

# 京都大学人文科学研究所共同研究実績・活動報告書

( 3 年計画の 2 年目)

## 1. 研究課題

東アジア古典文献コーパスの実証研究

Empirical Research on Digital Analysis of Classical Chinese Texts

## 2. 研究代表者氏名

安岡孝一

YASUOKA Koichi

## 3. 研究期間

2016 年 04 月 - 2019 年 03 月 (2 年度目)

## 4. 研究目的

2010 年以來、我々が構築を続けてきた漢文コーパスは、MeCab を用いた形態素解析手法を、漢文処理に適用するものである。この漢文コーパスでは 4 階層の品詞体系を採用しており、その第 2 層は「名詞」「代名詞」「数詞」「動詞」「前置詞」「副詞」「助動詞」「助詞」「感嘆詞」の 9 種類の品詞で構成される。すなわち我々は、従来の漢文文法等で見られた「形容詞」を廃止しているのだが、これが動詞類全体にどのような影響を及ぼしているのかは、必ずしも十分に検討できていない。本共同研究では、漢文コーパスにおける動詞類の実証研究をおこなう。すなわち、実際のコーパスにおいて「動詞」「前置詞」「副詞」「助動詞」の 4 つのふるまいを研究し、さらに下層の意味素性と小素性についても、現在の品詞体系の妥当性を検証する。

In our recent research we have developed a method to analyze classical Chinese texts. In our method, we use our original morphological analyzer based on MeCab. We proposed our original four-level word-class system for classical Chinese on the MeCab-based analyzer. We designed the top level of the word-class system to represent the predicate-object structure of classical Chinese. The second level is the ordinary word-class of classical Chinese, consisting: “noun,” “pronoun,” “numeral,” “verb,” “preposition,” “adverb,” “auxiliary verb,” “particle,” and “interjection.” The third and fourth levels are word-subclasses to describe detailed behavior of the words in classical Chinese texts. In other words, we excluded “adjective” from the second level of our word-class system. But we did not precisely examine the effect of lack of “adjective” for our morphological analyzer. In this

research we will examine the effectiveness of our four-level word-class system, especially examine “verb,” “preposition,” “adverb,” and “auxiliary verb” in the second level. We will also examine the validity of the third and fourth levels of our word-class system.

#### 5. 本年度の研究実施状況

平成 29 年度前半は、S 式によって漢文のメタ文法(文法の文法)を記述する手法に対し、Edwin George Pulleyblank の『Outline of Classical Chinese Grammar』の各例文を記述することで、その有効性の検証をおこなった。その結果、この手法は有効であるものの、作業のためのオーバーヘッドが大きく、Treebank 等の「量産化」には不適當な側面があることが判明した。そこで、平成 29 年度後半は、Universal Dependencies を用いて、漢文の依存文法解析をおこなう手法をも検討した。これに伴い、S 式によるメタ記述と、Universal Dependencies による依存文法記述の、両方を同一の例文に適用し、その得失(というか得手不得手)の比較をおこない続けている。

#### 7. 本年度の研究実施内容

2017-04-21 2017 年度前半活動方針

2017-05-12 「晋克衛磁洛州」をどうするか

2017-06-02 Outline of Classical Chinese Grammar の例文検討

2017-06-16 Outline of Classical Chinese Grammar の例文検討

2017-07-07 Universal Dependencies Tools

2017-07-21 Outline of Classical Chinese Grammar の例文検討

2017-09-08 Outline of Classical Chinese Grammar の例文検討

2017-09-22 漢文 Universal Dependencies(仮)

2017-10-06 Outline of Classical Chinese Grammar の例文検討

2017-10-20 SVG による CoNLL-U ヴューワ

2017-11-17 漢文 Universal Dependencies(仮)

2017-12-01 MeCab-Kanbun to CoNLL-U

2017-12-15 MeCab-Kanbun to CoNLL-U

2018-01-12 SVG による ictree ヴューワ

2018-01-27 人文科学とコンピュータ第 116 回研究発表会

2018-01-28 人文科学とコンピュータ第 116 回研究発表会『古典中国語 Universal Dependencies への挑戦』

## 10. 共同利用・共同研究の参加状況

区分	機関数	参加人数				延べ人数			
		総計	外国人	大学院生	若手研究者	総計	外国人	大学院生	若手研究者
所内	1	5 (1)	1 (0)	0	1	40 (6)	11 (0)	0	(6)
学内	0	0	0	0	0	0	0	0	0
国立大学	1	1 (0)	0	0	0	5 (0)	0	0	0
公立大学	0	0	0	0	0	0	0	0	0
私立大学	2	3 (0)	0	0	0	15 (0)	0	0	0
大学共同利用機関法人	0	0	0	0	0	0	0	0	0
独立行政法人等公的研究機関	0	0	0	0	0	0	0	0	0
民間機関	0	0	0	0	0	0	0	0	0
外国機関	0	0	0	0	0	0	0	0	0
その他	0	0	0	0	0	0	0	0	0
計	4	9 (1)	1 (0)	0	0 (1)	60 (6)	11 (0)	0	(6)

※( )内には、女性数を記載

## 11. 本年度 共同利用・共同研究を活用して発表された論文数

参加研究者がファーストオーサーであるものを対象

総論文数	2(2)
国際学術誌に掲載された論文数	0

※( )内には、拠点外の研究者による成果(内数)を記載