

# 食の描写表現の相違を用いた言語圏ごとの料理の感じ方の比較手法の研究

03-160988 中元雪絵  
指導教員 坂田一郎 教授

## 1. 研究背景

グローバル化に伴い、食産業でも海外進出が増加しているが、進出先の地域社会へのローカライゼーション戦略を打ち出し成功する事例が多く見られるようになった。ローカライゼーション戦略の肝となるのは、地域ごとの文化的差異を理解することである。これまで文化に関する研究は、アンケート調査がほとんどであったが、近年生活に密着した情報発信手段であるソーシャルメディアのデータに注目が集まっている。Sajadmanesh らは、レシピデータを用い、味や材料からみた各国食習慣の距離を算出した[1]。この研究では文化の距離を世界規模で議論することに成功している一方で、各文化の間に具体的にどのような差異があるのか検出することは難しい。

## 2. 目的

本研究では、ソーシャルメディアを用いて地域ごとの対象の捉え方の違いを明確化することを課題とし、言語的表現の違いを分析することでこれら差異を得ることが可能であるかを検証する。具体的には、描写表現に着目し、描写表現の共起頻度の違いを用いて文化間差異を自動的に比較する手法を提案する。提案手法を用いて Twitter の食に関する英日言語の感想データを比較することにより、文化間の感じ方の差異を検出し、提案手法の有効性を考察する。

## 3. 提案手法

Fig. 1 は Twitter から得られる麺についての英語と日本語のツイートの例である。英語では麺に対して「hot」や「spicy」が共起し、日本語は「真冬の夜中」という場面設定や「旨い」が共起する。このような描写表現の相違を、異なる言語圏に対して定量的に評価し、比較することで文化的な差異を検出する。

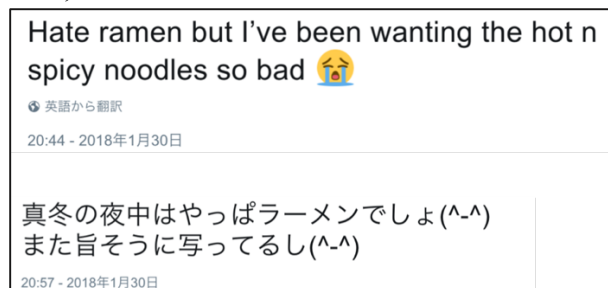


Fig. 1 英語と日本語のツイートの例

提案手法の概要は Fig. 2 のようになっている。

まず、感想データに出現する描写表現を抽出する。本研究では形容詞を描写表現とみなし、テキスト解析を行い Twitter の投稿データから形容詞のみを取得した。

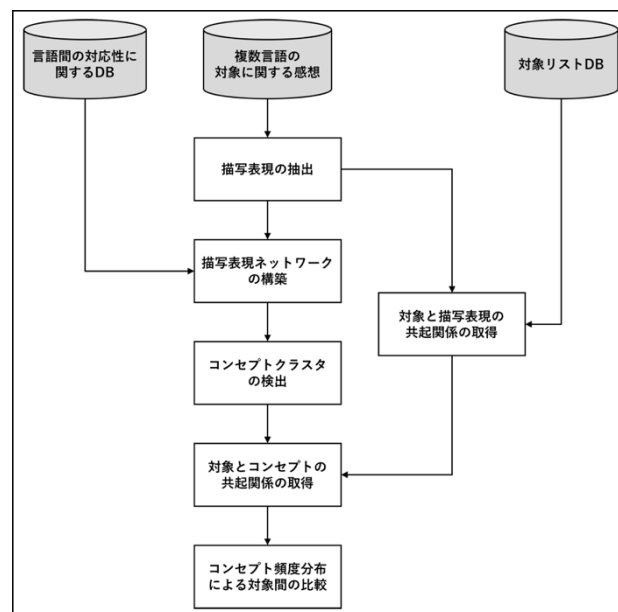


Fig. 2 提案手法の概要

続いて、各描写表現をノードとみなし、近い関係のある描写表現同士に重みを付与して繋ぐことで、描写表現ネットワークを構築する。本研究では、ConceptNet[2]を用いて関係のあるペアと重みを取得した。ConceptNet は異なる言語間を繋ぐリンクを持つため、このネットワークは全言語の描写表現に対し意味の近いものを結んだネットワークとなる。

描写表現ネットワークをクラスタリングし、得られたひとつひとつのクラスターを概念とみなす。各概念には、意味の近い表現が、言語間問わず集合する。

各分析対象（クエリ）と描写表現との共起関係を取得し、これを先ほど検出した概念クラスター単位で集計する。結果として、対象と各概念との共起確率が求まる。これを概念頻度分布と定義し、概念頻度分布の差を用いて、言語圏ごとの差を比較する。

## 4. 実験と結果

### 4.1. Twitter の食データに対する提案手法の適用

飲み物に関するクエリを英語・日本語に対して設定し、Twitter からそれぞれ感想データを取得した。各クエリに対して提案手法を用いて概念頻度分布を算出し、提案手法の妥当性の検証を行った。

### 4.2. 全体傾向の言語間比較

取得した全ての英語ツイートと日本語ツイートをを用い、描写表現を用いて英語圏と日本語圏の全体傾向の差異を正しく抽出できるか評価した。

ツイートの極性分析の結果、英語の極性平均は

0.266, 日本語は 0.0137 となった。また, 形容詞の種類による分析を行った結果, 英語では対象の属性や性質を描写する形容詞が多く, 日本語では自身の感情や評価に関する形容詞が多く用いられるという結果が得られ, 英語圏でポジティブな形容詞を用いて料理を描写する投稿が多い傾向に対し, 日本語圏でネガティブな評価表現で不満を言う投稿が多いという示唆を得た。これは, 日本語圏で自分に関する言及が多く, 英語圏で食事の状況に関する投稿が多いという先行研究の結果と一致する[3,4]。

