

# 翻訳者の行動パターンを再現する能動的推論モデルの構築

溝脇孝哲<sup>1</sup> 山田優<sup>1</sup> Michael Carl<sup>2</sup> Aishvarya Raj<sup>3</sup>

Devi Sri Bandaru<sup>2</sup> Xinyue Ren<sup>2</sup> Yuxiang Wei<sup>4</sup>

<sup>1</sup>立教大学 <sup>2</sup>Kent State University <sup>3</sup>London Interdisciplinary School <sup>4</sup>Saint Francis University

{22wv006d, masaru.yamada}@rikkyo.ac.jp {mcarl6, dbandaru, xren8}@kent.edu

aishvarya.raj@lis.ac.uk yuxiang.wei@link.cuhk.edu.hk

## 概要

翻訳は、原文を読みながら訳文を入力する同時並行的な作業であるが、読み方と入力の進め方には個人差（スタイル）がある。近年、視線移動とキー入力の記録から翻訳者の行動の特徴が記述されてきた一方で、翻訳者の行動やスタイルの違いを「シミュレーションして再現する」枠組みは十分ではない。そこで本稿は、視線移動とキー入力を行動として選択できる生成モデルを構築し、実データから推定した感覚運動レベルの行動の傾向を組み込んだ。そのモデルを使い、行動の好みに関するパラメータを変えてシミュレーションを行った。その結果、翻訳開始直後に現れる読みパターンの違いが、実データに近い形で再現されることを示した。

## 1 はじめに

翻訳は、原文理解と訳文入力と同時に進む複雑な認知活動である。近年、CRITT Translation Process Research Database (TPR-DB) [2]に代表される翻訳プロセスデータベースの整備により、視線移動とキー入力を同期させた翻訳者の行動ログが大量に蓄積されてきた。これらのデータは、翻訳者の認知負荷を分析するだけでなく、作業方略の個人差や翻訳スタイルを定量的に捉える手がかりを与える。典型的な翻訳スタイルの例として、図1のような Head-starter と Context-planner という2つのタイプ[3]がある。Head-starter とは、原文を読むのとほぼ同時に訳文の産出を行うスタイルである（図1上）。一方、Context-planner は、まず文脈を読んだうえで訳出を始めるスタイルである（図1下）。

既存研究の多くでは、このようなスタイルや行動パターンの分類・記述が観測データの特徴量にもとづき体系的に行われてきたが、翻訳者の行動をシミュレーションして、スタイルの違いを再現・検証できるモデルはまだ少ない。それに関連して近年では、能動的推論[1]を翻訳者の認知プロセスの説明に導入しようとする試みが進んでいる。能動的推論は、知覚と行動選択を同一のベイズ推論として扱い、生物は観測と予測の不一致（驚き）を低減するように行動を選択するという一般理論を与える。翻訳研究では、感覚運動・認知・情動などの階層をモデル化し、翻訳中の行動を選ぶエージェントとして翻訳者を捉える提案などがなされてきた[4,5]。

