科学研究費助成事業　　研究成果報告書

平成　２６　年　　６　月　２３　日現在

機関番号：　３４３１５
研究種目：　若手研究(A)
研究期間：　2011～2013
課題番号：　２３６８２００６

研究課題名（和文）日本人英語学習者のための可変長で不連続性を許した学術連語項目リストの階層的構築

研究課題名（英文）Hierarchically Creating Variable-Length and Non-Contiguous Academic Expression Lists for Japanese Learners of English

研究代表者
田中　省作（Tanaka, Shosaku）

立命館大学・文学部・教授

研究者番号：００３２５５４９

交付決定額（研究期間全体）：（直接経費）　　10,600,000円、（間接経費）　　3,180,000円

研究成果の概要（和文）：　本研究は，語と文の中間的な要素である広義の連語項目（緩やかな語の連なり，MWE）リストを，日本人大学生の英語学習を念頭に中高教育との接続性を考慮し，各種分野のコーパスに基づき分野階層的に構築を試みた．　MWEには，統語情報を考慮し可変長で不連続なものも対象としたような方法を導入し，本研究で使用したプログラム群については，Webで公開した．コーパスが整備できれば比較的容易に構築が可能である．

研究成果の概要（英文）：This project created academic multi-word expression (MWE) lists for Japanese learners of English considering the junior high school and high school curricula in Japan, hierarchically. The method adopted for this construction can identify variable-length and non-contiguous expressions, such as MWEs. The programs that can be used with the method are released on the Internet; moreover, they allow us to create MWE lists with relative ease only if we prepare a corpus as the source data for the method.

研究分野：　人文学

科研費の分科・細目：　言語学・外国語教育

キーワード：　連語項目　コーパス　言語処理

１．研究開始当初の背景

近年，言語習得や英語教育研究において，“on the whole” や “take *X* into account/consideration”といった語と文の間に位置づけられる語の緩やかな連なり（本研究では広義に捉え，MWE と記す）は，重要な言語知識の一つとして注目されている. MWE にも，語彙や構文同様，分野に強く依存するものがある．たとえば，“let *X* and *Y* be *A* and *B*, respectively”は, 数学および周縁の分野では有用かつ重要な MWE である．そのような分野の特徴語や MWE の抽出に，近年の語彙研究では JACET8000 といった成果を上げたコーパスに基づく試みがある．その基本アイディアは「語の *n* 連鎖（*n*-gram）ごとに，特定分野コーパスと均衡（一般英語）コーパス内の頻度を統計的に比較し，特定分野に有意に頻出する *n*-gram を『特定分野の MWE』とする」というものである．このような方法論には，次のような問題がある．

(1) 均衡コーパスを基準としていること

従来研究の多くは，特定分野と比較する均衡コーパスとしてイギリス英語を代表するコーパス（British National Corpus: BNC）を採用している．しかし，日本の中高英語教育を経た学習者にとって, BNC に代表されるような英語が一般的であるとは言い難い．したがって，BNC と特定分野の対比に基づいた MWE リスト（以後，簡単に表現リストと記す）は，日本人の英語経験や知識から乖離したものとなり，それらの教育的位置づけが難しくなる．

さらに，たとえば「情報工学」といった狭小な特定分野を直接標的とし，表現リストを作成した場合, BNC と特定分野間の差が非常に大きいために,「情報工学」の表現リストには学術/工学/情報工学などそれぞれの MWE が混在し，非段階的で混沌としたものとなる．

(2) MWE を固定長で連続的な表現として捉えていること

多くの従来研究は，技術的な簡便さから MWE を *n*-gram で粗く捉え，整備している．このように MWE を考えてしまうと，*n* が異なる *n*-gram は素直に比較できないため, 結局, MWE の長さ（*n*）を事前に固定せざるを得ない．

その上，*n*-gram は連続的な表現しか捕捉できないので，“take *X* into account”といった不連続に関係づけられる MWE は，必然的に洩れてしまう．

２．研究の目的

本研究の主目的は，前節で述べた 2 つの問題の解決を目指し，特定分野もしくはそれに対応する組織等で，大学生以上の日本人英語学習者のための学術的な MWE を階層的に整備することである．そのために，次のような 3 つの小テーマを設定した．

(1) 可変長で不連続性を許した MWE 抽出法の導入
(2) 日本の中高英語との連続性，および分野の階層性を考慮した表現リストの生成法の検討
(3) 新しい MWE の発見と基本的性質の解明

