

## 2021年度森泰吉郎記念研究振興基金報告書

慶應義塾大学大学院政策メディア研究科 修士課程1年 中井瑞

### 研究課題名 多次元尺度による文法構造の史的変化の分析

研究概要 ヒトの言語はヒトだけが持ちうるユニークな文化である。しかし世界には8000を超える言語が知られており、それらは400以上の語族に分けられている。つまり人類集団の中で言語は、ヒト言語としての普遍性を保ちつつも多様性を生み出していると言える。今まで150年にわたりこれら多様な言語の歴史的関係性を推定する試みが行われてきたものの、統計学が発達した現在もこれらは容易ではない。しかし、新たな巨大データベースと文法理論の登場により未発見の事実が数多く眠っている可能性が示唆された事から、これらの探索と新たな推定法の確率を試みた。

研究成果 今年度の研究活動により、語族内・語族間の基礎語彙 (ASJP) と文法構造 (WALS, AUTOTYP) のデータを基に距離モデルを用いて言語間距離を推定した。その後、各種系統関係を推定し、任意の指標を用いてマッピングを行った。更に、高い推定精度を得られたと見られる幾つかの言語に関して、同一意味文について深層構造の表層化過程についてのモデル図を作成し、文脈自由文法に従い文型ごとに生成規則を与えた。

\*本研究課題は未発表データが含まれるため、詳細な結果およびデータの公開を控えさせていただきます。

謝辞 森泰吉郎記念研究振興基金はデータ解析に向けた計算機の購入のために支出させていただきました。又、研究における知識や技術を増やすだけでなく、より効率的に研究を行うことができました。ご支援をいただきありがとうございます。