

文化差理解支援における可視化効果の検証

吉野 孝^{1,a)} 宮部 真衣^{2,b)}

概要：多言語間コミュニケーションにおいて、同一の単語を用いて会話をしている場合でも、相手の文化について十分に理解していないために、誤解が生じる可能性がある。我々は、文化差の検出手法を提案し、文化差可視化システムの開発を行っている。文化差可視化の目的は、文化差理解支援である。文化差の可視化の方法としては、お互いの言語における語句の説明を、文字（説明文）を用いて行う方法や画像を用いて行う方法が考えられる。本稿では、文化差理解支援における説明文および画像による文化差可視化の効果の検証を行った。

キーワード：文化差検出，可視化，多言語間コミュニケーション

Effect of Visualization on Cultural Understanding Support System

TAKASHI YOSHINO^{1,a)} MAI MIYABE^{2,b)}

Abstract: There is a possibility that the misunderstanding is caused in multilingual communications, because people cannot understand enough other culture even when talking by using the same word. We propose cultural difference detection methods and develop a cultural differences visualization service. The purpose of cultural differences visualization is to support cultural differences understanding. There are two methods to visualize cultural differences using words and phrases and using images for each word in each language. In this paper, we verified the effect of visualization of cultural difference understanding supports between using words and phrases and using image.

Keywords: cultural difference detection, visualization, multilingual communication

1. はじめに

多言語間コミュニケーションにおいて、同一の単語を用いて会話をしている場合でも、相手の文化について十分に理解していないために、誤解が生じる可能性がある [1]。文化の違いに基づく誤解を回避するためには、用いている単語に文化差があることを、話者に認識させる必要がある。しかし、相手の文化に関する十分な知識が必要となるため、文化差の有無の判断は容易ではない。我々はこれまでに、文化差検出に関する検討を行って

きた。語句に関する文化差の有無については、Wikipedia の項目数に基づく手法、記事内の国名・言語名数に基づく手法、記事における執筆者の意図に基づく手法を統合することにより、比較的高い精度での判定を可能にした [2]。さらに、日本語版および中国語版 Wikipedia における相互言及の有無に着目することで、従来手法よりも精度よく文化差を検出でき、文化差の検出を日本語と中国語以外へ適用できる手法を提案した [3]。

文化差可視化の目的は、文化差理解支援である。我々は、文化差検出結果をもとに、文化差の可視化システムの構築を行っている [2], [4]。文化差の可視化方法としては、お互いの言語における語句の説明を、文字（説明文）を用いて行う方法や画像を用いて行う方法が考えられる。

本稿では、文化差理解支援における、説明文および画像による文化差の可視化の効果の検証を行った。

¹ 和歌山大学 システム工学部
Faculty of Systems Engineering, Wakayama University,
Wakayama-shi, Wakayama 640-8510, Japan

² 東京大学 知の構造化センター
Center for Knowledge Structuring, The University of
Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

a) yoshino@sys.wakayama-u.ac.jp

b) mai.miyabe@gmail.com

2. 関連研究

関連研究として、まず、異文化間コミュニケーションにおける、文化差に関する研究を示す。Cho らは、異文化話者らがコンピュータとネットワークを介してコミュニケーションを行う際に用いる絵文字に着目した。絵文字は、異文化間で普遍的に解釈されないという問題がある。そこで、その問題を解決するために、解釈に文化差のある絵文字の検出における工学的な手法の適用可能性について検討した [5]。検討の結果、従来の工学的な手法では、人の文化差判定を近似することは困難であることを示した。Koda らは、アバターを介したコミュニケーションにおける、異文化間での表情の解釈に着目した。アバターの表情に関するユーザの解釈について実験を行った結果、表情の解釈が文化によって大きく異なることを示した [6]。文化差に関しては、これまでにいくつかの検討が行われているが、文化差判定は容易ではない。

Wikipedia の複数の言語版を利用した研究を示す。藤原らは、自国の言語版の Wikipedia だけでは情報量が不足する場合の補完のために、多言語版の Wikipedia に対して、リンク構造解析を用いることで、差異情報を抽出する方法を提案している [7]。松浦らは、日本語と外国語での同一ニュースに関する変遷を分析するために、Wikipedia を用いている [8]。吉岡は、機械翻訳システムの精度向上のために、Wikipedia の言語間リンクを用いた中日の翻訳辞書の作成方法を提案している [9]。

このように、Wikipedia は知識抽出分野で資源として注目を集めており、様々な利用が検討されている。しかし、これまでに、Wikipedia の多言語データを利用した文化差検出に関する試みは行われていない。

