

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 20 日現在

機関番号：37402

研究種目：基盤研究(B)

研究期間：2011～2013

課題番号：23320101

研究課題名(和文) 高機能インターフェイスを備えたデジタルディケンズレキシコン作成とその活用研究

研究課題名(英文) The Development of Dickens Lexicon Digital with a Multifunctional Information Retrieval System and its Practical use for Linguistic Research

研究代表者

堀 正広 (HORI, Masahiro)

熊本学園大学・外国語学部・教授

研究者番号：20238778

交付決定額(研究期間全体)：(直接経費) 8,100,000円、(間接経費) 2,430,000円

研究成果の概要(和文)：本研究の目的は、山本忠雄博士(1904-91)がその著Growth and System of the Language of Dickens: An Introduction to A Dickens Lexicon(1946年東京大学より文学博士の学位を授与、1953年学士院賞受賞)において構想し、自ら収集した6万枚のカードに基づき、Dickens Lexicon Digitalを作成することである。18世紀と19世紀英国小説の電子テキストを含めた多機能搭載型のLexiconとして、インターネット上で公開する。完成版の公開は2017年3月を目標だが、部分的な公開は2014年7月に行う。

研究成果の概要(英文)：Our project to create The Dickens Lexicon Digital is based on approximately 60,000 index cards which the late Dr. Tadao Yamamoto (1904-91) compiled. Our Dickens Lexicon is neither book-based nor document-based, but is rather designed as a web-based reference resource. The Dickens Lexicon Digital with a Multifunctional Information Retrieval System, including the e-texts of 18th and 19th century fiction is scheduled to be available on the web by March 2017, although it will be partially functional and available on the web, from July 2014.

研究分野：言語学

科研費の分科・細目：英語学

キーワード：レキシコン デケンズ コンピュータ 電子テキスト 言語・文体 18世紀・19世紀小説 イデオモ インターフェイス

1. 研究開始当初の背景

故山本忠雄博士(1904—1991)の *Growth and System of the Language of Dickens: An Introduction to A Dickens Lexicon* (1943)は、1946年(昭和21年)に東京大学より文学博士の学位を授与された。この博士論文は、戦後になって関西大学(英語学会)の故八鳥治一教授の尽力によって、1950(昭和25年)に出版され、1952年に第2版が、そして、2003年第3版の改訂版(溪水社)が出ている。また、1953年(昭和28年)には、英語英文学研究者としては初めての学士院賞を受賞された。

この *Growth and System* は、副題にあるように *Dickens Lexicon* の作成を最終目標として書かれた Introduction である。山本博士は *Lexicon* の作成のために様々な共同研究を試みられたがいずれも満足のいく結果を得られなかった。その後は、山本博士独力で用例の収集に当たられ、*Pickwick* からもう一度作品の年代順にカードが集められた。残念ながら、*Lexicon* は完成されることなく1991年に他界された。山本博士が他界された後、カードの所在はわからなかった。

山本博士の死後、博士自ら集められたカードの所在は不明のままであったが、1997年広島時代の教え子にご遺族から博士の蔵書を寄贈したいとの申し出があった。その蔵書の整理中に博士の書斎の一角に膨大な数のカードを収納した大きな木製のボックスが発見され、ご遺族の許可をいただき、カードを借用し、当時の共同研究者の一人であった黒瀬保を代表として、広島大学及び熊本大学出身の研究者、総勢20名からなる *Dickens Lexicon* 編纂プロジェクトチームが、1998年に結成されることとなった。その後、メンバーに Web アプリケーション開発の専門家と統計処理の専門家、さらに *Dickens* を専門とする定義入力担当者が加わった。

2. 研究の目的

本研究は、山本忠雄博士(1904-91)(元広島大学・神戸大学教授)がその著 *Growth and System of the Language of Dickens: An*

Introduction to A Dickens Lexicon (1946年東京大学より文学博士の学位を授与、1950年に出版され、1953年学士院賞受賞、2003年第3版の改訂版(溪水社)刊行)において構想した *Dickens Lexicon* を、コンピュータ技術を駆使して様々な機能を搭載し、*Dickens* の言語文体研究のみならず18世紀、19世紀の英国小説の言語文体研究に寄与することを目的とした、研究者20数名によるプロジェクトである。本 *Lexicon* は冊子体資料ではなく、インターネット上で公開する予定である。本 *Lexicon* の付加機能として *Dickens* の全作品および18、19世紀を代表する作家の主要作品のテキストを収録する。また、本 *Lexicon* を使用して書かれた研究論文を公表する。