なお，ここで整備した表現リストや，開発したプログラム群を，可能な限り公開することも副次的目的である．言語資源等の要件を満たせば，各所で比較的容易に同様の処理が可能となる．

３．研究の方法

(1) 方針

①学術分野に関する情報が付された論文データ

学術分野に関する情報が付されたような論文データには PERC コーパスなどがあげられるが，ほとんどは著作権や使用条件の問題で，直接本研究で解析，利用することはできない．また，特定機関で展開されている研究分野に基づき，分野別に同種データを構築した事例として，京都大学 田地野彰教授らが推進した京都大学学術論文コーパス(田地野他, 2008)がある．そのスキームは本研究が意図するデータ整備のモデルとなるものである．しかし，同規模のデータを本研究で一から構築することは容易ではない．本研究では, 推進者の専門領域である情報学分野に限定した小規模な論文データを蓄積し，まずはそれをパイロット・データとした．

また，このようなデータの代替として，本研究では，近年，大学をはじめとした多くの研究機関が構築し，自組織の研究者らが執筆した論文や記事等の著作物を電子的に蓄積, 公開している機関リポジトリを活用することとした．具体的には，機関リポジトリから得られる組織情報を，学術分野に粗く対応付け，各組織でアーカイブされている論文を，当該分野の論文データとみなし，表現リストの作成に活用する．なお，機関リポジトリの言語資源としての活用は，本課題に遅れて別課題としても採択されることとなった．連携関係の下，両課題を進めた．

②可変長で不連続性を許した MWE 抽出

可変長で不連続な MWE の抽出には，名古屋大学 松原茂樹准教授らが提案した方法(松原他, 2008)を基本とした．詳細は次項目(2)で述べるが，本研究では，抽出した表現リストを最終的に関連分野の英語識者が確認・編纂することを念頭に置き，文中の依存構造や組

織の階層関係を勘案した方法に改変し，適用した．

③表現の階層的整備

機関リポジトリから得られる組織情報を参照し，その階層性を表現リストに反映させた．なお，表現リストは主に日本人英語学習者の活用を想定し，この階層の最上部に日本の中高英語を置き，それに対応する著作物は，中高の英語教科書や参考書とした．たとえば，「A 大学 B 学部 C 学科」の場合，「A 大学の論文データ vs. 中高英語データ」に基づく「A 大学の表現リスト」,「B 学部の論文データ vs. 中高英語+B 学部をのぞく A 大学の論文データ」に基づく「B 学部の表現リスト」,「C 学科の論文データ vs. 中高英語+C 学科をのぞく A 大学の論文データ」に基づく「C 学科の表現リスト」という具合に，3 つのリストを作成する．「A 大学⇒B 学部⇒C 学科」の方向性は，「A 大学内における EGAP（一般学術目的の英語）から ESAP（特定学術目的の英語）」におおむね対応する．

このような表現リストの間で，組織の階層関係を考慮し，調整を加える．上部組織の表現リストである一定の上位部分に列挙される表現は，それよりも下部の組織の表現リストでは含めないよう，後処理を行う．

(2) 手順

機関リポジトリに含まれる英語著作物を事前に組織階層別に分け，それぞれで次のように英語学術表現リストを生成した．

①チャンク構造と依存関係の同定

各文を構文解析し，チャンク構造と依存関係を同定する．本研究で注目するチャンク構造は，補文標識（LC）と内部に句構造を含まない基本名詞句（NC）とした．各語は動詞の分詞形を除き原形表記に統一した後に，名詞・動詞といった浅い品詞レベルで細分化する．冠詞・数字・記号はそれぞれ<D>・<C>・<S>に置換する．

たとえば，“This paper shows a new method to solve it.”のチャンク構造は，

$[_{NC}$ <D> paper_N $]$ show_V
$[_{NC}$ <D> new_J method_N $]$
$[_{LC}$ to_T $]$ solve_V $[_{NC}$ it_P $]$ <S>

となる．ここで，x_p は原形が x で品詞が p の語，$[_Y \boldsymbol{y}]$は語列 $\boldsymbol{y}$ が Y 句で，N・V・J・P・T はそれぞれ名詞・動詞・形容詞・代名詞・TO を表す．なお，文構造を成していないものは分析対象から除く．