画像検索に関するシステム開発や研究は、数多く行われている。画像検索の検索結果として表示する画像は、その画像が見つかったページのタイトルや本文、画像のファイル名などの説明文の情報をもとに判断している。また、画像内容を表す特徴量を考慮して、画像検索を行うシステムも提案されている。例えば、獅々堀らは、検索語句に適した画像を Web 上から検索するために、画像知識 DB を用いたフィルタリングおよびユーザの画像選択のフィードバックを用いることで、既存の画像検索システム (Google 画像検索) による、画像の検索精度の改善を試みている [10]。出原らは、ユーザ自身がよく知らない物事についての情報を得るために画像検索を行う場合、検索結果として出力された画像が、正解かどうか分からない場合があることに着目した。検索結果の各画像の内容把握を助けるため、文章を元の Web ページから抽出し、補うことを試みている [11]。また、Fujita らは辞書の定義に適した画像の選択方法として、同義語・上位語・関連語を用いて検索語を拡張する手法を用い、辞書の見出しにあった高精度な画像選択方法を提案して

いる [12]。

本稿では、特に、文化差理解支援という観点から、可視化の効果の検証を行う。

3. 文化差の定義

本章では、本研究で検出・可視化の対象とする「文化差」について定義する。

文化差を定義するためには、まず、「文化」の定義が必要である。「文化」(Culture) の定義は、日本と欧米では異なり、一概に定義することは困難である [13]。例えば、今日欧米で用いられる「文化」は、「知識、信仰、芸術、道徳、慣習、その他社会の一員としての人間によって獲得される能力や習慣を包含する複合体である」と定義づけられている [13]。このような「文化」を単純に「測る」ことは困難であるが、コミュニケーション支援に文化差検出手法を適用するためには、何らかの尺度を考える必要がある。そこで本稿では、特に「知識」の面から「文化」を捉えることとし、まず、文化差検出手法の第一歩として、形式知化された知識の違いで文化差を測ることとした。

次に、「第 1 種の文化差」と「第 2 種の文化差」を定義する。「第 1 種の文化差」のある内容は、一方の文化圏で発生したり、存在したりしている「もの」や「こと」で、別の文化圏にも伝わっている場合も、基本的には、もとの文化圏の内容を指しているものである。例えば、日本の地名のいくつかは、海外にも伝わっているが、その地名の持つ背景 (歴史的あるいは文化的) などは正確には理解されない。「第 2 種の文化差」のある内容は、どちらの文化圏にも存在するが、それぞれの文化圏で意味の異なるものである。例えば、「醤油」は日本と中国のどちらにも存在するが、日本の「醤油」と中国の「醤油」は異なる^{*1}。

4. 可視化検証用のデータセット

可視化検証用のデータセットとして、語句および語句を説明する説明文、語句を表す画像が必要となる。本章では、可視化検証用のデータセットの作成手順について述べる。

4.1 検証用の語句の収集

本評価で用いる検証用の語句は、著者らが独自に作成したものを用いた。可視化検証用の語句は、次の手順で収集した。

Step 1 本学のシステム工学部および大学院システム工学研究科の学生に依頼し、下記に該当する日本語の

^{*1} 中国の醤油は、味やにおいが日本の醤油よりも強い。中華料理での使われ方は、香りや味よりも、色づけに重点を置いている。日本の醤油は、中国から日本に伝わり、日本で独自に改良されたと言われており、製造方法も異なる。また、世界各国に輸出あるいは現地生産されており、広く受け入れられている。

語句を収集した．

- 日本独特のもの（と思っている語句）
- 日本にも海外にもあるが、日本と海外とは違うもの（と思っている語句）

ただし、地名、人名、固有名詞は、ほとんど第1種の文化差のある語句に該当するため、入力しないように依頼した．

Step 2 収集した語句が Wikipedia の日本語版と中国語版の両方に、記事として存在するか調べた．

Step 3 日本に3年以上住んでいる本学の中国人留学生5名に依頼し、Step 1 で収集した語句のうち Wikipedia に記事（日本語版と中国語版の両方）として存在するもの（Step 2 で調べた結果によって抽出したもの）を、次の3種類に分類してもらった．

- 中国にはない（第1種の文化差）
- 中国にもあり、日本と同じ（文化差なし）
- 中国と日本のものは違う（第1種の文化差または第2種の文化差）

「中国と日本のものは違う」を選択した場合には、どのように違うかを簡単に記述してもらった．

Step 1 で、日本語の語句が、2,200 語句以上集まった．Step 2 で、入力語句の5%程度について、入力された順に Wikipedia の記事の有無を調べ、114 語句をデータセットとした．

表1に「文化差あり」に分類された語句（83語）を、表2に「文化差なし」に分類された語句（31語）を示す．本稿では、Step 3 において、「第1種の文化差」と「第2種の文化差」と分けられた語句を、「文化差あり」として扱う．また、表3に「文化差あり」に分類された語句の理由を示す．なお、語句の分類は、分類作業を行った各中国人留学生の経験に基づいた意見をもとにしているため、事実と異なる可能性がある．表3に示した分類結果は、5名の留学生による分類結果を集計し、最も多い分類結果（2名以上が分類）を代表値とした場合の結果をもとに分類した．「日本と中国との違いについて」の内容は、留学生らの記述をもとにしている．