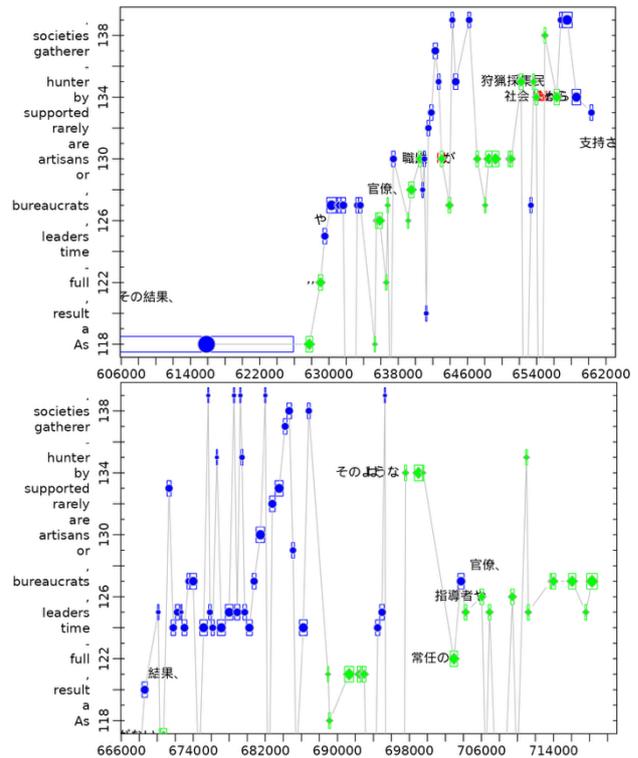


図1 翻訳者の行動を可視化したグラフ（上：Head-starter、下：Context-planner）。X軸は時間（ミリ秒）、Y軸は原文の単語。点はfixation、線はsaccade、黒文字は訳文の入力、赤文字は削除を表す。

シミュレーションして、スタイルの違いを再現・検証できるモデルはまだ少ない。それに関連して近年では、能動的推論[1]を翻訳者の認知プロセスの説明に導入しようとする試みが進んでいる。能動的推論は、知覚と行動選択を同一のベイズ推論として扱い、生物は観測と予測の不一致（驚き）を低減するように行動を選択するという一般理論を与える。翻訳研究では、感覚運動・認知・情動などの階層をモデル化し、翻訳中の行動を選ぶエージェントとして翻訳者を捉える提案などがなされてきた[4,5]。

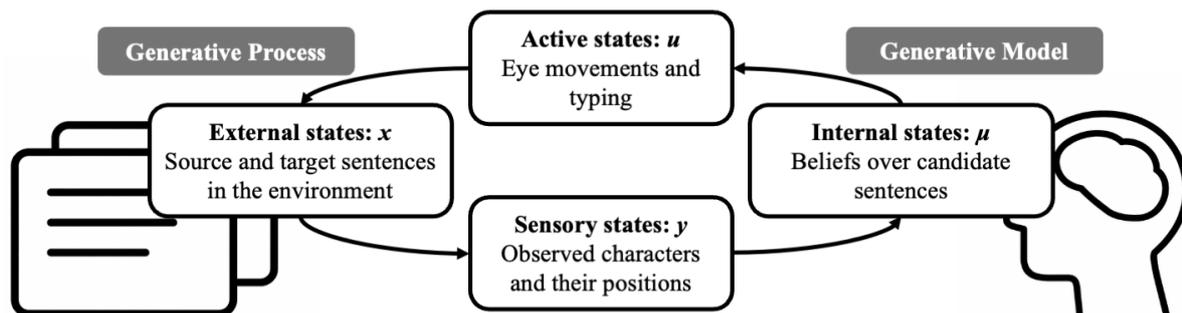


図 2 能動的推論にもとづき[8]翻訳プロセスをシミュレーションするモデルの概念図

本稿では、能動的推論にもとづき、視線移動とキー入力の協調をシミュレーション可能な形でモデル化し、翻訳スタイルの違いを再現する方法を検討する。リサーチクエスチョン (RQ) は次の通り。RQ1: 能動的推論の枠組みで、翻訳プロセスはどのようにモデル化できるか。RQ2: モデルによるシミュレーションで、翻訳スタイルの特徴を再現できるか。

なお、本稿のモデルは機械翻訳のような訳文生成そのものが目的ではなく、翻訳スタイルの違いとして原文と訳文の間のプロセスを再現することを目指す。以降、2章でモデルの概要と構築方法を整理し、3章で結果を報告する。