さらに, 英語全体と日本語全体のコンセプト頻度分布を比較することで, 色への描写が英語圏にて多いことを検出した。これはアメリカでカラフルな食品が多い事実と合致する。

以上より, 描写表現を用いて食文化に対する知見を得ることが可能であり, 描写表現, 中でも形容詞から文化的差異を検出する有効性を示している。

#### 4.3. ネットワーク図による関係性の可視化

二つの対象について, コンセプト頻度分布の JS 距離を用いて距離の近さを定めた。この値を用いて英語のクエリ集合, 日本語のクエリ集合それぞれについてネットワーク図を描画し, 対象間の関係を正しく評価できていることを確認した。

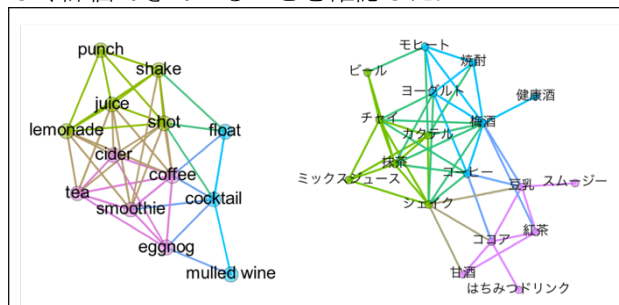


Fig. 3 ネットワーク図 (左: 英語, 右: 日本語)

英語で juice, lemonade, tea など果物が用いられる飲み物が近く, 日本語でココア, 甘酒, はちみつドリンク, 紅茶と温かい飲み物が近い位置にある。全ての飲み物の関係を正しく評価できている訳ではないが, 私たちの感性に合致する結果が見えていることから, 感性の評価にあたり, コンセプト頻度分布の JS 距離を用いることは妥当であると考えられる。

#### 4.4. 同一対象への感じ方の言語間比較

クエリのうち, 同一対象を扱っている 4 つのペアについて差の大きいコンセプト頻度分布を比較した。英語全体と日本語全体における差よりもさらに大きい差を持つクラスタの内容を取り上げ, 提案手法により文化的差異が得られることを検証した。

ここでは「coffee」と「コーヒー」の差に対する分析の結果を取り上げて説明する。Fig. 4 は coffee とコーヒーの頻度分布の差を表し, 値が正のクラスタは英語で共起が多く, 負のクラスタは日本語で共起が多いことを示す。

英語ではコンセプト《暖かい, warm》や, コンセプト《遅い, boring》が日本語に比べて多く共起した。

アイスコーヒーの起源は日本であり, 英語圏では冷たいコーヒーを飲む習慣があまりない背景から, コーヒーに対する「熱い」という印象がアメリカの方が強い可能性がある。

また, コンセプト《遅い, boring》の共起傾向から, アメリカ人にとってコーヒーが「boring」な文脈に存在することが多く, 「exciting」なコーヒーに対する欲求がある可能性が示唆される。実際に, 「boring」でないコーヒーを作る方法を紹介する記事は多く存在する。

日本語では, コンセプト《lovely, 可愛い》が多く共起した。これは日本でラテアートが独自の変化を遂げたためだと思われる。アメリカにもラテアートは存在するが, 日本のようにキャラクターを描く文化は一般的ではなく, コーヒーやラテアートに対して「可愛い」という一般認識はないと予測できる。

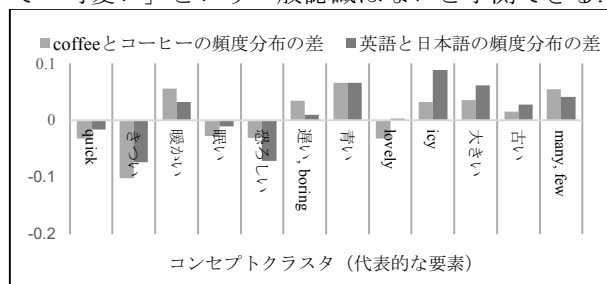


Fig. 4 coffee とコーヒーの頻度分布の差

他のペアに対しても, 同じように文化的な感じ方や捉え方の差異を検出することが可能であった。これにより, 提案手法の有効性を示した。

#### 5. 結論

本研究では, ソーシャルメディア上の描写表現を用いて人の感性に基づく文化間比較の行うことの有用性を示した。

また, 複数言語の描写表現をコンセプト単位で分類した上で, 共起頻度の分布を用いて比較を行うことで, 異なる言語圏に関する文化的差異を明確化する手法の開発に成功した。

#### 参考文献

- [1] SAJADMANESH, Sina, et al. Kissing Cuisines: Exploring Worldwide Culinary Habits on the Web. In: Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 2017. p. 1013-1021.
- [2] SPEER, Robert; CHIN, Joshua; HAVASI, Catherine. ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In: AAAI. 2017. p. 4444-4451.
- [3] VIDAL, Leticia, et al. Using Twitter data for food-related consumer research: A case study on “what people say when tweeting about different eating situations”. Food Quality and Preference, 2015, 45: 58-69.
- [4] ACAR, Adam; DEGUCHI, Ayaka. Culture and social media usage: Analysis of Japanese Twitter users. International Journal of Electronic Commerce Studies, 2013, 4.1: 21.