4.2 検証用の説明文の作成

本評価で用いる検証用の説明文は、Wikipedia の説明文を用いている．可視化検証用の語句の説明文は、4.1 節で収集した、文化差判定済みの114語を対象として、次の手順で作成した．

Step 1 該当語句の Wikipedia 記事から、導入部^{*2}を抽出する．今回の検証実験の評価者は日本人であるため、中国語版は機械翻訳を用いて日本語へ翻訳する^{*3}．

^{*2} [http://ja.wikipedia.org/wiki/Wikipedia:スタイルマニュアル_\(レイアウト\)](http://ja.wikipedia.org/wiki/Wikipedia:スタイルマニュアル_(レイアウト))

^{*3} 言語グリッド [14] を介して J-Server を利用した．

表1 「文化差あり」に分類された語句（83語）

Table 1 A part of experimental words set classified to have a cultural difference. The number of words is 83.

バレンタインデー	羽根突き	アニメ	醤油	地震	忍者
インターネット	公衆便所	おたく	宗教	豆腐	教員
ファッション	コンビニ	マナー	赤飯	剣道	俳句
おせち料理	中華まん	ビール	貨幣	元日	舞妓
カレーライス	カラオケ	七五三	煎餅	柏餅	漫画
自動販売機	タクシー	わさび	神社	納豆	漫才
クリスマス	ニート	親子丼	神主	祝日	味噌
おにぎり	ちくわ	軍隊	神道	政府	萌え
パチンコ	こたつ	昆布	牛丼	老人	禅
すき焼き	ちまき	空気	相撲	饅頭	茶
就職活動	天ぷら	団子	大人	妖怪	能
ひな祭り	天津飯	干支	浴衣	節分	祭
ラーメン	年賀状	漢字	寿司	首相	侍
アイドル	和菓子	七夕	刺身	落語	

語句の分類は、日本に3年以上滞在している中国人留学生5名が行った．各自の経験に基づいた意見をもとにしているため、事実と異なる可能性がある．

表2 「文化差なし」に分類された語句（31語）

Table 2 A part of experimental words set classified not to have a cultural difference. The number of words is 31.

ようかん	化粧品	ケーキ	映画	仕事	母親	鬼
ステーキ	外国人	ゲーム	観光	子供	虹	
固定電話	公務員	花火	教授	戦争	銃	
新婚旅行	時刻表	学歴	発酵	新聞	神	
携帯電話	障害者	友情	父親	大豆	夫	

語句の分類は、日本に3年以上滞在している中国人留学生5名が行った．各自の経験に基づいた意見をもとにしているため、事実と異なる可能性がある．

Step 2 導入部がない記事の場合には、Wikipedia の解説本文の最初の部分を中心に抜き出す．最初の部分が、記事の内容を示すものとして不適切な場合には、他の部分を抜き出す．

Step 3 日本語の説明文と中国語の説明（機械翻訳語）を比較して、文章量に大きく差がある場合には、同程度（差が2~3倍程度以内）になるように、文章量を調整する^{*4}．

表4に、各言語の説明文の分量を示す．平均文字数は中国語の方が約20%多い．また、最長文字数および最短文字数も中国語の長い．今回、収集した語句は、もともと「日本独特のもの（と思っている語句）」を想定して集めた．日本に関する記事の場合、日本語の Wikipedia の説明の導入部は、比較的簡潔に記載されており、解説本文が充実している記事が多い．逆に、中国語の Wikipedia において、日本に関する記事の場合、解説本文そのもののしかない場合が多いため、分量が多くなる傾向があった．

^{*4} 多くの場合は、日本語の説明が中国語の説明に比べて少ないため、追加した．

表 3 「文化差あり」に分類された語句とその理由の一部

Table 3 A part of experimental words set and the reason classified to have a cultural difference.

ID	語句	分類結果	日本と中国との違いについて
6	干支	日本とは違う	日本では、亥(いのしし)だが、中国では豚。
28	刺身	中国にはない	中国では、あまり生の食べ物は食べない。
45	煎餅	日本とは違う	中国では煎餅は主食。日本ではお菓子。
53	天津飯	中国にはない	中華料理に、ご飯の上に卵を載せる料理はない。
57	ラーメン	日本とは違う	日本のラーメンとは味が違う。
100	パチンコ	中国にはない	中国ではギャンブルは禁止されている。
108	納豆	中国にはない	中国に納豆はない。
112	醤油	日本とは違う	中国の醤油は塩辛い。
113	おにぎり	中国にはない	中国では、冷めたご飯は食べない。
114	饅頭	日本とは違う	日本の饅頭の中には、餡がある。

「中国にもあり、日本と同じ」と分類されたものは省いている。
分類は、日本に3年以上滞在している中国人留学生5名が行った。
判定結果は、留学生各自の経験に基づいた意見をもとにしているため、事実と異なる可能性がある。2名以上が分類したものを結果とした。
「日本と中国との違いについて」の内容は、留学生らの記述およびWikipediaの内容をもとにしている。

表 4 各言語の説明文の分量

Table 4 Number of explanation characters in each language.

