS—Oに固定、Obj>Subj: 語順をO—Sに固定、Swap: S—OとO—Sの関係を逆転、に加え、項省略あり/なしの**仮想日本語**コーパスを作成
- ベースラインとして、語順規則をランダムに再構築した文法×500を設計し、項省略あり/なしの合わせて1,000の言語を作成 (仮想言語群の妥当な分布を推定するため)
- 記憶および予測の処理負荷を、平均依存関係長と平均5-gramサブライザルで測る

## 仮説：記憶と予測の処理負荷低減が文法相関を形作る？

- 記憶と予測が文の処理負荷の要因 [Gibson, 2000; Lewis & Vasishth, 2005; Isono, 2024; Hale, 2001; Levy, 2008; Futrell et al., 2020]

### 記憶に関する理論：Dependency Locality Theory [Gibson, 2000]

- 依存関係の距離が遠いほど、作業記憶での手続きで負荷が生じる
- “long-before-short”語順 [Hawkins, 1994; Yamashita & Chang, 2001]の選好も説明可能：



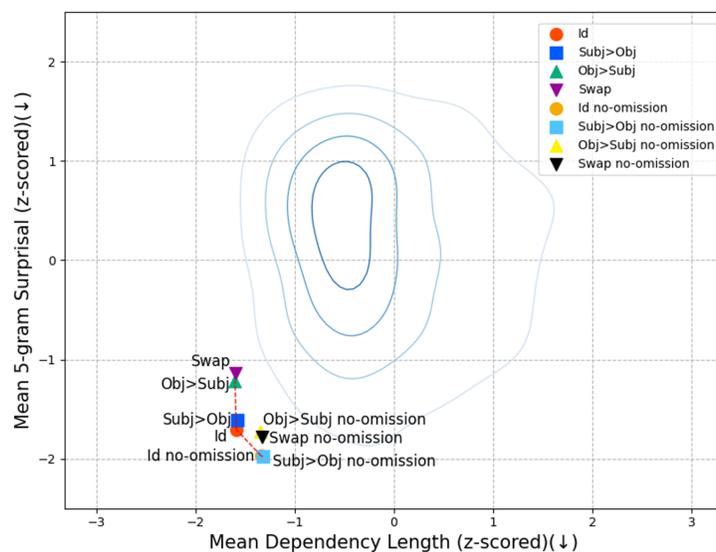
### 予測に関する理論：Surprisal Theory [Hale, 2001; Levy, 2008]

- 単語の処理負荷は、その予測のしにくさに比例する：  
 $difficulty(w) \propto -\log P(w | Context)$
- 項省略は、後続動詞のsurprisalが大きくなるように行われる [石月ら, 2022]

- 記憶と予測には**トレードオフ**関係がある [Futrell et al., 2020; Hahn et al., 2021]
- 自然言語の語順は、このトレードオフのもとで、効率的なように形作られている [Gildea & Jaeger, 2015; Hahn et al., 2021; Hahn & Xu, 2022]
- 自由語順と項省略は、記憶と予測の処理負荷に大きな影響を与える

➡ 自由語順と項省略の相関関係は、記憶と予測のトレードオフにより説明できる？

## 結果：文法相関は処理負荷低減の最適解



記憶と予測における処理負荷の分布。青い分布は、ベースライン言語での処理負荷の推定値 (カーネル密度推定) を表す

最も効率的な言語L*	λの範囲
Subj>Obj no-omission	[0.00, 0.11]
Id no-omission	[0.12, 0.52]
<b>Id</b>	<b>[0.53, 0.97]</b>
Obj>Subj	[0.98, 1.00]

$$\Omega_L(\lambda) := \lambda \cdot \text{deplen}_L + (1 - \lambda) \cdot \text{surp}_L, \quad (0 \leq \lambda \leq 1)$$

$$L^* = \arg \min_{L \in \mathcal{L}} \Omega_L(\lambda)$$

- 「現行日本語」は、**記憶**の処理負荷を減らす方を重視した最適化の結果
- 予測を減らす方を重視した場合、項省略なし・語順固定の方向へ
- LLMのような「予測偏重」型モデルが普及すると、項省略はなくなる！？固定語順へ！？

【参考文献】 [1] Hale. *MIT Working Papers in Linguistics*, 1980. [2] Oku. *CLS*, 1998. [3] Saito. *NELS*, 2002. [4] Takahashi. *The Oxford Handbook of Japanese Linguistics*, 2008. [5] Gibson. *Images, Language, Brain*, 2000. [6] Lewis & Vasishth. *CogSci*, 2005. [7] Isono. *Cognition*, 2024. [8] Hale. *NAACL*, 2001. [9] Levy. *Cognition*, 2008. [10] Futrell et al. *CogSci*, 2020. [11] Hawkins. *A Performance Theory of Order and Constituency*, 1994. [12] Yamashita & Chang. *Cognition*, 2001. [13] 石月ら. 言語処理学会, 2022. [14] Gildea & Jaeger. *arXiv*, 2015. [15] Hahn et al. *PsychoRev*, 2021. [16] Hahn & Xu. *PNAS*, 2022. [17] Asahara et al. *LREC*, 2018. [18] Ishizuki et al. *LREC-COLING*, 2024.