

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 8 日現在

機関番号：13901

研究種目：挑戦的萌芽研究

研究期間：2013～2015

課題番号：25580112

研究課題名(和文) 漢字の認知的意味空間 中国, 韓国, ベトナム, モンゴルの日本語学習者の比較検討

研究課題名(英文) Cognitive space of kanji semantics - A contrastive study of Chinese, Korean, Vietnamese and Mongolians learning Japanese

研究代表者

玉岡 賀津雄 (Tamaoka, Katsuo)

名古屋大学・国際言語文化研究科・教授

研究者番号：70227263

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本科研では、(1)中国語・韓国語日本語学習者用のために日韓中の3言語間の品詞性および関連する情報データベースとWeb上検索エンジン(<http://kanjigodb.herokuapp.com>)を完成した。(2)11年分の新聞コーパスを使って、常用漢字2136字の特性および二字熟語のデータベースとWeb上検索エンジン(www.kanjidatabase.com)を公開した。(3)ネット上で使える部首と漢字の関係を実験するための漢字意味空間実験ソフトを作成した。(4)日本語と中国語の語彙が、非選択的に活性化されるが、ターゲット文の処理には影響しないことを脳波実験で示した。

研究成果の概要(英文)：In this grant, we have done four things: first, we created a database of two-kanji compound words for native Chinese- and Korean-speaking students who are learning Japanese, available as a search engine(<http://kanjigodb.herokuapp.com>). Second, using a corpus from the Mainichi newspaper (2000-2010), we created a Web-accessible search engine for 2136 joyo-kanji and their compounds (www.kanjidatabase.com). A paper explaining these tools was accepted by Psychological Research (Tamaoka et al., in press). Third, we created a Web-accessible experimental software for kanji semantic space. Finally, we published an ERP study of Chinese-and-Japanese bilinguals, indicating non-language selective word activation of both languages, but a cut-off mechanism for sentential processing of a target language (Tamaoka et al., 2016).

研究分野：計量言語学

キーワード：漢字データベース Web検索エンジン 日本語学習者 二字漢字語 非言語選択的活性化 漢字の意味空間 漢字特性 部首

1. 研究開始当初の背景

漢字は意味的な最小単位(形態素)であり、その組み合わせによって多様な語彙を形成する。そのため、漢字形態素の意味的なネットワークが適切に形成されることが、日本語学習者の語彙的な知識を拡張させると予想される。そのために、漢字データベースとその検索エンジンの作成および二言語間の語彙の意味的な結合を実験で考察することを当初の目的とした。

2. 研究の目的

本研究では、二言語間の漢字のネットワークを利用して、日本語学習者の漢字語彙処理能力の向上を試みた。中国語を母語とする日本語学習者であれば、日本語の漢字と形態的に類似しているため、両言語間の意味的空間によるネットワークの形成が容易に行われると予想される。韓国語では、漢字を表記形態としては使用していない。しかし、両言語では、漢語語彙が多く、日本語と韓国語は、音韻的に類似した語彙が多い。音韻的な単位も形態的な単位と結合することで、日本語の漢字の意味的なネットワークを形成すると予想される。非漢字圏のベトナム語母語話者の場合でも、韓国語母語話者と同じように、伝統的に漢字を単位として音韻的な表現(中国語から借用された語彙)が多数存在し、日本語の漢字を学習することで、ベトナム語の形態素と日本語の漢字との意味的な結合関係が形成され、意味的なネットワークを形成することが可能であると予想される。モンゴル語については、韓国語やベトナム語よりさらに、日本語の漢字形態素との関係が希薄になるので、漢字の意味的空間やネットワークの形成は、日本語の学習にともない学習対象である日本語の1言語内で行われると予想される。多様な日本語学習者の漢字学習を促進するために、漢字の多様な特性のデータベースとそのWeb上での検索エンジンを開発し、さらに2言語間の意味的なネットワークの関係を実験で調べることにした。

3. 研究の方法

本挑戦的萌芽課題では、2つのアプローチをとった。1つ目は、コーパスによるアプローチである。まず、大規模コーパスから2,136字の常用漢字の諸特性を調べた。また、中国語または韓国語と日本語の2言語間(日中と日韓)の漢字特性について辞書などを使って検討し、漢字二字からなる語彙の特性をデータベースにした。2つ目は、反応時間パラダイム、視線計測、脳波測定などの実験的アプローチである。特に、日中バイリンガル話者を対象に、多様な実験を実施した。日本語母語話者にも実験を行い、第1言語での基準の指標とした。

