

京都大学人文科学研究所共同研究実績・活動報告書

(4年計画の 3年度目)

1. 研究課題

(和文) 東アジア古典文献コーパスの研究

(英文) Study on Corpus of Ancient Chinese

2. 研究代表者

(氏名) 安岡孝一

3. 研究期間

平成 20年 4月 から 平成 24年 3月 まで

4. 研究目的 (400字程度)

わが国の文化的活動は、奈良時代より随時もたらされた大量の中国の漢文と、それを訓読という手法によって解釈した歴史なしには語るができない。特に、近世、明治～昭和初期における漢文訓読という営為は、伝統的な日本の思想基盤を見直す作業であったとともに、近代国家としての思想基盤を再構築する過程において、重要な位置を占めるものであった。

本共同研究の構想は、その前身となった共同研究班「漢字情報学の構築」(2004年4月～2008年3月、班長：安岡孝一)で培われた知識、特に、コンピュータを用いた漢文解析に関する研究の中から生まれた。その中で、われわれは、日本における訓読が、元となる漢文とその読み下し文との間でそれぞれ異なる文法構造を橋渡しするための手法であることから、漢文の文法解析において有効に働きうるはずだ、という感触を得た。すなわち、訓読という手法を情報学的な視点から再検討し、訓読漢文コーパスともいべきものを作成して、それを他の一般的な漢文(白文)に適用することで、漢文の意味構造を解析可能になるだろう、という予想である。

本共同研究では、明治～昭和初期に日本国内で作成された訓読漢文テキストをコーパス化し、それを基に、漢文の意味構造を解析するシステムの研究・開発をおこなっている。すなわち、これまで漢文を読むための技法に過ぎなかった訓読を、コンピュータによる文法解析メソッドの一つとして、情報学的視点から捉えなおしている。この結果、文法的な構造化がおこなわれずに単なる文字列のままに放置されている大量の漢文に対して、その意味構造を解析することが可能となり、漢文の内容理解への大きな足がかりになると考えられる。

5. 本年度の研究実施状況 (400字程度)

平成22年度は、漢文コーパスの制作作業の準備、および制作環境の構築を中心に、共同研究をおこなった。なお、本研究班では、参加者全員が文献や書籍を見ながら論じ合うというスタイルを取っているため、特定の発表者等は記さないことにする。

2010年4月16日

科学研究費補助金の配分と今後の研究計画について

2010年5月7日

制作環境プラットフォームの設計

2010年5月21日

富山房の漢文大系

國譯一切經

言選Web

2010年6月4日

爾雅

中国哲学書電子化計画

2010年7月2日

Mac OS X上での作業環境構築(1)

2010年7月16日

Mac OS X上での作業環境構築(2)

2010年9月3日

Book Scanners

Mac OS X上での作業環境構築(3)

2010年9月17日

Mac OS X上での作業環境構築(4)

2010年10月1日

Mac OS X上での作業環境構築(5)

2010年10月15日

『漢文大系』のPDF化

Mac OS X上での作業環境構築(6)

2010年11月5日

Mac OS X上での作業環境構築(7)

2010年11月19日

Mac OS Xでの作業環境構築(マシン間のコピー)

2010年12月3日

emacs上でのmecab-kanbun自動処理

2010年12月17日

automatorによるemacs立ち上げ自動化

2011年1月21日

Osaks Symposium on Digital Humanities 2011での発表に向けて

2011年2月4日

Osaka Symposium on Digital Humanities 2011投稿原稿読み合わせ

2011年2月18日

『文字と非文字のアーカイブズ』シンポジウムとの連携に関して

6. 研究成果の概要 (400字程度)

平成20年度に開始した本研究班は、平成22年度に至り、全国共同利用・共同研究拠点「人文学諸領域の複合的共同研究国際拠点」の一翼を担うこととなった。また、科学研究費補助金(基盤研究(B))「形態素解析のための品詞情報つき古典文献コーパスの構築」(平成22～24年度)を獲得し、これまで手弁当でおこなってきた漢文コーパス構築作業を、外部資金によって進めることができるようになった。

平成22年度は、漢文コーパスの構築のための基礎作業として、『漢文大系』(富山房)の全文画像を構築し、さらに全文テキスト化のための目次情報を構築した。これに並行して、品詞処理のためのプロトタイプを設計し、それと同時に、散文と韻文を分離する手法の開発をおこなっている。

7. 共同研究班に関連した公表実績 (出版、公開シンポジウム、学会分科会、電子媒体など)

平成20～21年度の本共同研究班の研究成果を、2010年7月7～10日ロンドンKing's Collegeで開催のDigital Humanities 2010において、パネルセッション「The Origins and Current State of Digitization of Humanities in Japan」として発表した。

また、平成22年度の本共同研究班の研究成果を、2011年3月28～29日開催の国際シンポジウムOsaka Symposium on Digital Humanities 2011において発表すべく、3件のextended abstractを投稿したところ、見事3件とも採択された。しかし、2011年3月11日に発生した東日本大震災により、Osaka Symposium on Digital Humanities 2011そのものが2011年夏に開催延期となってしまったため、現時点では平成22年度の成果を、十分には公表しえていない。

ただし、共同研究班での議論、および参考とした資料・プログラム等は、逐一<http://kanji.zinbun.kyoto-u.ac.jp/~yasuoka/kyodokenkyu/>で、WWWに公開している。