## 2 モデル

### 2.1 モデルの概観

本研究のモデルでは、翻訳プロセスを2階層の部分観測マルコフ決定過程として記述する[6,7]。上位層は比較的遅い時間スケールで、「原文センテンスと読んでいるチャンクの位置」および「訳文センテンスと入力しているチャンクの位置」を追跡し、下位層はより速い時間スケールで、「原文チャンクと視線の語位置」および「訳文チャンクとカーソルの語位置」を更新する。これらは隠れ状態 (図 2: External states と Internal states) となり、語またはチャンクの観測 (図 2: Sensory states) にもとづきベイズ的に状態を推論する。下位層の推論結果は上位層へのボトムアップな観測になり、上位層の予測は下位層のトップダウンな事前分布になる。推論した状態にもとづき、エージェントは行動を選択する。行動は視線移動とキー入力の2要因 (図 2: Active states) で、下位層では感覚運動レベルの視線移動タイプとキー入力タイプ[9]の組み合わせから選択する。上位層では3つの心的状態タイプ (Flow / Orientation / Hesitation) [10]を行動の候補とする。Flow (F) とは、

先読みが最小限で、滑らかに訳文を入力していく状態である。Orientation (O) とは、訳し始める前 (または途中) に、原文を先読みして内容を把握し、計画を立てている状態である。そして Hesitation (H) は、予想と入力合わないなどの「驚き」で訳出の流れが中断され、読み返しや修正が増える状態を指す。各層でこれらの行動の候補について期待自由エネルギー (G) が算出され、G が小さい行動ほど選ばれやすくなる。

### 2.2 パラメータの推定

本モデルのポイントは、行動に応じた隠れ状態の遷移確率の一部を、プロセスデータから推定した分布で実装する点にある。これにより、現実の翻訳行動に近いダイナミクスをモデルに埋め込める。分布の推定には、TPR-DB に含まれる英日翻訳タスク (ENJA15[11]) の人手翻訳セッションを用いる。1セッションには、6、7文程度の原文を翻訳した際の視線・キー入力情報が含まれており、合計で (図 1 で示した2セッションを除いた) 23セッションを使用した。このデータから、行動のラベルごとに「次の視線位置またはカーソル位置がどれだけ進む/戻るか」を表す差分の分布を推定し、それを離散状態空間上の遷移確率に変換した。例えば視線については、連続する fixation 間の語位置の差  $\Delta w$  を集計し、視線移動タイプごとにその確率を推定した。キー入力についても同様に、カーソルの移動量  $\Delta u$  を集計し、確率分布を推定した。

### 2.3 シミュレーションの条件設定

翻訳スタイルは、同じ翻訳環境でも「どの行動を好むか」という事前バイアスとして現れると考えられる。そこで本稿では、スタイルの違いを行動の事前分布として操作する。本稿では以下の2条件を設定して、シミュレーション結果を比較する。

- 条件① 常に Flow (F) 行動を好む。
- 条件② 好み時間が時間に応じて Orientation (O) → Hesitation (H) → Flow (F) へと変化する。

### 3 シミュレーションの結果

#### 3.1 可視化した行動の比較

条件① (Flow 固定) でのシミュレーションの一例を可視化したものが図3 上部である。視線 (青い点と線) は原文をほぼ直線的に移動し、訳文の入力 (黒文字) も比較的早い時間から同時に行われている。この条件では、上位層が Flow 行動を好むため、文脈の先読みをせずに、読み進めながら並行的に訳文を入力する行動が選択されやすくなっている。このような行動パターンは、実データの Head-starter スタイル (図1 上) と類似している。

これに対して、条件② (O→H→F) の例が図3 下部のグラフである。視線は、まず比較的素早く後半部分まで移動した後で、前半部分に戻る。その後、原文を再度順に読み進めつつも、訳文入力が進んでいる。この条件では、開始直後は上位層で文脈を把握する行動が支持され、その後徐々に Flow 的な行動へと移っている。このような「先に読む→入力する」という時間構造が生成され、Context-planner の特徴 (図1 下) と整合的な行動パターンが得られた。

#### 3.2 量的な比較

3.1 節では、条件の違いによって Head-starter と Context-planner 的な2つのスタイルが現れることを視覚的に確認したが、ここではその差を視線パターンの比率で量的に比較する。