4. 研究成果

本課題では、4つのことを行った。第1に、

日本語では、漢字2字で構成される熟語は数多くあり、国語辞典の見出し語のうち、約70%を占めている。現代の韓国語では、漢字をほとんど使用していないものの、漢字表記由来の語が多く存在している。また、中国語では漢字が母語として使用されている。そのため、韓国語または中国語を母語とする日本語学習者は日本語の漢字語の習得において有利であると言われている。しかし、漢字表記を基本とする語彙であっても、品詞や文法的な用法が異なる場合がある。そこで、韓国語または中国語を母語とする日本語学習者のために、日本語、韓国語および中国語の3言語で、書字が類似した二字漢字語(日韓中同形二字漢字語)の品詞情報を記録したデータベースを作成した。これは、『ことばの科学』(2014年)の第27号の特集号に掲載した。

国際交流基金・日本国際教育支援協会(2007)の『日本語能力試験出題基準』(2007,改訂版)の<文字・語彙>の4級,3級,2級の語彙から漢字2字で構成される語(以下,二字漢字語)を全て抽出した。その総計は2,060語になった。これらの語彙について、日本語国語辞書『新明解国語辞典』『角川新国語辞典』『岩波国語辞典』『明鏡国語辞典』『旺文社詳解国語辞典』の5種類の辞典に掲載された品詞情報を調べて、「名詞」「動詞」「形容詞」「副詞」などの品詞を記入した。まず、1985年から1998年までの14年分の朝日新聞記事から抽出した約36万語の語彙使用頻度データ(天野・近藤,2000)から得た情報である。次に、2000年から2010年までの11年分の毎日新聞のコーパスにおける語彙の出現頻度を、形態素解析エンジンMecab 0.991を使って独自に検索して記録した。さらに、『日本語能力試験出題基準』(2007)における各語の配当級を記録した。次に、2種類の新聞コーパスから語彙の使用頻度を記載した。データベースは以下のように構成されている。

- 1 見出し語の番号
- 2 日本語の見出し語の漢字表記
- 3 日本語の読み表記
- 4 新明解国語辞典(第7版)【机上版】
- 5 角川新国語辞典(初版)
- 6 岩波国語辞典(第6版)
- 7 明鏡国語辞典(第2版)
- 8 旺文社詳解国語辞典(初版)
- 9 朝日新聞(1985-1998)の使用頻度
- 10 毎日新聞(2000-2010)年の使用頻度
- 11 日本語能力試験出題基準の級
- 12 韓国語の漢字表記
- 13 漢字のハングル表記
- 14 動詞能動-hada
- 15 動詞受動-doeda
- 16 形容詞-hada
- 17 接尾辞「-的(-jeog)」
- 18 中国語の漢字表記(簡体字)
- 19 中国語の読み表記(ピンイン)

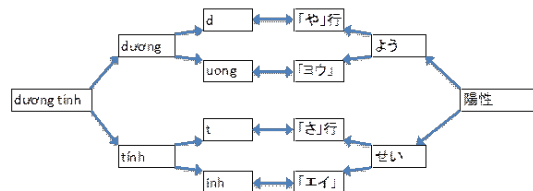
- 20 現代漢語辞典(第5版)
- 21 現代漢語規範辞典(第1版)
- 22 『漢語辞典』と『規範辞典』品詞一致

さらに、このデータベースをWeb上で自由に利用可能な検索エンジンを完成した。この検索エンジンは、同形二字漢字語の日韓中の3言語間の品詞性およびそれに関連する情報(朴・熊・玉岡, 2014; 熊・玉岡, 2014)を、整理したものである。この検索エンジンは、中国語あるいは韓国語を母語とする日本語学習者や研究者を対象としている。エンジンの構築には、MongoDB, Flask, AngularJSなどの最新の技術を使用した。朴・熊・玉岡(2014)は、日韓中3言語の8冊の国語辞書をもとに、日本語能力試験の4級から2級までの2,060語の二字漢字語の品詞情報を中心としたデータベースを作成している。しかし、このデータベースは23列(ID番号を除く)の情報からなるExcelの表形式のデータであり、そのままでは閲覧・検索がし難い。そこで、このデータベースを広く一般に利用してもらえるように、多様な検索機能を備えた検索エンジンを開発し、Web上で公開した(<http://kanjigodb.herokuapp.com>)。本検索エンジンでは、データを6種類のデータのグループに分け、分かりやすいカードの形で表示した。そして、(1)単語を個別に確認するのに適した「シンプル検索」および(2)一定の特徴を持つ単語を選び出すのに適した「詳細検索」、の2つの検索機能を構築した。この説明論文(于・玉岡, 2015)を『ことばの科学』(2015年)の第29号に掲載した。