	日本語の説明文	中国語の説明文 (機械翻訳)
文字数の平均(文字)	276	335
文字数の標準偏差(文字)	174	228
最長文字数(文字)	898	1388
最短文字数(文字)	28	52

4.3 検証用の画像の収集

本評価で用いる検証用の画像は、Google 画像検索の結果を用いている。可視化検証用の語句を表す画像は、4.1 節で収集した、文化差判定済みの114語を対象として、次の手順で収集した。

Step 1 該当語句を Google 画像検索^{*5}を用いて検索し、画像を10件取得する^{*6}。中国語の検索は、検索言語を中国語(簡体字)にして、同様に Google 画像検索^{*7}を用いて検索し、画像を10件取得する。日本語の語句に対応する中国語は、予め Wikipedia の多言語リンクを利用して抽出する。

Step 2 すべての画像を確認し、不適切な画像^{*8}を削除する。また、検証用語句として、意図する内容以外の意味を表す画像については除外した。

Step 3 該当する語句を表す画像を3件選択する。選択の基準は、類似の画像を選択するのではなく、可能

な限り多様性のある画像を選択する^{*9}。該当する語句を表す画像が、画像検索結果に含まれていない場合には、検索範囲を広げて該当する語句を表す画像を選択する。日本語の語句を表す画像と中国語の語句を表す画像が同じにならないように、異なる画像を選択する。

5. 文化差の判定実験

5.1 仮説

文化差の可視化は、文化差理解の支援を目的としている。つまり、文化差理解支援により得られるべき効果としては、まず理解度の向上がある。また、文化差判定が必要となる場面には、対話などの即時性の求められる場面も含まれる。そのため、文化差理解の支援では、理解に要する負荷の軽減効果も求められる。

そこで、文化差理解支援による効果の検証項目として、「理解度」と「理解負荷」を取り上げる。本実験では、各項目の評価指標として、以下のものを用いる。

- 理解度：記述内容の正確性の評価
- 理解負荷：文化差理解にかかる時間の評価および主観による評価

本実験では、次の3つの仮説を立てた。

仮説1：画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定を短時間に行える。

仮説2：画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定の主観的な負荷は低い。

仮説3：文化差の理解は、説明文の提示の方が画像の提示よりも、正確に理解できる。

5.2 実験の詳細

仮説を検証するために、次の3種類の評価を行った。

- (1) 語句のみによる文化差の判定
- (2) 説明文による文化差の判定
- (3) 画像による文化差の判定

被験者グループとして、「説明文」グループおよび「画像」グループの2つを作成した。各グループは、それぞれ10名で、本学のシステム工学部および大学院システム工学研究科の学生である。

「説明文」グループ：(1) 語句のみによる文化差の判定を行ったあと、(2) 説明文による文化差の判定を行う。

「画像」グループ：(1) 語句のみによる文化差の判定を行ったあと、(3) 画像による文化差の判定を行う。

図1、図2、図3に、それぞれ「語句」、「説明文」、「画像」による文化差の判定の実験システムの画面を示す。図1では、各語句をみて、文化差があるかどうかを判定してもらう。

^{*9} 多様性がない場合には類似の画像とする。

^{*5} <http://www.google.co.jp/imghp?hl=ja>

^{*6} リンク先に画像がない場合に取得できない場合もあった。

^{*7} <http://www.google.co.jp/imghp?hl=zh-cn>

^{*8} 公序良俗に反する画像や刺激の強い画像。

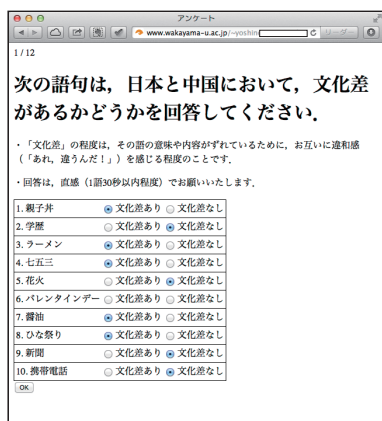


図 1 語句のみによる文化差判定用の画面例

Fig. 1 Screenshot of cultural differences evaluation using 'a word.'