例文中に含まれる依存関係は，次の通りである．

1: <D> → paper_N
2: paper_N → show_V
4: <D> → method_N
5: new_J → method_N
6: method_N → show_V
7: to_T → method_N
8: solve_V → to_T
9: it_P → solve_V
10: <S> → show_V

ここで，“x→y”は x が y に係っていること，最左の番号は x の文中での出現位置を表している．

②チャンク構造を考慮し，n-gram を生成

文の前後に文頭・文末を表す特殊記号@を $n-1$ 個付加し，n-gram を生成する．その際，NC と LC を跨ぐ場合には，それらの語列を一旦<NC>,<LC>という 1 記号に置換した列も別途考え，それぞれで n-gram を生成する．

さきほどの例で $n=3$ の場合，

@ <D> paper_N
<D> paper_N show_V
paper_N show_V <D>

に加え，

@ <NC> show_V
<NC> show_V <NC>
show_V <NC> <LC>

なども生成される．n も 2〜10 という具合にある一定の範囲で動かし，これらを累積的に計数する．

計数は，次の 2 つの観点で行う．一つは，全ての n-gram をそれぞれ素直に計数するもので，次項③の連接確率の算出に用いる頻度（単純頻度）である．もう一つは，依存関係を考慮して計数するもので，次のような条件を満たす n-gram $\boldsymbol{x}$ のみ計数対象とする．

- $\boldsymbol{x}$ に内容語が存在する．
- $\boldsymbol{x}$ 内の全ての非内容語の係り先が，$\boldsymbol{x}$ 内に存在する内容語である．

たとえば，上記の“<D> paper_N show_V”や“<NC> show_V <NC>”は上記条件を満たし，計数対象とする．その一方で，“paper_N show_V <D>”は<D>が n-gram 内で依存関係を結んでいないため，計数対象とはならない．また，上記の例ではないが，“<NC> of_P <NC>”は内容語を含んでいない，“show_V <NC> of_P”で of_P が show_V ではなく<NC>に係る場合は，やはり計数対象とはならない．このようにして得られる頻度を，依存関係を考慮した頻度とよぶ．

③スコアリング

生成された各 n-gram $\boldsymbol{x}$ に対して，次のようにスコアを与える．

$$\text{score}(\boldsymbol{x}) = \log[f_d(\boldsymbol{x})]\ell(\boldsymbol{x})(H_L(\boldsymbol{x})+1)(H_R(\boldsymbol{x})+1)$$

$f_d(\boldsymbol{x})$ は $\boldsymbol{x}$ の依存関係を考慮した頻度，$\ell(\boldsymbol{x})$ は $\boldsymbol{x}$ に内含される語・基本句数である．$H_L(\boldsymbol{x}), H_R(\boldsymbol{x})$ は前後に連接する語のエントロピーで，次のように与える．

$$H_\alpha(\boldsymbol{x}) = -\sum\nolimits_y P_\alpha(y \mid \boldsymbol{x}) \log P_\alpha(y \mid \boldsymbol{x})$$

$\alpha \in \{L, R\}$ で，$P_L(y \mid \boldsymbol{x})$ は y が n-gram $\boldsymbol{x}$ に左連接する確率，$P_R(y \mid \boldsymbol{x})$ は右連接する確率で，それぞれ次のように与える．

$$P_L(y \mid \boldsymbol{x}) \triangleq f(y\boldsymbol{x}) / f(\boldsymbol{x})$$
$$P_R(y \mid \boldsymbol{x}) \triangleq f(\boldsymbol{x}y) / f(\boldsymbol{x})$$

なお，$f(\boldsymbol{x})$ は $\boldsymbol{x}$ の単純頻度である．

④フィルタリング

score($\boldsymbol{x}$)に基づき n-gram $\boldsymbol{x}$ を学術表現の候補として抽出していく．その際，次のいずれかの条件が成立する $\boldsymbol{x}$ は，抽出の対象外とする．

- $f(\boldsymbol{x})<T$ である．
- 自組織よりも上部組織 i の表現リストで上位 β_i%までに抽出済である．
- score($\boldsymbol{x}'$) > score($\boldsymbol{x}$) かつ $\boldsymbol{x}$ を内容語に関して完全に包含するような $\boldsymbol{x}'$が抽出済である．