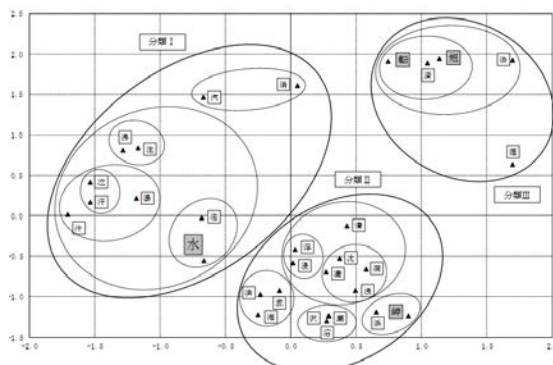
さらにこのデータベースに、ベトナム語の発音や意味を追加した。日本語の漢字語は、ベトナム語の形態素と類似している場合がある。例えば、ベトナム語の漢越語の *duong tinh* 「陽性」1音節 *duong* 「陽」の語頭は *d* で、語尾は *uong* である。一方、日本語の漢字音読みの語頭は「や」行から始まり、語尾規則を応用し、日本語の漢字音読みの語尾

は「ヨウ」で終わり、語頭と語尾の音韻を合わせ、推測すると「ヨウ」になり、漢字で表す「陽」の音読みにあたる。2音節(2つ目の漢字)の *tinh* 「性」の語頭は *t* で、語尾は *inh* である。日本語の「性」の発音は *sei* で、規則にあっている。以下は、*duong tinh* 「陽性」のベトナム語の音節と日本語の漢字の関係を図にしたものである。



第2に、2000年から2010年までの毎日新聞のコーパスを利用して、新しい常用漢字2,136字の漢字特性および二字熟語のデータベースを完成して、Web上で検索エンジン(www.kanjidatabase.com)を公開した。さらに、このデータベースについて英語で紹介した論文は、*Psychological Research* に採択され、2016年に掲載予定である(Tamaoka, et al., 2016)。これにより、世界中の研究者がこのデータベースにアクセスできるようになる。

第3に、ネット上で使える部首と漢字の関係を実験するための漢字意味空間実験ソフトを作成した。これにより、このソフトは、世界中どこでも実験に使用できるようになった。このサイト上の実験プログラムは、英語などのヨーロッパ言語はもちろん、ベトナム語、モンゴル語など漢字を母語として使用していない日本語学習者の漢字の習得のプロセスを考察するのに適している。例えば、サンズイを中心とした漢字の意味空間であれば、以下のように「水」から意味的に離れていく距離を描くことができる。さらに、クラスタ分析を使うことで、意味的に類似した漢字をグループ化することもできる。非漢字圏の日本語学習者であれば、漢字学習を始めて、徐々に漢字の学習が進んでいく様子を描くことも可能である。

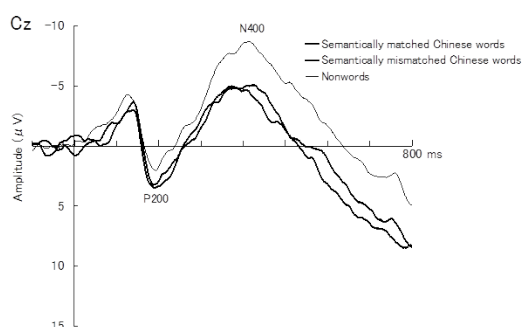


第4に、第1言語の語彙が、学習対象の第2言語の日本語を処理している場合でも、非選択的に活性化されることを脳波(ERP);

event-related potentials) 実験で実証した (Tamaoka, et al., 2016)。具体的には、以下の1つの文について3つの条件をセットとして、実験刺激を作成した。