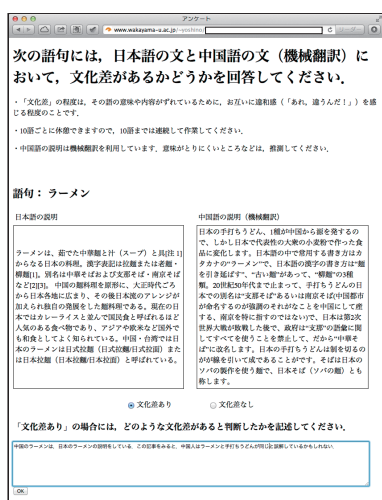


図 2 説明文による文化差判定用の画面例

Fig. 2 Screenshot of cultural differences evaluation using 'explanatory notes.'

図 2 では、日本語の説明文と中国語の説明文を読んで、文化差があるかどうかを判定してもらう。中国語の説明は、機械翻訳を用いているため、文字化けの発生や内容を十分に読み取れない場合がある。その際には、読み取れる範囲と自分の知識で判定するように伝えている。判定は、数分程度で行うよう伝えている。また、各語の判定にかかる時間を測定している。文化差があると判定した場合には、どのような文化差があると判定したかを記述してもらう。計測する時間は、記述の時間も含まれている。また、休憩が適宜とれるように、10 語毎に休憩画面を表示している。

図 3 では、日本の画像と中国の画像を閲覧することにより、文化差があるかどうかを判定してもらう。「説明文」と同様に、文化差があると判定した場合には、どのような文化差があると判定したかを記述してもらう。計測する時間は、記述の時間も含まれている。また、休憩が適宜とれるように、10 語毎に休憩画面を表示している。



図 3 画像による文化差判定用の画面例

Fig. 3 Screenshot of cultural differences evaluation using 'images.'

表 5 各グループの判定時間および正解率

Table 5 Result of the judgment time and the accuracy rate in each group.

	「説明文」 グループ	「画像」 グループ
判定時間の平均 (秒)	112	29
判定時間の標準偏差の平均 (秒)	162	49
語句の正解率	71.9%	64.9%
正解率	34.2%	36.0%
「文化差あり」語句の正解率	18.1%	16.9%
「文化差あり」判定数の平均 (個)	38.7	22.4

6. 実験結果と考察

6.1 文化差の判定時間

表 5 に、各グループの判定時間および正解率を示す。判定時間の平均は、「説明文」および「画像」による文化差評価を行った 114 語句の平均である。

「画像」の文化差評価にかかる時間は、「説明文」の文化差評価にかかる時間の約 4 分の 1 であった。判定時間の標準偏差の平均は、「画像」に比べ、「説明文」の方が大きい。説明文を用いた文化差評価は、大きなばらつきがあったことがわかる。

判定時間の観点から、文化差の理解にかかる負荷は画像の負荷が少ないことがわかった。

「仮説 1: 画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定を短時間に行える」は、成立した。

6.2 文化差の正解率の分析

表 5 の「語句の正解率」は、各グループが「説明文」あるいは「画像」を用いた文化差判定の前に行った、語句の文化差判定（図 1）の正解率である。「正解率」は、各グループの「説明文」あるいは「画像」を用いた文化差判定の正解率である。

「説明文」あるいは「画像」を用いた文化差判定の正解率は、それぞれ、34.2%、36.0%であり、語句の正解率は 71.9%、64.9%と比べて低かった。

「文化差あり」語句の正解率（表 3）はさらに低く、「説明文」あるいは「画像」の正解率は、それぞれ 18.1%、16.9%であった。

文化差理解度の観点から、「語句の正解率」と比べて、「説明文」あるいは「画像」の正解率は低かった。「語句」は、被験者のこれまでの経験に基づいた評価である。「説明文」あるいは「画像」は、提示された説明文あるいは画像をもとに、判定した結果である。説明文あるいは画像の提示の結果、正解率が低下することがわかった。

「仮説 3: 文化差の理解は、説明文の提示の方が画像の提示よりも、正確に理解できる」は、成立しなかった。

6.3 文化差の正解率の分析（「第 1 種の文化差」を「文化差なし」と見なした場合）

今回、3 章で述べた、第 1 種の文化差（もとの文化圏の内容を指しているもの）に関して、適切に評価されなかった可能性がある。「第 1 種の文化差」はもとの文化圏の内容における「歴史的あるいは文化的」など、正確には理解しにくい文化差である。今回の評価で提示した、「説明文」および「画像」だけでは、「歴史的あるいは文化的」な文化差の情報を提示するためには、不十分だった可能性がある。

今回の「説明文」の判定実験においては、該当語句の概要を提示していた。また、「画像」の判定実験において、第 1 種の文化差を持つ画像は、両方とも、ほとんどの場合には、適切な画像を示していた。つまり、「歴史的あるいは文化的」な差を、容易に判別することは困難である可能性がある。

そこで、第 1 種の文化差の取り扱いについて、再検討を行った。表 6 に、第 1 種の文化差を「文化差なし」とみなした場合の各グループの正解率を示す。「説明文」グループ（語句）および「画像」グループ（語句）は、それぞれ「説明文」グループおよび「画像」グループの、「語句」のみによる文化差の判定結果である。表 6 の「文化差あり」は、第 2 種の文化差の判定をしていることになる。

正解率をみると、「語句」「説明文」「画像」に、大きな違いは見られないが、第 1 種の文化差を「文化差あり」と判定した表 5 の正解率と比べ、大きく改善していることがわかる。多くの被験者は、提示された「第 1 種の文

表 6 第 1 種の文化差を「文化差なし」とみなした場合の各グループ正解率

Table 6 Result of the accuracy rate in each group when first cultural difference is considered as no cultural difference.