具体的には、「翻訳開始から訳語『その結果』の次の語を入力するまで」に含まれる fixation を、3種類の視線パターン (L: Linear / R: Re-fixation / S: Scattered) に分類[9]し、各タイプの比率を計算する。L は視線がほぼ一直線に進む動き、R は同じあたりを繰り返し注視する動き、S は読む流れがまとまらない不規則な動きである。ここで  $n_L$ 、 $n_R$ 、 $n_S$  をそれぞれ視線パターンの出現回数、 $n_{fix}=n_L+n_R+n_S$  を fixation の総数とすると、比率  $p_L$ 、 $p_R$ 、 $p_S$  は  $p_L=n_L/n_{fix}$ 、 $p_R=n_R/n_{fix}$ 、 $p_S=n_S/n_{fix}$  で与えられる。比較の方法として、図1 で用いた TPR-DB の2セッション (Head-starter と Context-planner) を観測

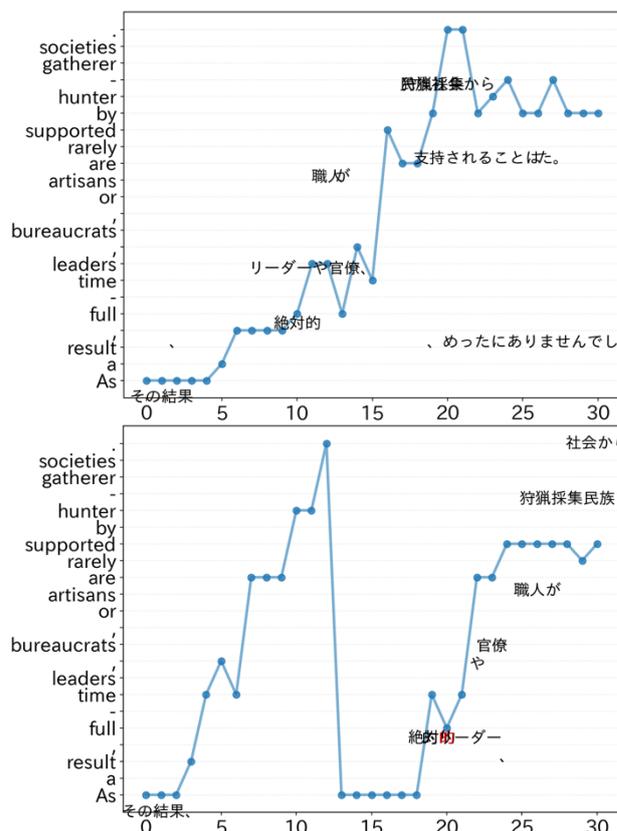


図3 シミュレーション結果。実データ (図1) と基本的には同じ形式のグラフだが、X軸は離散ステップ (合計30)。上が条件① (Flow 固定)、下が条件② (O→H→F) の一例。

値とし、各条件のシミュレーション 1000 試行分から同じ区間の fixation を取り出して得た  $p_L$ 、 $p_R$ 、 $p_S$  の分布と比較した。

まず観測値を見ると (表1)、Head-starter では当該区間の fixation が1回のみでタイプはSであった<sup>i</sup>。一方、Context-planner では fixation が33回観測され、最も多いのはLであった ( $p_L=0.49$ )。入力開始までにLが一定割合を占めており、図1に見られた「先に文脈を読む」という特徴と一致する。

次に、これらの観測値をモデルの生成分布に位置付けて比較した (表2)。観測値 (Obs.) はいずれも分布の95%区間に収まり、少なくとも外れ値とみなされるほど極端な値はなかった。Head-starter は  $n_{fix}=1$  のため極端な値になるが、特に  $p_L=0$  は分布の主要なモードと一致し (約67パーセント)、条件①シミュレーションが「最初にLがほとんど生じない」挙動を取りやすいことがわかる。一方、

<sup>i</sup> アノテーションのアルゴリズム上、1回のみ fixation はSと判定される[9]。

Context-planner の pL はモデルの分布の中ほど(約 61 パーセンタイル) に位置し、pS は上位、pR は下位に寄っていた。したがって、Context-planner の最初の読み部分に見られる「L が相対的に多い」という傾向を、条件②シミュレーションの生成分布の範囲内として説明できる。ただし pR が分布の下位であることから、このモデルは R を多めに生成している。