- (1) 友達からのプレゼントは手表でした。
- (2) 友達からのプレゼントは公里でした。
- (3) 友達からのプレゼントは戸治でした。

(1)の条件では、中国語の「手表」は日本語の「時計」であり、日本語の文そのものは中国語を日本語で解釈すると意味が成り立つ。(2)の条件では、「小里」は日本語では「キロメートル」の意味になり、意味が成り立たない。(3)の条件の「戸治」は日中ともに無意味語である。脳波をこれらのターゲット語の位置から測定すれば、仮に中国語の意味が活性化されそのまま日本語文の理解に使われるなら、「友達からのプレゼントは時計でした」と解釈されるであろう。しかし、(2)の条件では、日本語では意味をなさない文となり、N400が観察されることが予想された。(3)の条件は無意味語であるため、やはりN400が予想される。実験の結果は以下の図のようになった。



中国人日本語学習者の文処理における脳波において、N400が観察されたのは、(3)の無意味条件だけであった。これは、(1)と(2)の中国語で意味のある語は、日本語文を処理していても自動的に活性化されることを意味している。しかし、(2)のような日本文に当てはまらない文でも語の活性化が起こっていると思われるものの、N400が惹起されないため、日本語での文理解にまでは至っていないのではないかと考えられる。

同様に、第1言語の中国語でも以下のように同じ条件の文を作成して、脳波を測定した。

- (4) 这位是我的友达。(「友達」)
- (5) 这位是我的风邪。(「風邪」)
- (6) 这位是我的中克。(無意味語)

結果は、日本語の場合と同じで、無意味語の条件のみでN400が観察された。

以上のように、中国語でも日本語でも、非選択的に活性化された語は、ターゲット文の意味的な処理そのものには影響しないとい

う結果が得られた。文の意味処理のレベルでカット・オフ機能が働くのではないかという仮説を提案した。

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 21 件)

- ① 玉岡賀津雄 (2016). 共起表現研究のためのコーパス検索入門. レキシコンフォーラム, 7, 239-264. 査読有.
- ② Tamaoka, K., Makioka, S., Sanders, S., & Verdonshot, R. G. (2016, on-line fast). www.kanjidatabase.com: a new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychological Research*. 査読有. doi: 10.1007/s00426-016-0764-3
- ③ Tamaoka, K., Miyatani, M., Zhang, C., Shiraiishi, M., & Yoshimura, N. (2016). Language-non-selective lexical activation without its use for sentential interpretation: An event-related potential (ERP) study on the processing of L1 Chinese and L2 Japanese sentences. *Open Journal of Modern Linguistics*, 6, 148-159. 査読有. doi: 10.4236/ojml.2016.62015
- ④ 難波えみ・玉岡賀津雄 (2016). 様態と結果の副詞的表現と動詞の共起パターンに関するエントロピーと冗長度を指標にした検討. *計量国語学*, 30(4), 195-209. 査読有.
- ⑤ 早川杏子・魏志珍・初相娟・玉岡賀津雄 (2016). 日本語聴解能力測定のためのテスト開発と信頼性の検討—中国語および韓国語を母語とする日本語学習者のデータによる評価—. *関西学院大学日本語教育センター紀要*, 5, 31-45. 査読無.
- ⑥ Shibasaki, H., Tokimoto, S., Ono, Y., Inoue, T. & Tamaoka, K. (2015). English reading comprehension by Japanese high school students: structural equation modeling including working memory and L1 literacy *Open Journal of Modern Linguistics*, 5, 443-458. 査読有. doi: 10.4236/ojml.2015.55039
- ⑦ 大和祐子・玉岡賀津雄・初相娟 (2015). 中国人日本語学習者の人を表す接尾辞の派生語習得に影響する要因. *小出記念日本語教育研究会*, 23, 33-45. 査読有.
- ⑧ Goss, S., & Tamaoka, K. (2015). Predicting lexical accent perception in native Japanese speakers: An investigation of acoustic pitch