(3) 実験

情報学関係の小規模な論文パイロット・データ，学術分野付きの論文コーパス（未公開），複数の研究大学の機関リポジトリに対して，本手法を適用した．

ここでは，九州大学の機関リポジトリへの適用事例について簡単に述べる．2012 年 7 月時点の九州大学機関リポジトリ QIR に含まれる英語著作物 5,838 点のうち，形態素数が 2,000〜10,000 の論文 2,965 点を対象とした．延べ形態素数は 15,146,153 である．これらを学部・研究科レベルに相当する 27 部局に細分化し，九州大学全体に加え，それぞれの部局の表現リストを作成する．中高英語の著作物には，平成 14 年度版検定済中高英語教科書（中学 7 シリーズ，高校 28 シリーズ）の本文部分を活用した．延べ形態素数は 736,933 である．n=3〜7 で，最低頻度 T を 5，$\beta_{\text{中高英語}}$=10, $\beta_{\text{九州大学}}$=1 とした．なお，形態素解析には TreeTagger，構文解析には Charniak parser を利用した．

その結果，中高英語・九州大学全体（概ね学術共通に対応）・部局毎の表現リストが，実時間で生成された．紙面の関係上，ほんの一部になるが，九州大学全体の上位 5 位を示すと，

based on <NC>
such as <NC>
a/the set of <NC>
due to <NC>
according to <NC>

であった（品詞情報は省略）．なお，部局に対応する分野は，それぞれ排他的ではない上に，九州大学の組織の特殊性が反映されている．その結果，それら全てをそのまま素直に学術分野下の小分野と対応づけられない場合もあり，表現リストを洗浄するような後処理の必要性が明らかとなった．

松原茂樹・酒井祐太・小澤俊介・杉木健二 (2010) 学術論文からの英語表現集の自動生成, 第 7 回情報プロフェッショナルシンポジウム, pp.41-44.
田地野彰・寺内 一・金丸敏幸・マスワナ紗矢子・山田 浩 (2008) 英語学術論文執筆のための教材開発に向けて: 論文コーパスの構築と応用, 京都大学高等教育研究開発推進センター, Vol.14, pp.111-121.

４．研究成果

本研究は，主に次のような成果を上げた．今後の展望，課題も併記する．

(1) 学術表現の階層的構築方法の提案とその試行

可変長で不連続性を許すような MWE 抽出法を導入し，中高英語から学術分野，さらにその下位の小分野に至るまで，階層的に表現リストを構築する方法を示し，実際に試作を行った．九州大学の機関リポジトリを対象としたものについては，部分的な公開を予定している．

(2) 大規模な論文データの代替としての機関リポジトリの活用と基本的傾向

学術分野に関する情報が直接付与され，本研究に適用できる大規模な論文データの取得が難しいことや，推進者が他課題で電子図書館研究にかかわりがあったことから，本研究の一部では機関リポジトリをその代替として活用した．機関リポジトリから得られる表現のうちスコア上位となる表現は，比較的学術分野共通に挙げられるものと一致している．部局別に細分化されても，伝統的な学術分野で，対応する部局の論文が十分に蓄積されている場合にも同様の傾向が見られた．

このように機関リポジトリの言語資源としての応用可能性を示唆した．一方，本研究においては，生成した表現リストの精査から，機関リポジトリが反映している研究機関の特殊性を排除する必要があることが明らかとなった．このような処理が可能となれば，言語資源としての機関リポジトリの有用性や汎用性はさらに高まるものといえる．本研究の直接的な後継タスクとして，他の機関リポジトリから生成される表現リストとの重ね合わせなどによる洗浄化を検討している．

(3) MWE の基本的性質としての依存構造

MWE 抽出のフィルタリングにおいて，当初は品詞レベルまでの情報で MWE の言語的妥当性あるいは自然さを担保していた．最終年度，MWE 内での内容語を中心とした依存関係を強く意識し，頻度ベース（依存関係を考慮した頻度）であくまでも優先関係ではあるが，MWE の妥当性や自然さをある程度とらえることができたと考えている．ただし，このような方式の導入が最終年度後半期だったこともあり，定量的な評価はできていない．

(4) プログラム群の一部公開

本研究で開発したプログラムについては，可能な限り Web 等で公開を行った．特に，コ

ーパス研究等を志向する研究者らが抵抗なく活用できるよう，データ形式等はやや洗練さに欠け，冗長性は許しつつも，単純なテキスト・エディタ等で編集しやすいようにした．

なお，3.(2)で述べた依存関係を考慮した制約処理については，最終年度の後半期に導入した手続きであったため，これらを内在した関連プログラムは公開できていない．また，機関リポジトリに関連するプログラム群については，他課題で整理，公開を進める．