表 1 実データ (観測値) における L/R/S の要約

Style	n_fix	nL	nR	nS	pL	pR	pS
Head-starter	1	0	0	1	0.0	0.0	1.0
Context-planner	33	16	5	12	0.49	0.15	0.36

表 2 シミュレーション (1000 試行) による予測分布と観測値の比較。Med. はシミュレーションの中央値、95% PI は [2.5%, 97.5%] の予測区間、Obs. Pct. は観測値の百分位。

Cond.	Style	Stat	Obs.	Med.	95% PI	Obs. pct.
①(Flow)	Head-starter	pL	0.00	0.00	[0.00, 1.00]	66.7%
①(Flow)	Head-starter	pR	0.00	0.62	[0.00, 1.00]	9.2%
①(Flow)	Head-starter	pS	1.00	0.20	[0.00, 1.00]	100%
②(O→H→F)	Context-planner	pL	0.49	0.42	[0.00, 0.79]	60.7%
②(O→H→F)	Context-planner	pR	0.15	0.37	[0.10, 0.77]	7.7%
②(O→H→F)	Context-planner	pS	0.36	0.20	[0.00, 0.48]	87.7%

さらに図 4 では、主要な差が現れる pL について、各条件の分布をヒストグラムとして可視化し、観測値を破線で重ねた。条件①では pL が 0 近くに集中しやすい一方、条件②では 0.3~0.6 付近に山があり、条件間で「最初の読み進め量」が大きく変わることがわかる。またシミュレーションにばらつきはあるものの、その傾向が Head-starter や Context-planner などの実際の翻訳プロセスデータとも整合している。

以上より、能動的推論モデルのシミュレーションで、先読みと入力開始タイミングの差という主要な行動の違いが再現される可能性がある。また、視線パターンの比率という簡易的な指標でもその傾向が確認できた。

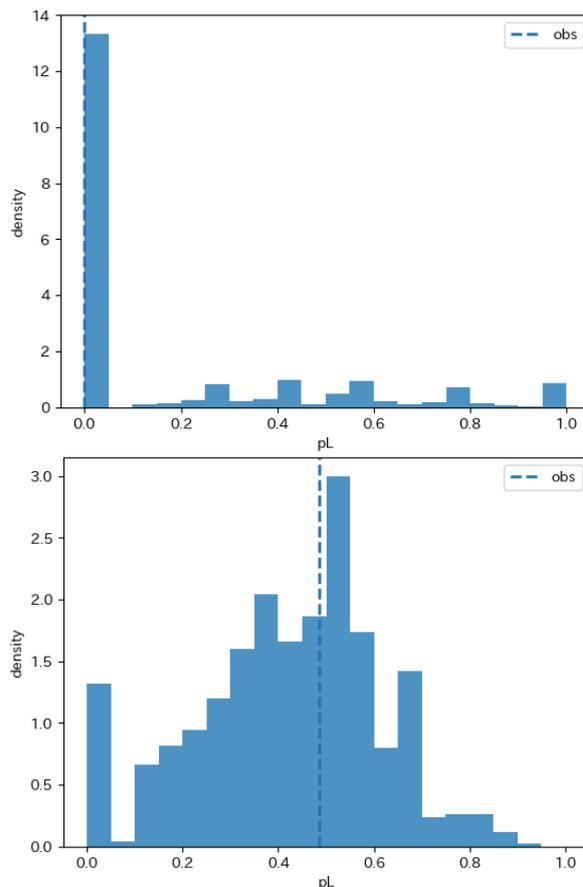


図 4 pL のヒストグラムと観測値。上が Flow 固定条件と Head-starter、下が O→H→F 条件と Context-planner。

## 4 おわりに

本稿では、翻訳プロセスにおける視線移動とキー入力をシミュレーションする能動的推論モデルを構築し、行動の事前分布を操作して、翻訳スタイルの違いを再現する可能性を提示した。また、実際の翻訳プロセスデータ由来の確率分布をモデルに組み込むことで、データに根ざしたシミュレーションが可能となることを示した。

