

# 文化の内挿・外挿・一貫性・時間の矢・拡散・マタイ効果～和歌の 10 世紀間の進化から～

田中琢真<sup>1</sup>

<sup>1</sup> 滋賀大学大学院データサイエンス研究科

## 概要

文化の歴史は人類以上に長い。文化史には素朴だが重要な問い多数がある。

1. 文化は内挿できるか？ すなわち、ある時代の文化は前後の時代の文化の中間か？
2. 文化は外挿できるか？ すなわち、次の時代の文化は予測できるか？
3. 文化の時系列はその逆と区別できるか？ 言い換えれば、時間の矢は観察可能か？
4. 文化的産物が後の産物に与えた影響の評価は、時代を超えて通用するか？
5. 文化の進化の駆動力は何か？

本研究は、古典和歌の系統樹ネットワークを構築し、これらの問いに答える試みである。7 世紀から 16 世紀の 20 万首を超える勢田勝郭氏の和歌データベースを使った。

系統樹での親子関係は、それぞれの歌についてその歌より古い歌の中で BERT の文ベクトルが最も近いものを親として決めた（最古の歌集の歌には親は設定しない）。この親子関係は、国文学者の認めた本歌取りと統計的に有意な一致を示した。

最初の勅撰集（古今集）をラベル 0、最後の勅撰集（新続古今集）をラベル 1 として  $k$ -nn 分類を行うと、間の勅撰集の平均ラベルは時代とともに滑らかに 0 から 1 に移行し、内挿・補間ができた。しかし、古今集と新古今集で同様に分類すると、新古今集より後の勅撰集は中間的な値になり、外挿・将来予測はできなかった。つまり、ランダム性が強い。

系統樹ネットワークについて、親と子の近さを測る指標  $\bar{r}$  を定義した。すると、時間順序を逆順にしたりランダムにしたりして構築するより、本来の時間順序で構築した方が  $\bar{r}$  が小さくなった。これは本来の時間順序による系統樹ネットワークが妥当であり、歌に対

応するベクトルの空間配置で時間の順方向と逆方向を識別できることを示す。すなわち、和歌の進化には時間の矢があった。

勅撰集  $i$  までの歌だけで学習して構築した系統樹ネットワークでのある歌の勅撰集  $i$  までの子の数は、全データセットで学習して構築した系統樹ネットワークでのその歌の勅撰集  $i$  より後の子の数と正の相関があった。子の数が本歌取りのされやすさ、つまり歌の影響力の指標だと仮定すれば、それぞれの歌の将来の影響力を予測できることになる。

初出が古く、勅撰集に選ばれた歌は、初出が同じ歌集だが選ばれなかった歌よりも（その勅撰集までの子の数が同じでも）その後の子の数が多い傾向があり、マタイ効果 (rich-get-richer) の存在が示唆された。典型的には子の数が 1.1–1.5 倍になった。

これらの結果は、和歌のマタイ効果とランダムウォークで進化したことを示す。重みが時間的に減衰する混合正規分布から歌が生成されるモデルは以上の結果を再現した。

Takuma TANAKA (2025) “Mean-reverting self-excitation drives evolution: phylogenetic analysis of a literary genre, *waka*, with a neural language model.” *Humanities and Social Sciences Communications* **12**, 394.