3. 研究の方法

DL 編纂作業は、以下の手順で行われた

編纂作業第一段階

まず、アルファベット順にカードを並べ枚数を確認する。53,807枚に達したカードを担当者14人の作業員で分担し、FileMaker というデータベースソフトを使い入力した。

編纂作業第二段階

編纂第一段階終了後、引用箇所の間違い、引用文の長さの不統一、山本博士が利用した辞書の特定、語義の未定義などの問題に直面した。編纂第二段階では、まず、引用文の見直しについて修正した。第一段階で入力したすべてのカードの引用文を適切な長さにし、出典を再確認するという作業を行った。

編纂作業第三段階

現在、編纂作業は第三段階に入っている。山本博士の残されたカードのおよそ3分の2は、語義が未定義のままになっているため、第三段階では、未定義の語義を定義する作業を行っている。この作業は、ディケンズの専門家が担当している。博士は、*The Oxford English Dictionary (OED)*の他に *The Concise Oxford Dictionary of Current English (COD)*の第二版と *The Pocket Oxford Dictionary of Current English (POD)*の

初版を頻繁に使用していた。その結果、第三段階終了後は図1のようになる。

本博士によるメモやその他役立つ情報なども記載される。

Entry: board **Word Class:** *n.*

Definition: The company of persons who meet at a council-table.

Dictionary: OED (NED)

Quoted in the OED? No
(if quoted) given as:

Title: B (Sketches by Boz) **Chapter:** Our Parish, 01

Text: Her previous admiration of “the board” (who all sit behind great books, and with their hats on) fades into nothing before her respect for her lace-trimmed conductor; and her account of what has passed inside, increases—if that be possible—the marks of respect, shown by the assembled crowd, to that solemn functionary.

cf. ‘Bow to the board,’ said Bumble. Oliver ... seeing no board but the table, fortunately bowed to that. (*Oliver Twist*, II)

図1 Board の編集第3段階終了図

4. 研究成果

現在、Web上で、The Dickens Lexicon Digital のテストバージョンが共同研究者間での利用が可能となっている。現在の機能としては、まず、キーワード欄に検索したい語句を入力すれば、すぐに検索結果を見ることが出来る。検索条件には、見出し語、定義、作品名、章、引用、OEDの引用の有無、コメントがあり、それぞれの項目にチェックを入れる事で検索が可能となる。また、検索条件として、カード、テキストのいずれかを選べば、検索範囲を指定する事もできる。カードはアルファベット順に並んでいるが、作品順に並び替える事もできる。例えば、“do the honours of the house”のカードをクリックすると、このイディオムのカードを見る事ができる。OEDからの定義、テキストのタイトル、チャプターだけでなく、実際に引用されているテキストも読む事ができるため、前後の文脈も確認する事ができる。また、Cf.の項目では、山



図2 A test version of the Dickens Lexicon

小説の電子テキストに関しては、ディケンズのテキストだけではなく、他の作家の作品も検索できるような機能を提供している。現在のテストバージョンでは、ジェイン・オースティンの作品のみが利用可能となっているが、今後は18世紀から19世紀の小説の電子テキストを網羅的に収録し、

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

(雑誌論文) (計 15 件)

Tabata Tomoji, Stylometry of Collaboration: Pinpointing style changes in the text of mixed authorship, 『人文学データのマイニング』, 査読有, 322 巻, 2014, pp. 35-44.

堀正広, 新渡戸稲造の英文から我々は何を学ぶか, 日本英文学会第 85 回大会 Proceedings, 査読無 (招待発表), 2013, pp. 29-30

Hori, Masahiro and Joseph Tomei, How is Bushido Read by Modern Native Speakers of English?-With Special Reference to the English of Bushido-, 『海外事情研究』, 査読有, 第 40 巻第 2 号 (通巻 81 号) 熊本学園大学付属海外事情研究所, 2013, pp. 21-36.