- sensitivity and working memory. *Japanese Psychological Research*, 57(2), 143-154.
査読有.
doi: 10.1111/jpr.12076
- ⑨ Verdonschot, R. G., Lai, J., Feng, C., Tamaoka, K., & Schiller, N. O. (2015). Constructing initial phonology in Mandarin Chinese: syllabic or sub-syllabic? A masked priming investigation. *Japanese Psychological Research*, 57(1), 61-68.
査読有.
doi: 10.1111/jpr.12064
- ⑩ 早川杏子・玉岡賀津雄 (2015). 改訂版・構造分類による日本語文法知識テストの開発ー中国人日本語学習者のデータによるテスト評価ー. *ことばの科学*, 29, 5-24.
査読無.
- ⑪ 于ショウイン・玉岡賀津雄 (2015). 日韓中同形二字漢字語の品詞性ウェブ検索エンジン. *ことばの科学*, 29, 43-61.
査読無.
- ⑫ 小森和子・玉岡賀津雄・斉藤信浩・宮岡弥生 (2014). 第二言語として日本語を学ぶ中国語話者の日本語の漢字語の習得に関する考察. *中国語話者のための日本語教育研究*, 5, 1-16.
査読有.
- ⑬ 張セイイ・木山幸子・玉岡賀津雄 (2014). 中国語を母語とする日本語学習者による流行語の理解. *Language Education & Technology (外国語教育メディア学会機関誌)*, 51, 31-49.
査読有.
- ⑭ Tamaoka, K. (2014). The Japanese writing system and lexical understanding. *Japanese Language and Literature (The American Association of Teachers of Japanese, AATJ)*, 48, 431-471.
査読有.
- ⑮ Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R.G. (2014). Is pitch accent necessary for comprehension by native Japanese speakers? An ERP investigation. *Journal of Neurolinguistics*, 27, 31-40.
査読有.
doi: 10.1016/j.jneuroling.2013.08.001
- ⑯ 朴ソンジュ・熊可欣・玉岡賀津雄 (2014). 同形二字漢字語の品詞性に関する日韓中データベース. *ことばの科学*, 27, 53-111.
査読無.
- ⑰ 熊可欣・玉岡賀津雄 (2014). 日中同形二字漢字語の品詞性の対応関係に関する考察. *ことばの科学*, 27, 25-51.
査読無.
- ⑱ 朴ソンジュ・熊可欣・玉岡賀津雄 (2014). 同形二字漢字語の品詞性に関する日韓中データベース概要. *ことばの科学*, 27, 3-23.
査読無.
- ⑲ 宮岡弥生・玉岡賀津雄・酒井弘 (2014). 日本語の文法能力テストの開発と信頼性: 日本語学習者のデータによるテスト評価ー. *広島経済大学研究論集*, 36(4), 33-46.
査読無.
- ⑳ Tamaoka, K., & Kiyama, S. (2013). The effects of visual complexity for Japanese kanji processing with high and low frequencies. *Reading and Writing*, 26(2), 205-223.
査読有.
doi: 10.1007/s11145-012-9363-x
- ㉑ Verdonschot, R. G., Nakayama, M., Zhang, Q., Tamaoka, K., & Schiller, N. O. (2013). The Proximate Phonological Unit of Chinese-English Bilinguals: Proficiency Matters. *PLoS ONE*, 8(4), e61454.
査読有.
doi: 10.1371/journal.pone.0061454
- ㉒ Verdonschot, R. G., La Heij, W., Tamaoka, K., Kiyama, S., You, W. P., & Schiller, N. O. (2013). The multiple pronunciations of Japanese kanji: A masked priming investigation. *The Quarterly Journal of Experimental Psychology*, 66, 2023-2038.
査読有.
doi: 10.1080/17470218.2013.773050
- [学会発表] (計 14 件)
- ① Tamaoka, K., Yu, S. & Mansbridge, M. P. (2015). An eye-tracking investigation of the processing of nominative/genitive alternations in Japanese. AMLaP (Architectures and Mechanisms for Language Processing) 2015, September 3-5 2015, University of Malta Valletta Campus, Valletta Malta.
- ④ Kiyama, S., Kunimi, M., Verdonschot, R., Tamaoka, K., & Nakai, T. (2015). The dmPFC gray matter density enhances elderly's positivity effect in emotional sentence retrieval. June 14-18, 2015, Hawaii Convention Center, Hawaii, the U.S.A.
- ⑤ Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2013). The new web-accessible Japanese Joyo kanji database. 2013 Psychonomic Society Annual Meeting. November 14-17, 2013, Sheraton Centre Hotel in Toronto, Ontario, Canada.
- ⑥ Tamaoka, K., Mansbridge, M. P., Lianzhen, Z., & Verdonschot, R. G. (2013). When I am in Paris, I bought some suitcases?: The insensitivity of native Japanese speakers to English tense inconsistency. International Conference on Multilingualism - Linguistic Challenges and Neurocognitive Mechanisms. 24-25 October, 2013. McGill University, Montreal, Canada.
- ⑦ Mansbridge, M. P., Hayakawa, K., Tamaoka, K. (2013). Animacy effects on the processing

of Japanese sentences among second language speakers. AMLaP (Architectures and Mechanisms for Language Processing) 2013, September 2-4, 2013, Aix-Marseille University, Marseille, France.