ただし、本稿の結果は 1 文・少数例にもとづく予備的検証にとどまるため、翻訳セッション全体や多数の文、他言語への一般化は今後の課題である。加えて、現状は手動で設定しているパラメータの一部を実データから推定する方法の整備や、訳文の完成などの目標をモデルに明示的に組み込むことも必要である。さらに、量的評価を拡充し、翻訳スタイル差の説明に必要な要因をより厳密に同定していきたい。

## 謝辞

本研究は、独立行政法人日本学術振興会 特別研究員 (DC2) 奨励費 24KJ2065 の助成を受けたものです。

## 参考文献

- [1] Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. Active inference: A process theory. *Neural Computation*, Vol. 29, No. 1, pp. 1-49, 2017. doi:10.1162/NECO\_a\_00912
- [2] Michael Carl, Moritz Schaeffer, and Srinivas Bangalore. The CRITT Translation Process Research Database. In Michael Carl, Srinivas Bangalore, and Moritz Schaeffer (eds.), *New Directions in Empirical Translation Process Research: Exploring the CRITT TPR-DB*, pp. 13-54, Springer International Publishing, 2016. doi:10.1007/978-3-319-20358-4\_2
- [3] Michael Carl, Barbara Dragsted, and Arnt Lykke Jakobsen. A Taxonomy of Human Translation Styles. *Translation Journal*, Vol. 16, No. 2, 2011. <http://translationjournal.net/journal/56taxonomy.htm>
- [4] Michael Carl. Models of the Translation Process and the Free Energy Principle. *Entropy*, Vol. 25, No. 6, Art. 928, 2023. doi:10.3390/e25060928
- [5] Michael Carl. An Active Inference Agent for Modeling Human Translation Processes. *Entropy*, Vol. 26, No. 8, Art. 616, 2024. doi:10.3390/e26080616
- [6] Karl Friston, Richard Rosch, Thomas Parr, Cathy Price, and Howard Bowman. Deep temporal models and active inference. *Neuroscience & Biobehavioral Reviews*, Vol. 90, pp. 486-501, 2018. doi:10.1016/j.neubiorev.2018.04.004
- [7] Francesco Donnarumma, Mirco Frosolone, and Giovanni Pezzulo. Integrating large language models and active inference to understand eye movements in reading and dyslexia. *Physics of Life Reviews*, Vol. 55, pp. 61-78, 2025. doi:10.1016/j.plrev.2025.08.008
- [8] Thomas Parr, Giovanni Pezzulo, and Karl J Friston. A Recipe for Designing Active Inference Models. In Thomas Parr, Giovanni Pezzulo, and Karl J Friston (eds.), *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*, pp. 105-124, *The MIT Press*, 2022. doi:10.7551/mitpress/12441.003.0010
- [9] Michael Carl, Takanori Mizowaki, Aishvarya Ray, Masaru Yamada, Devi Sri Bandaru, and Xinyue Ren. The Behavioural Translation Style Space: Towards Simulating the Temporal Dynamics of Affect, Behaviour, and Cognition in Human Translation Production. *SKASE Journal of Translation and Interpretation*, Vol. 18, No. 2, pp. 212-239, 2025. doi:10.33542/JTI2025-S-11
- [10] Michael Carl, Yuxiang Wei, Sheng Lu, Longhui Zou, Takanori Mizowaki, and Masaru Yamada. Hesitation, orientation, and flow: A taxonomy for deep temporal translation architectures. *Ampersand*, Vol. 12, Art. 100164, 2024. doi:10.1016/j.amper.2024.100164
- [11] Michael Carl, Akiko Aizawa, and Masaru Yamada. ENJA15: A Free Corpus of English-Japanese Translation Process. The Eighth Asia-Pacific Translation and Interpreting Forum (APTIF 2016), Xi'an, China, 2016. <http://www.aptif8.org/>