	「説明文」 グループ （「語句」）	「画像」 グループ （「語句」）	「説明文」 グループ	「画像」 グループ
正解率（全体）	64.9%	54.4%	67.5%	62.3%
誤判定（文化差あり）	9.8%	48.8%	68.3%	80.5%
誤判定（文化差なし）	49.3%	43.8%	12.3%	13.7%

化差」のある内容を見て、「文化差なし」と判定していることがわかった。

「文化差あり」の誤判定をみると、「説明文」「画像」グループの誤判定の割合が、「語句」のみの判定に比べて、大きいことがわかる。「文化差なし」の誤判定をみると、「説明文」「画像」グループの誤判定の割合は、「語句」のみの判定に比べて、大きく減少していることがわかる。これは、第 2 種の文化差の判定も、十分な精度で判定されていないことがわかる。

「説明文」の提示および「画像」の提示による文化差の可視化は、今回の実験では文化差理解支援への十分な貢献は見られなかった。

ただし、今回の実験における文化差判定の基準は、検証用の語句収集時の文化差判定における文化差の基準と同一ではなかった。具体的には、4.1 節で述べたように、検証用の語句収集時には、3 種類への分類と違いの記述が、中国人留学生により行われた。つまり、中国人留学生の目線から判断されたものであることから、ここで文化差があると判断された語句は、細かい違いにも着目された可能性がある。一方、実験においては、文化差の判定基準として、「日本人と中国人がお互いに違和感を感じるかどうか」を目安にするよう指示を与えた。今回抽出した語句は、もともと「日本独特」であるという点を起点として収集しており、日本と中国の類似部分が大きければ、多少の違いが存在した場合でも、「文化差なし」と判定された可能性がある。すなわち、検証用語句の収集における文化差の判定基準と、実験時における判定基準が異なった可能性がある。今後、それぞれの判断基準を合わせた上で、可視化の効果を再検証する必要があると考えられる。

6.4 文化差の判定のしやすさ

表 7 に、被験者の主観による評価として、文化差の判定のしやすさに関するアンケート結果を示す。アンケートは、5 段階リッカートスケール（1: 強く同意しない、2: 同意しない、3: どちらともいえない、4: 同意する、5: 強く同意する）を用いた。検定の結果、「説明文」グ

表 8 文化差の評価実験における被験者の文化差判定の理由

Table 8 Subjects' judgment reason in the experiment of judgment cultural difference.

・「説明文」の提示実験参加者の判定理由 ・中国語の機械語翻訳の文法や意味がとれないところがあった（8名が同内容の回答）。 ・文章だけでは比較するものが想像できず、違いがあるかどうか、判断できませんでした（評価 1）。 ・理由を記述するときに日本語と中国語の文章の粗探しのような説明しかできなかった（評価 2）。 ・日本独自のもののなか、それとも世界中で認知されているもののなかで判断しづらい項目が多数あった（評価 2）。 ・誤訳なのか、もしくは訳は正しくて文化に差があると考えられるのか判断が難しい（評価 2）。
・「画像」の提示実験参加者の判定理由 ・写真の場合は的確な写真の表示がされていないとどう判断しているのかわからなかった（評価 3）。 ・明確に違うものは簡単でしたが、言葉には言い表せない、なんか違うように感じるというものもあり、その判断が難しかった（評価 3）。 ・はっきり出ていたものは文化差が出ていた（評価 4）。 ・どの程度違えば文化差があると言ってよいのかあいまいになってしまったから（評価 2）。 ・その単語のイメージが自分の中にある場合、少し判断に困りました（評価 4）。 ・そもそも日本のもので想像しづらい（概念的に広い？パチンコ、団子など）ものがあり、判断しづらいと感じた（評価 2）。 ・判断のしづらいものもあったが、大多数のものはあまり迷うことなく文化差の判断ができたため（評価 4）。

ループと「画像」グループにおいて、有意水準 5%で有意差が見られた。文化差の判定は、画像の方が説明文に比べてしやすいと被験者は感じていることがわかった。

「仮説 2: 画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定の主観的な負荷は低い」は、成立した。

6.5 文化差の評価実験における被験者のコメント

表 8 に、文化差の評価実験における被験者の文化差判定の理由を示す。「説明文」「画像」の両方とも、被験者は文化差の判定に迷う点があり、特に、「説明文」の場合には、機械翻訳の翻訳精度の問題で、意味がとれないところがあり、判定できない場合もあることがわかった。被験者の理由を見ると、「説明文」の文化差の判定が難しいことがわかった。

