

博報堂教育財団 第15回、16回「日本研究フェローシップ」

成果報告書

I. 研究成果概要

氏名（フリガナ） 在住国名	石原 俊一（イシハラ シュンイチ） オーストラリア
所属・役職	オーストラリア国立大学、准教授
招聘回（招聘研究期間）	第15回（2021年9月1日～ 2022年2月28日）
受入機関	国立国語研究所
招聘研究テーマ	日本語の書き言葉に見られる個人性の言語学的研究：テキストメッセージを中心に
研究目的	本研究は、書き言葉に見られる著者の個人性に関する言語学的基礎研究であり、日本語における。著者認識の手法を数理モデルで構築し、法科学分野へ適用する事を長期的な目的とする。人間が発する言葉、作成した文章には、言語情報以外にも多様な情報が含まれる。本研究では、書き言葉、特に現代社会のコミュニケーション手段として欠かせないテキストメッセージに注目し、文章中に著者の個人性がどの程度具現化されるかを、言語学的要素・構成をもとに計量的に評価し、明らかにする。また、先行研究における英語著者の場合と比較して、日本語著者に見られる個人性の特徴が、こういった傾向を示すかも考察する。
研究成果概要	
<p>1. どのように研究を進めたか（具体的に）</p> <p>日本語で書かれた関連分野の先行研究を集め、どのような特徴量が一般的に日本語著者認識に有効であるのかを知見を得ることから始まり（1ヶ月）、同時に著者認識のための計量的な基礎研究および長期的には数理モデルを構築しかつ実証するために必要な日本語テキストデータベースを集めることを同時に行った（1ヶ月）。大小異なった規模のデータベース、有償無償を含め、4つを手に入れることができた。これらのデータベースは本来著者認識のために構築されたものではないため、データベースを再構築する必要があり、その作業に2ヶ月程度を費やした。研究の後半には、先行研究で有効だと報告されている特徴量がどの程度実際に有効であるかを比較検討するために、再構築したデータベースを用い、計量的調査を行った。それと同時に、それらの特徴量を使用し簡単な著者認識実験も行った。これら2つの結果の概要は研究会で発表したものである。これらに加え、国立国語研究所所属の先生方、および科学警察研究所の연구원の方々と将来に向けての研究の方向性について有益な意見交換をすることが出来た。</p>	
<p>2. 研究によりどのような知見が得られたか（具体的に）</p> <p>先行研究から及び関連研究分野の先生方と意見交換することで、日本語での法科学著者認識の研究が進んで</p>	

いないことを再認識することができた。それと同時に日本語ではどのような特徴量が著者認識には有効なのか、先行研究から情報が得られた。その特徴量の有効性を入手したデータベースを基に比較し、異なった特徴量の相対的有効性の知見を得ることが出来た。具体的には、「空白」および「記号」の使用は高い個人性情報を含んでいることが実証できた。単語の頻度ベクトルだけを特徴量と使用した著者認識実験の結果から、かなりの精度で著者認識は可能であるとの結果が得られた。これらの研究はかなりコントロールされた状況下での実験であったため、今後は法科学的にさらに現実性のある環境での実験が必要である。

3. 研究成果（予定を含む）

○論文（題目、掲載誌、発行者、掲載月、内容の概略（200字以内））

- ・ 招聘研究テーマからの論文発表は特に無し。
- ・ 研究が進むにつれ研究結果を国際論文誌に発表していく予定であるが、具体的な物は現段階では特に無し。

○口頭発表（題目、イベントの名称、日・場所、内容の概略（200字以内））

- ・ 口頭発表（法科学的著者認識:DNA鑑定を踏まえて、国立国語研究所NINJALコロキウム、2021年10月12日、本講演では、テキストメッセージなどを分析し、その結果を科学的証拠として裁判などで提示する場合の方法論、法科学的著者認識と一般的著者認識の違い、法科学的著者認識の課題、研究傾向などについて議論した。）
- ・ 口頭発表（Analysing linguistic text evidence with the likelihood ratio framework, Germanic Society for Forensic Linguistics Lecture Series, 2021年12月16日、本講演は尤度比を使った法科学的著者認識の有用性を議論するために行われた。）

○その他の活動

- ・ 現在、2大学の先生から客員講演・講義の依頼があるが詳細は検討中である。
- ・ 直接当該研究とは関係はないが、私の母校の山梨県立甲府西高校からオーストラリアの高校生とのオンラインでの交流イベント開催の可能性の問い合わせがあり、私の知り合いが教師として勤めている高校を紹介した。2月9日に最初のイベントが開催され、盛況のうちに終わったようである。2週間おきに続けていく予定だと聞いている。

4. 今後の活動予定

入手したデータベースのうち、再処理をしてないデータベースの処理を完了させる。

実用に向けた科学的及び工学的基礎実験を続けて行く。

幸いにも、大学院生の数名が興味を示してくれているので色々研究が進みそうである。