Nishio Miyuki, Dickens's Artistry of Reporting Verbs, *On-line Proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA)*, 査読無, 2013

永崎研宣, 人文学分野とサイバーインフラストラクチャ ~ デジタル・ヒューマニティーズにおける現状と課題 ~, 『情報の科学と技術』, 査読有, Vol. 63, No.9, 2013, pp. 369-376

Tabata Tomoji, Approaching Dickens's Style through Random Forests, Digital Humanities 2012 Conference Abstracts, 査読有, 2012, pp. 388-391

田畑 智司, テクストマイニングからテキスト分析へ: Collins との共著作品における Dickens の文体, 電子化言語資料分析研究, 査読無, 2011-2012, 2012, pp. 3-18

島(西尾)美由紀, 「Many's the NP」構造についての一考察 Dickens の作品を中心として, 『近代英語研究』, 査読有, 28 号, 2012, pp. 51-70

Kiyonori Nagasaki, Toru Tomabechi, Dorji Wangchuk, Koichi Takahashi, Jeff Wallman, A. Charles Muller, Approaches to the Treatment of Primary Materials in Digital Lexicons: Examples of the New Generation of Digital Lexicons for Buddhist Studies, Digital Humanities 2012 Conference Abstracts, 査読有, 2012, pp. 61-64

Tabata Tomoji, 'Key' Words and Stylistic 'Signatures': Textometry with Random Forests, 統計数理研究所共同研究レポート 『マイニング技術を応用したテキスト分析研究』, 査読有, 278 巻, 2012, pp. 46-64

堀正広・田畑智司・今林修・島(西尾)美由紀・地村彰之, The *Dickens Lexicon* Project と *Lexicon* の利用法, 大阪大谷大学英文学会 『英語英文学研究』, 査読無, 第 38 号,

2011 . Pp. 143-202

(学会発表) (計 25 件)

Kiyonori Nagasaki, North American Coordinating Committee on Japanese Library Resources [NCC] Meetings, (Philadelphia, USA), Digital Resources of Japanese Texts: from a Viewpoint of Digital Humanities, 査読無 (招待講演), 27 March, 2014

Tabata Tomoji, Humanities Studies in the Digital Age and the Role of Buddhist Studies (University of Tokyo), Opening up a New Perspective for Text Analysis in the Digital Age, 査読無 (招待講演), 17 November 2013

Kouguchi Keisuke, PALA 2013 (Poetics and Linguistics Association) (University of Heidelberg, Germany), Stylistic Use of Participant Items in *A Tale of Two Cities*, 査読有, 3 August, 2013

Tabata Tomoji, PALA 2013 (Poetics and Linguistics Association) (University of Heidelberg, Germany), Stylometry of Collaborations: Dickens, Collins and their collaborations, 査読有, 1 August 2013

Nishio Miyuki, PALA 2013 (Poetics and Linguistics Association) (University of Heidelberg, Germany), Dickens's Artistry of the Reporting Verbs, 査読有, 31 July, 2013
今林修, 近代英語協会第 30 回大会 (愛知大学), 英語英文学研究と「選択」としての文体, 査読有, 2013.7.6

堀正広, 日本英文学会第 85 回大会 (東北大学) 『新渡戸稲造の英文から我々は何を学ぶか 語彙力が文法力かレトリックか, あるいは』, 査読無 (招待発表), 2013.5.25
堀正広, 英語コーパス学会春季シンポジウム (大阪大学) 『3つの柱: テクストの読み, 言語理論, コーパスの利用』, 査読有, 2013.4.27

堀正広, 言語文化レトリック研究会, 『異文化翻訳の諸問題: 『五輪書』, 俳句, 禅公案, *Bushido*』, 査読無 (招待講演), 2013.3.20

田畑 智司, 計量的言語研究の諸相 (北海道大学大学院メディア・コミュニケーション研究院), Key Words and Textometry: Are key words really "key" words?, 査読無 (招待講演), 2012.9.19

Hori Masahiro, Dickens Society, Collocational and Idiomatic Creativity in the Pickwick Papers, 査読有, 14 September, 2012

Imahayashi, Osamu, Tomoji Tabata, and Miyuki Nishio, Dickens Fellowship Annual Conference 2012 (University of Portsmouth), 査読有, 12 August, 2012.