表 5 の「文化差あり」判定数は、「説明文」の方が「画像」より多いが、被験者の文化差判定の理由から、日本語と中国語の文章の違いを、文化差として判定した場合があったと考えられる。

表 9 に、文化差の評価実験における被験者の感想を示す。「説明文」の提示実験参加者のコメントとして、「言葉では説明できない文化差があるのではないかなと思った」があった。また、日本語の説明では、暗黙的に説明が省略されている可能性も指摘された。今回、日本語の説明は Wikipedia の導入部のみであり簡潔に記載されて

表 9 文化差の評価実験に関する被験者の感想

Table 9 Subjects' impression in the experiment of judgment cultural difference.

・「説明文」の提示実験参加者の感想 ・文章だけでなく、具体的な物は写真もあればより判断がつけやすいと感じた。 ・文章同士を比較してしまっていると感じた。言葉では説明できない文化差があるのではないかなと思った（文章では同じ意味になっていても、日本のお米と海外のお米は違うのではないかな）。食べ物系の単語は「日本独特」「○○から伝わった」「○○が変化した」のような言葉があるので、判断しやすかった。 ・文化差について新しい発見があった（知らない知識が増えた）。 ・中国語機械翻訳文の方が、日本文化の起源や背景について詳しく書かれている場合が多く、日本語説明文より自国の文化に対して気付くことが多かったような気がします。日本語説明文は、自国のことなので割と省略がかかっていて（暗黙の了解？）、情報量が少し足りないように感じました。 ・食材や料理、または日本で発祥したものについては、あまり文化差がないように感じた。 ・日本固有の文化などは、文化差が生まれにくいと感じました。それに対し、神や宗教など、その国に古くから根付いているものに関しては、文化差が生まれやすいと感じました。 ・文化が国に伝来する際に、その国の特色に合うようアレンジを加えることがあるため、文化差が生じると考える。携帯電話、萌えなど比較的最近できた語句の文化差はあまりなさそうだと感じた。
・「画像」の提示実験参加者の感想 ・画像を見る前では、文化差があるだろうと思っていたものが、実際の画像を見ると文化差がないパターンが多かったように感じました。 ・意外と文化差が少ないと思いました。 ・たとえば検索画像が同じようなものを見せていても、その国で発祥したものよりも他国のものの方が、より一般イメージ化されたものがでてくるのではないかなと思った。例えば日本で生まれたものは、日本語で検索するとより様々な種類のものでくるのではないかな。 ・中国と日本は文化差がもっとあると思っていたが、想像していたよりは少なかった。 ・日本独特の文化のものは、ほとんどそのままの形で中国に伝わっているように感じました。 ・文化差というのは、その言葉の見た目や写真だけで判断できるものではないかなと思った。事象を指す言葉であれば「その事象がどのような扱いを受けているか」「社会的にどのような位置づけにあるものなのか」まで含めて文化差でないかなと思った。

いる場合が多い。実験参加者が省略されていると感じた部分に関しては、解説本文に記載されていると可能性が高いと考えられる。逆に中国語は、解説本文そのものしかない場合が多く、今回、その部分を説明文として用いたため、日本語に比べて、説明が多くなる傾向があった。

「画像」に関しては、「実際の画像を見ると文化差がないパターンが多かった」「意外と文化差が少ないと思いました」「想像していたよりは少なかった」という意見があった。今回は、Google 画像検索で上位 10 件から 3 件を選択したが、3 件では文化差を表現するために不十分であった可能性がある。

6.6 新しい文化差理解支援の可視化手法の提案

説明文および画像による文化差の可視化の効果の検証を行った結果、画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定を短時間に行え、

表 7 文化差の判定のしやすさに関するアンケート結果 (5 段階評価)
Table 7 Result of questionnaire survey on judgment of a cultural difference (5-point Likert scale).

質問項目	回答グループ	評価値 (人)					中央値	最頻値	有意確率
		1	2	3	4	5			
文化差の判定はしやすかった。	「説明文」グループ	2	8	0	0	0	2	2	0.008
	「画像」グループ	0	5	2	3	0	2.5	2	

・ 5 段階評価の評価値：1: 強く同意しない, 2: 同意しない, 3: どちらともいえない, 4: 同意する, 5: 強く同意する
・ 有意確率は, Mann-Whitney U-test を用いた。

文化差の判定の主観的な負荷は低いことを示した。しかし、現時点の「説明文」の提示および「画像」の提示は両方とも、文化差の理解に関して、十分に貢献しない可能性があることがわかった。