５．主な発表論文等

〔雑誌論文〕（計 12 件）

① 田中省作: 英語学術表現リストの階層的構築 -言語資源としての機関リポジトリの新しい活用-, 立命館文學, 第 636 巻, pp.87-97, 2014 年（査読無）
② 田中省作: ジニ係数に基づいたランダムフォレストにおける部分木の重要度, 統計数理研究所共同研究リポート, 321, pp.15-27, 2014 年（査読無）
③ 小林雄一郎, 田中省作, 阿部真理子: 情報量基準に基づく習熟度尺度の再検討, 統計数理研究所共同研究リポート, 321, pp.29-43, 2014 年（査読無）
④ Koyama, Y., Tanaka, S., Miyazaki, Y., Fujieda, M.: Development of a Corpus-assisted Writing System for Research Papers by Science and Technology Students, ILAC Selections -Autonomy in a Networked World-, pp.65-67, 2013 年（査読有）
⑤ 徳見道夫, 田中省作: 大学入試センターのリスニングテストと九州大学の標準化テスト成績の連関性, 英語英文学論叢, 62, pp.19-37, 2012 年（査読有）
⑥ 多田一馬, 田中省作: 談話標識 N-gram で近似した論理展開のジャンル分析, 統計数理研究所研究リポート, 276, pp.69-79, 2012 年（査読無）
⑦ 照井路子, 田中省作: 同義の文法構造のレジスタ分析 -「与格交替」を例に-, 統計数理研究所研究リポート, 276, pp.55-67, 2012 年（査読無）
⑧ 鄭 宗田, 田中省作: CLEC に基づく中国人英語学習者の「make」の使用状況分析, 統計数理研究所研究リポート, 276, pp.43-53, 2012 年（査読無）
⑨ 宮崎佳典, 田中省作, 小山由紀江: コーパスを用いた英語技術文書作成補助ツールの評価, 統計数理研究所研究リポート, 276, pp.1-21, 2012 年（査読無）
⑩ 田中省作, 安東奈穂子, 冨浦洋一: コーパス構築と著作権 −Web を源とした質情報付き英語科学論文コーパス-, 英語コーパス研究, 第 19 号, pp.31-42, 2012 年（査読有）
⑪ 田中省作: Web コーパスの言語情報処理基盤, 英語コーパス研究, 第 18 号, pp.97-111, 2011 年（査読有）
⑫ 田中省作, 柴田雅博, 冨浦洋一: Web を源とした質情報付き英語科学論文コーパスの構築法, 英語コーパス研究, 第 18 号, pp.61-71, 2011 年（査読有）

〔学会発表〕（計 19 件）

① 田中省作: 分類型ランダムフォレストにおける部分木の重要度, 言語研究と統計 2014, 2014 年 3 月 29 日, 統計数理研究所（東京都立川市）
② Kobayashi, Y., Tanaka, S., Tomiura, Y., Miyazaki, Y., Tokumi M.: Identifying Discipline-specific Expressions Based on Institutional Repository, Digital Humanities Australasia 2014, 2014 年 3 月 19 日, The University Club of Western Australia (Perth, Autstralia)
③ Watabe, T., Miyazaki, Y., Tanaka, S.: Design of a Mathematical Expression Corpus for Word Sense Disambiguation, First IIAI International Conference on Advanced Information Technologies 2013, 2013 年 11 月 29 日, Hotel Aryaduta Jakarta (Jakarta, Indonesia)
④ 田中省作: 基本句を考慮した n-gram の計数, 英語コーパス学会第 39 回大会, 2013 年 10 月 6 日, 東北大学（宮城県仙台市）
⑤ 田中省作, 宮崎佳典, 小山由紀江, 藤枝美穂: 分野依存性を考慮した用例提示型英文書作成支援ツールの開発, 教育システム情報学会 2013 年度第 2 回研究会, 2013 年 7 月 14 日, 千歳科学技術大学（北海道千歳市）
⑥ Miyazaki, Y., Tanaka, S., Koyama, Y.: A Tool Supporting Writing Technical Documents in English Using Corpora: Retrieving Functions by Cosine Similarity and Pattern Matching, The 7th International Multi-Conference on Society, Cybernetics and Informatics (IMSCI 2013), 2013 年 7 月 10 日, Doubletree by Hilton Orlando at SeaWorld (Florida, United States of America)
⑦ 田中省作, 冨浦洋一: 機関リポジトリを活用した英語学術表現リストの階層的構築, 言語処理学会第 19 回年次大会, 2013 年 3 月 14 日, 名古屋大学（愛知県名古屋市）
⑧ 田中省作: 言語資源としての機関リポジトリ, 「計量的言語研究の諸相」第 2 回講演会, 2013 年 3 月 9 日, 北海道大学（北海道札幌市）
⑨ 田中省作, 冨浦洋一, 宮崎佳典, 小林雄一郎, 徳見道夫: 機関リポジトリを活用した部局別英語学術表現リストの作成支援, 情報処理学会第 75 回全国大会, 2013 年 3 月 8 日, 東北大学（宮城県仙台