Jimura, Akiyuki, Yoshiyuki Nakao, Noriyuki, Kawano, The 18th Congress of the New Chaucer Society (Portland, Oregon, USA),

査読有, 23 July, 2012

堀正広, 言語文化レトリック研究会, 「認知文体論と古典作品の新たな読み」(大阪大学), 査読無, 2012.3.28

堀正広, メディア英語談話分析研究分科会とメディア英語教授法・教材研究分科会(関西外国語大学), 「ディスコース分析及び英語教育のためのコーパス利用」, 査読無(招待講演), 2011.12.24

Imahayashi Osamu, Middle and Modern English Corpus Linguistics 2011 (Osaka University), Adverbial Idioms in Dickens: A Corpus Stylistic Analysis, 査読有, 29 August, 2011

Tabata Tomoji, Language Individuation: A Symposium in Honour of John Burrows (University of Newcastle, NSW, Australia), Using Random Forests to identify Dickensian style, 査読有, 6 July, 2011

Hori Masahiro, Modern Linguistics Association in Korea (Dajeon, Korea), Chronological Study of English Collocations, 査読有, 28 May, 2011

Tabata Tomoji, Modern Linguistics Association in Korea (Dajeon, Korea), Text-Mining in Corpus Stylistics: Spotlighting Linguistic Variations in the Inaugural Addresses of U.S. Presidents, 査読有, 28 May, 2011

(図書)(計 11 件)

赤野一郎・堀正広・投野由紀夫編, 大修館, 『英語教師のためのコーパス活用ガイド』, 2014, 242

田畑 智司・岸江信介編, ひつじ書房, 『テキストマイニングによる言語研究』, 2014, 200

Akiyuki Jimura 他, Osaka Books Ltd, *Studies in Middle and Modern English: Historical Change*, 2014, 115-32

Akiyuki Jimura 他, Osaka Books Ltd, *Chaucer's Language: Cognitive Perspectives*, 2013, 27-45

Hori, Masahiro 他, Peter Lang, *Approaching Language Variation through Corpora*, 2013, 421

堀正広編, ひつじ書房, 『これからのコロケーション研究』, 2012, 271

Hori, Masahiro 他, Red Moon Press USA, Kaneko Tohta Selected Haiku With Notes and Commentary: Part2:1961-2012, 2012, 250

Hori, Masahiro 他, Red Moon Press USA, Kaneko Tohta Selected Haiku With Essays and Commentary Part 1: 1937-1960, 2012, 256

Hori, Masahiro 他, Red Moon Press USA, The Future of Haiku: An Interview with Kaneko Tohta, 2011, 137

Hori, Masahiro 他, Red Moon Press USA, Kaneko Tohta Selected kimonofuei: Poetic Composition on Living Things, 2011, 91

堀正広著, 研究社, 『例題で学ぶ英語コロケーション』, 2011, 224

6. 研究組織

(1) 研究代表者

堀 正広 (HORI, Masahiro)

熊本学園大学・外国語学部・教授

研究者番号: 20238778

(2) 研究分担者

地村 彰之 (JIMURA, Akiyuki)

広島大学・文学研究科・教授

研究者番号: 00131409

(3) 研究分担者

田畑 智司 (TABATA, Tomoji)

大阪大学・言語文化研究科(研究院)・

准教授 研究者番号: 10249873

(4) 研究分担者

高口 圭輔 (KOUGUUCHI, Keisuke)

安田女子大学・文学部・教授

研究者番号: 50195658

(5) 研究分担者

島(西尾)美由紀 (SHIMA (NISHIO),

Miyuki) 近畿大学・工学部・講師

研究者番号: 50549524

(6) 研究分担者

今林 修 (IMAHAYASHI, Osamu)

広島大学・文学研究科・准教授

研究者番号: 90278987

(7) 研究分担者

永崎 研宣 (NAGASAKI, Kiyonori)

一般財団法人人文情報学研究所・人文情

報学研究部門・人文情報学研究所 所長

研究者番号: 30343429