そこで、新しい文化差理解支援の可視化手法の提案する。これまでの検証実験をもとに、次の設計方針を考えた。

- (1) 機械翻訳を利用せず、母語の情報を提示する。
- (2) できるだけ多くの画像を提示する。
- (3) 具体的な違いの事例を示す。

7. おわりに

本稿では、文化差理解支援における、説明文および画像による文化差の可視化の効果の検証を行った。検証方法として、日本語と中国語の Wikipedia の説明文と Google 画像検索を用いた画像とを比較し、記述内容の正確性の評価、文化差理解にかかる時間の評価および理解負荷に関する主観評価を行った。

検証の結果、下記を明らかにした。

- (1) 画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定を短時間に行える。
- (2) 画像の提示による文化差の理解は、説明文の提示による理解より、文化差の判定の主観的な負荷は低い。

文化差の理解に関して、現時点の「説明文」の提示および「画像」の提示は両方とも、「歴史的あるいは文化的」な文化差の情報を提示するためには、不十分だった可能性がある。

今回抽出した語句は、もともと「日本独特」というところを起点にして集めた語句を利用しており、結果にその影響が含まれる可能性がある。例えば、中国人が判定すると、理解度や正解率などが変わる可能性もあるため、文化差の判断基準を十分に検討する必要がある。

今後、文化差の可視化の方法として、「説明文」および「画像」を組み合わせた、新しい文化差理解支援の可視化手法を実現し、その評価を行う。

謝辞 本研究の一部は、独立行政法人科学技術振興機構研究成果最適展開支援事業 (A-STEP) 探索タイプ「検索エンジンと機械翻訳を用いた多言語用語間における文化差検出サービス」および日本学術振興会科学研究費基盤研究 (B)(22300044) の補助を受けた。

参考文献

- [1] 藤井薫和, 重信智宏, 吉野 孝: 機械翻訳を用いた異文化間チャットコミュニケーションにおけるアノテーションの評価, 情報処理学会論文誌, Vol.48, No.1, pp.63–71 (2007).
- [2] 吉野 孝, 宮部真衣, 諏訪智大: Wikipedia を用いた文化差可視化サービスの開発, 情報処理学会, マルチメディア, 分散, 協調とモバイル (DICOMO2011) シンポジウム (2012)。
- [3] 吉野 孝, 宮部 真衣: 日本語版・中国語版 Wikipedia における相互言及の有無を用いた文化差検出方法の提案, 第 5 回 Web とデータベースに関するフォーラム (WebDB フォーラム 2012), B2-2, pp.1-8 (2012)。
- [4] 諏訪 智大, 宮部 真衣, 吉野 孝: Web ページにおける文化差可視化システムの開発, 平成 24 年度情報処理学会関西支部大会, F-01, pp.1-4 (2012)。
- [5] Cho Heeryon, 石田 亨, 山下直美, 稲葉利江子, 高崎俊之, 神田智子: 絵文字解釈における人間の文化差判定, ヒューマンインタフェース学会論文誌, Vol. 10, No. 4, pp. 427–434 (2008)。
- [6] Tomoko Koda and Toru Ishida: Cross-cultural study of avatar expression interpretations, SAINT 2006, pp. 130–136 (2006).
- [7] 藤原裕也, 灘本明代: Wikipedia の言語間比較による差異情報抽出手法の提案, 情報処理学会研究報告, Vol. 2011-DBS-152, No. 3, pp. 1–8 (2011)。
- [8] 松浦愛美, 江口浩二: 時系列対訳トピックモデルを用いた言語横断トレンド分析, 情報処理学会研究報告, Vol. 2010-DD-75, No. 11, pp. 1–5 (2010)。
- [9] 吉岡真治: Wikipedia を用いた中日カタカナ翻訳辞書の作成と言語グリッドへの応用, 電子情報通信学会技術報告, 人工知能と知識処理, Vol. 109, No. 424, pp. 43–46 (2010)。
- [10] 獅々堀正幹, 小泉大地, 柘植 寛, 北研二: 画像知識データベースを用いた WWW 画像検索システムの開発, 電子情報通信学会論文誌 D-I, J87-D-I(2), pp.154-163 (2004)。
- [11] 出原 博, 藤本 典幸, 竹野 浩, 萩原 兼一: WWW 画像検索における画像周辺の HTML 構文構造を考慮した画像説明文の抽出手法, 電子情報通信学会技術研究報告. データ工学 105(338), pp.19-24, (2005)。
- [12] Sanae Fujita and Masaaki Nagata. Enriching Dictionaries with Images from the Internet - Targeting Wikipedia and a Japanese Semantic Lexicon: Lexeed - : Proceedings of the 23rd International Conference on Computational Linguistics (Coling-2010), pp. 331–339 (2010).
- [13] 西田ひろ子: 異文化間コミュニケーション, 創元社 (2000)。
- [14] Toru Ishida. Language Grid: An Infrastructure for Intercultural Collaboration. IEEE/IPSJ Symposium on Applications and the Internet (SAINT 2006), pp.96-100, keynote address (2006).