市）

⑩ Koyama, Y., Tanaka, S., Miyazaki, Y., Fujieda, M.: Development of Corpus-assisted Research Paper Writing System for Science and Technology Students, The 5th Independent Learning Association Conference 2012 (ILAC2012), 2012 年 8 月 31 日, Victoria University of Wellington (Wellington, New Zealand)

⑪ Kobayashi, Y., Tanaka, S.: Syntax and Discourse in Good and Poor Scientific Articles, The 10th Teaching and Language Corpora Conference (TaLC10), 2012 年 7 月 13 日, University of Warsaw (Warsaw, Poland)

⑫ 田中省作: コロケーション研究における「相互情報量」, 名古屋大学大学院国際開発研究科公開講演会, 2012 年 3 月 29 日, 名古屋大学（愛知県名古屋市）

⑬ 田中省作, 小林雄一郎, 徳見道夫, 後藤一章, 冨浦洋一, 柴田雅博: 学校英文法の学参例文データベースとその応用, 情報処理学会第 93 回人文科学とコンピュータ研究会, 2012 年 1 月 27 日, 奄美市立奄美博物館（鹿児島県奄美市）

⑭ 小林雄一郎, 田中省作, 冨浦洋一: N-gram を素性とするパターン認識を用いた英語科学論文の質判定, 情報処理学会第 205 回自然言語処理研究会, 2012 年 1 月 21 日, 福岡大学（福岡県福岡市）

⑮ 小林雄一郎, 田中省作, 冨浦洋一: メタ談話標識を素性とするパタン認識を用いた英語科学論文の質判定, 人文科学とコンピュータシンポジウム, 2011 年 12 月 10 日, 龍谷大学（京都府京都市）

⑯ Miyazaki, Y., Tanaka, S., Koyama, Y.: Development and Improvement of a Corpus-based Web Application to Support Writing Technical Documents in English, International Conference of Computer Education (ICCE2011), 2011 年 12 月 2 日, Le Meridien Chiang Mai Hotel (Chiang Mai, Thailand)

⑰ 田中省作, 冨浦洋一, 徳見道夫: 学校文法に基づいた英文解析による言語データの頻度分析, 英語コーパス学会第 37 回大会, 2011 年 10 月 1 日, 京都外国語大学（京都府京都市）

⑱ 宮崎佳典, 田中省作, 小山由紀江: 技術文献コーパスを活用した英語技術文書作成支援 Web アプリケーション開発, 外国語教育メディア学会第 52 回全国研究大会, 2011 年 8 月 10 日, 名古屋学院大学（愛知県名古屋市）

⑲ 小林雄一郎, 田中省作, 冨浦洋一: ランダムフォレストを用いた英語科学論文の分類と評価, 情報処理学会第 90 回人文科学とコンピュータ, 2011 年 5 月 21 日, 同志社大学（京都府京都市）

〔図書〕（計 1 件）

① 岸江信介・田畑智司 (編)『テキストマイニングによる言語研究』 東京: ひつじ書房（「メタ談話標識を素性とするランダムフォレストによる英語科学論文の質判定」小林雄一郎・田中省作の共同執筆） (2014)

〔産業財産権〕

○出願状況（計 0 件）

○取得状況（計 0 件）

〔その他〕

http://www.cl.ritsumei.ac.jp/HAMWE/

６．研究組織

(1) 研究代表者

田中　省作（TANAKA, SHOSAKU）
立命館大学・文学部・教授
研究者番号：00325549