

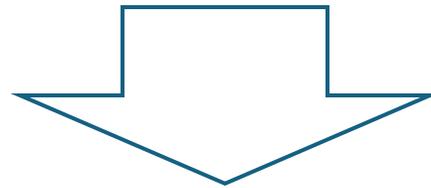
# 英日文芸翻訳における創造性の定量的分析 —人間翻訳とLLM翻訳の比較—

菅野公亮（関西大学）

# 背景・問い

## • 背景

- ① 文学作品へのMT活用への関心が拡大
- ② エラー分析や自動評価に偏った品質評価
- ③ 最近注目を浴びる「推論モデル」の影響

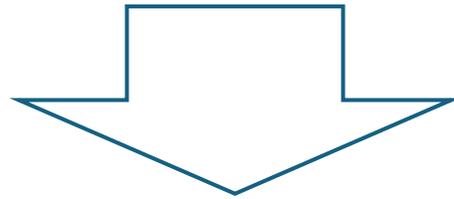


## • Research Questions

- ① プロの人間翻訳と比較して、最新のLLMによる翻訳はどの程度の創造性を示すのか？
- ② LLMの内部推論能力を高めることにより、翻訳における創造性は向上するのか？

# 創造性とは

- 心理学における**創造性**のコンセンサス  
Sawyer(2006)  
“個人や集団、または社会からの、**斬新かつ適切なもの**”

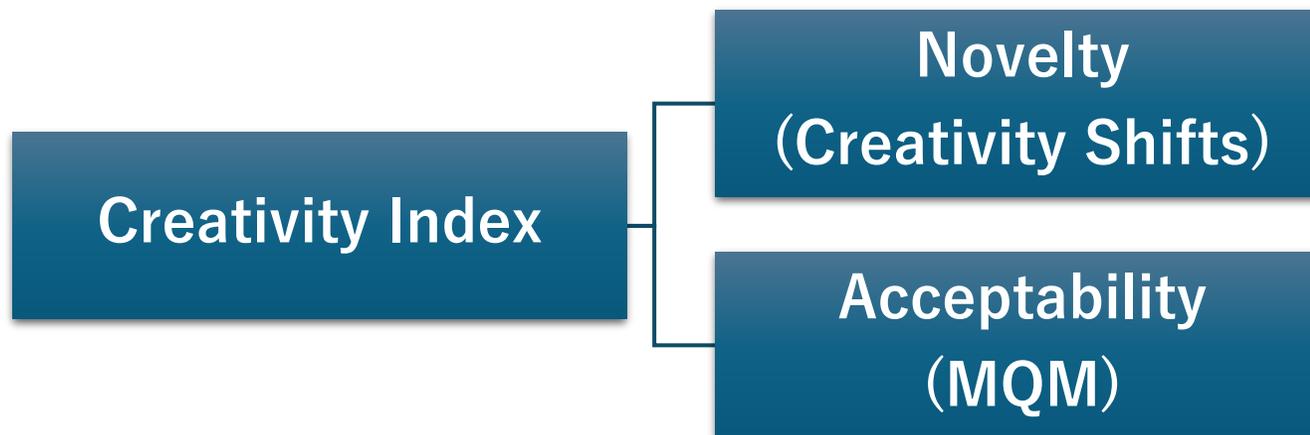


- **翻訳における創造性**  
Bayer-Hohenwarter & Kussmaul(2020)  
“創造的な翻訳とは、原文からの変更（**シフト**）を伴うことが多く、それにより**新しいもの**をもたらし、かつ設定された翻訳の目的（**スコポス**）に対して**適切である**翻訳”

# 評価指標

**Creativity Index (CI)** (Du et al., 2025)

⇒ **Novelty & Acceptability** の2軸から成る定量的指標



- **Acceptability (受容性) の測定**

⇒ MQMフレームワークによるエラー分析

💡 **MQM**…エラーの種類×重大度で体系的に翻訳を評価する枠組み

- **Novelty (新規性)**

⇒ **Creative Shifts (CSs) の特定・分類**

① 原文に“高い問題解決能力が必要とされる単位 (Units of Creative Potential: **UCP**) ”を設定

② 各UCPの訳出が以下のいずれかに該当する場合、**CSs**としてカウントする

**Abstraction**  
(抽象化)

**Concretisation**  
(具体化)

**Modification**  
(変更)

- **Creativity Indexの算出**

$$CI = \left( \frac{\#CSs}{\#UCPs} - \frac{\#error\ points}{\#words\ in\ ST} \right) \times 100$$

# データセット

原文: カート・ヴォネガット著『**2BR02B** (1962)』

💡 死が根絶され人々が永遠に生きられるものの、1人が生まれるには別の1人が死ななければならない未来のSF世界

## 人間翻訳

伊藤典夫氏訳「2BR02B」  
『バゴンボの嗅ぎタバコ入れ  
(2007)』早川書房所収

## LLM翻訳

	GPT訳①	GPT訳②
Model	gpt-5.2-2025-12-11	
Prompt	Translate the following text into Japanese	
Reasoning-effort	None	Medium
Temp	1.0	
Top P	0.98	

# 自動評価 (COMET)

- 両GPT翻訳は人間翻訳よりも有意にスコアが高かった。推論レベルによる有意差は認められなかった。

【統計手法】 対応のあるt検定およびボンフェローニ補正を実施 (有意水準  $\alpha = 0.0167$ )