- ⑧ Xiong, K., Mansbridge, M.P., & Tamaoka, K. (2013). Effects of L1/L2 differences in verbalization of cognates within passive voice processed by native Chinese speakers learning Japanese. AMLaP (Architectures and Mechanisms for Language Processing) 2013, September 2-4, 2013, Aix-Marseille University, Marseille, France.
- ⑨ 于劭贇・熊可欣・早川杏子・玉岡賀津雄 (2015). 学習中級段階からの日本語通訳者養成のための翻訳 N-back 訓練法—台湾人日本語学習者の縦断的教育効果測定による妥当性の検証—. 2015 年度日本語教育学会秋季大会. 2015 年 10 月 10 日から 11 日まで, 沖縄国際大学.
- ⑩ 張セイイ・早川杏子・玉岡賀津雄 (2015). 中国人日本語学習者による動詞由来の複合名詞の理解. 2015 年度日本語教育学会春季大会. 2015 年 5 月 30 日から 31 日まで. 武蔵野大学有明キャンパス.
- ⑪ 熊可欣・玉岡賀津雄 (2015). 中国人日本語学習者による日中同形語の統語情報への非選択的アクセス. 2015 年度レキシコン・フェスタ 3. 2015 年 2 月 1 日, 国立国語研究所 2F 講堂.
- ⑫ 難波えみ・玉岡賀津雄 (2015). 副詞と動詞の共起表現の多様性に関するコーパス研究. 2015 年度レキシコン・フェスタ 3. 2015 年 2 月 1 日, 国立国語研究所 2F 講堂.
- ⑬ 難波えみ・玉岡賀津雄 (2014). 副詞と動詞の共起頻度によるエントロピーと冗長度の算出. 計量国語学会第 58 回大会. 2014 年 9 月 20 日(土), 東洋大学白山キャンパス 6 号館 2F.
- ⑭ 熊可欣・玉岡賀津雄・早川杏子・Verdonschot, R. G. (2014). 日中バイリンガルの同形同義語の統語情報への非選択的アクセス. 日本第二言語習得学会第 14 回年次大会. 2014 年 5 月 31 日, 関西学院大学西宮上ヶ原キャンパス.

[図書] (計 3 件; 分担執筆)

- ① 玉岡賀津雄・小泉政利 (2016). 第 8 章 言語理解. 小泉政利 (編), ここから始める言語学プラス統計分析 (pp. 129-145). 東京: 共立出版.
- ② Tamaoka, K. (2015). Chapter 18: Processing of the Japanese language by native Chinese speakers. In Mineharu Nakayama (Ed.), Handbooks of Japanese Psycholinguistics (pp. 283-332), Berlin, Germany: De Gruyter Mouton.
- ③ 熊可欣・玉岡賀津雄 (2015). 第 1 章の 1 (t 検定)と 2(方差分析). 館岡洋子・于康 (編) 日語教育—研究方法と応用 (pp.

24-43). 北京: 高等教育出版.

[その他]

ホームページ

① 同形二字漢字語の品詞性に関する日韓中データベースの Web 検索エンジンをつきサイト <http://kanjigodb.herokuapp.com/>

② 常用漢字 2,136 字およびそれらの二字熟語の Web 検索エンジンをつきサイト(英語) <http://www.kanjidatabase.com/>

③ 研究業績を掲載したサイト <http://www.lang.nagoya-u.ac.jp/~ktamaoka/scholarly/index.html>

6. 研究組織

(1) 研究代表者

玉岡 賀津雄 (TAMAOKA, Katsuo)
名古屋大学・大学院国際言語文化研究科・教授

研究者番号: 70227263

(2) 研究分担者

木山 幸子 (KIYAMA, Sachiko)
三重大学・教養教育機構・特任講師
研究者番号: 10612509