※使用モデル(Unbabel/wmt22-cometkiwi-da), N=123

	Mean	SD
人間翻訳	0.777	0.094
GPT(None)	0.807	0.076
GPT(Medium)	0.815	0.065

	p値	Cohen's d
人間 vs. GPT(None)	$p < 0.001$	$d=0.316$
人間 vs. GPT(Medium)	$p < 0.001$	$d=0.400$
None vs. Medium	$p=0.034$ (n.s.)	$d=0.193$

# 分析結果 (Creativity Index)

- 2548語の原文から48文(602語)、54個の**UCP**を分析対象とした
- ※全ての翻訳においてCritical Errorは確認されなかった

## 💡 MQMでのError Points

Neutral→0  
Minor→1  
Major→5  
Critical→15

	CSs	Errors	Error Points	CI
人間翻訳	17	6	6	30.48
GPT訳 (None)	9	12	20	13.34
GPT訳 (Medium)	8	13	21	11.32

# Acceptability (MQM) の例

💡 GPT訳は、特に語の翻訳や文体・コロケーションに苦戦した

	テキスト
原文	(男性が悲嘆に沈み存在感を失っている場面) <b>His camouflage was perfect, since the waiting room had a disorderly and demoralized air, too.</b>
人間翻訳	カムフラージュは完璧といえた。待合室そのものが秩序と機能の乱れを感じさせたからだ。
GPT(None)	待合室のほうも同じく、 <u>散らかり、意気消沈した空気</u> をまとっていたから、彼のカムフラージュは完璧だった。
GPT(Medium)	待合室そのものにも、乱れ、 <u>士気のくじけた空気</u> があったから、彼の <u>擬装</u> は完璧だった。

※demoralize  
①風紀を乱す  
②士気をくじく

# Novelty (CSs) の例①

- UCP = "this daub"
- 直訳的なGPT訳とは異なり、人間翻訳はパラフレーズを行っている

	テキスト
原文	(望まぬ作品を描かされている画家が、その出来を称賛されて…) "You think I'm proud of <b>this daub</b> ?" he said.
人間翻訳	「このインチキ仕事をわたしが誇りにしていると思うかい？」
GPT(None)	「これが誇らしいと思ってるのか？」と言った。
GPT(Medium)	「こんな塗りが誇らしいと思うか？」彼は言った。

※ daub…  
塗りたくり、下手な絵

# Novelty (CSs) の例②

- **UCP** = "**Easy-Go**" (人を殺害する施設のあだ名)
- 直訳的なGPT訳とは異なり、人間翻訳はオノマトペを用いたあだ名らしい翻訳を行った

	テキスト
原文	It was the telephone number of an institution whose fanciful sobriquets included: (中略) " <b>Easy-Go</b> "
人間翻訳	「イージー・ゴー(ぽっくり)」
GPT(None)	「気楽に行こう」
GPT(Medium)	「気楽に行け」

※sobriquet…あだ名

# Novelty (CSs) の例③

- UCP = "dunk people"
- 人を殺害する施設で働く“Duncan”という女性の姓と“dunk people”が掛かった英語の言葉遊びに対しての方略の違い

	テキスト
原文	(人を殺害する施設で働く女性が画家に自己紹介する場面) "My name's Leora Duncan." She waited. "And you dunk people," he said.
人間翻訳	「人間を薬漬けにする仕事だな」
GPT(None)	「で、人を“沈める”わけだ」
GPT(Medium)	「それで、人を“沈める(ダンクする)”んだな」

# まとめ

- **人間翻訳は両方のGPT訳と比較して高い創造性示した**  
(約2倍のCSsと3分の1のエラ一点)
- **LLMの内部推論能力を高めても創造性の向上は確認されなかった**

※しかし、自動評価ではGPT翻訳が人間翻訳を上回った  
⇒意味的一致 (COMET) とCreative Shiftsのトレードオフが示唆される

💡 本実験ではGPTの素の翻訳性能を測定するため、Zero-shotプロンプトを採用した。今後は指示設計を組み込んだプロンプトが創造性を向上させる要因となり得るかを検証したい。

# References

- Bayer-Hohenwarter, G., & Kussmaul, P. (2020). Translation, creativity and cognition. In F. Alves & A. L. Jakobsen (Eds.), *The Routledge handbook of translation and cognition* (pp. 310–325). Routledge. <https://doi.org/10.4324/9781315178127-21>
- Du, S., Guerberof-Arenas, A., Toral, A., Gerrits, K., & Marco Borillo, J. (2025). Optimising ChatGPT for creativity in literary translation: A case study from English into Dutch, Chinese, Catalan and Spanish. In *Proceedings of Machine Translation Summit XX: Volume 1* (pp. 578–591). European Association for Machine Translation. <https://aclanthology.org/2025.mtsummit-1.44/>
- Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation*. Oxford University Press.
- Vonnegut, K. (1999). 2BR02B. In *Bagombo Snuff Box*. G. P. Putnam's Sons. (Original work published 1962)
- ヴォネガット, K. (2007). 2BR02B (伊藤典夫 訳). *バゴンボの嗅ぎタバコ入れ*. 早川書房. (原著出版 1999)