

中国語での飲食情報提供のためのシズルワードの日中対訳表現の分析

石橋 明大[†] 平林 (宮部) 真衣[‡] 吉野 孝[†]
[†]和歌山大学システム工学部 [‡]東京大学大学院医学系研究科

1 はじめに

日本政府観光局 (JNTO) が発表している国籍/目的別訪日外客数のデータ [1] によると, 2015 年~2021 年の期間では中国からの訪日人数が最も多く, その中でも 2015 年~2020 年の期間では, 観光客としての訪日人数が最も多かった. また観光庁の調査 [2] によると, 訪日外国人が期待していたことは「日本食を食べること」が最も多く, 観光客にとって飲食の重要性は高い. このように, 今後は特に訪日中国人観光客向けの, 飲食店の Web サイトの多言語化や店内への多言語メニューの設置などといったインバウンド対策が重要になる.

我々は「サクサク」や「かりかり」のような, 食品のおいしさを表す表現である「シズルワード」[3]に着目する. シズルワードには, 翻訳が省略されやすい擬態語やオノマトペ [4] が多く含まれるため, 正確な多言語化は容易ではない. これまでに進藤らはシズルワードの日英対訳表現について, 辞書から取得できない「隠れた対訳表現」の分析を行った [5]. この分析では, シズルワードの「多くが複数の対訳表現を持つ」「辞書に対訳として掲載されていない表現が存在する」といった特徴に着目している. 日中対訳表現は, 1 語に翻訳することが難しく, 必ずしも同義語が存在しない [6] ため, この特徴による問題が英語と同様に発生する可能性がある.

そこで本研究では, 類似語の対訳は飲食情報提供の際に利用できる可能性があると考え, これを「類似翻訳表現」と定義する. 本稿では, 中国語の飲食情報提供のため, シズルワードの日中対訳表現における, 既存対訳表現および類似翻訳表現について分析する.

2 関連研究

シズルワードの外国語翻訳に関する研究として, 日英対訳表現を分析した研究がある [5]. この研究では, 英語表記のレシピサイトや, 飲食店の Web サイトから取得した文章を用い, 辞書では取得できないシズルワードの対訳表現である「隠れた対訳表現」を取得している. 本稿では日中対訳を対象とし, シズルワードの類似翻訳表現の取得を試みる.

また, Kato らはメディアごとに登場するシズルワードの傾向を分析した [7]. この研究では, レシピサイトでは食感系や味覚系, 一般の Web サイトでは情報系のシズルワード, マイクロブログでは独特なシズルワードが多く用いられていることを明らかにした¹⁾. 本稿ではこの研究結果を参考にし, 特に擬態語やオノマトペを多く含む食感系や味覚系のシズルワードを取得しやすい, レシピサイトからテキストデータの収集を行う.

3 シズルワードの既存対訳表現の分析

3.1 分析対象とするシズルワード

本稿では「おいしさを感じる言葉」[8]に掲載されている系統別シズルワードランキングを参考に, 分析対象とするシズルワードを選定した. 系統別シズルワードランキングの上位 30 語²⁾から, Kato らの先行研究 [7]において, 分析にあたって有用ではないとされた「美味」を除いた, 計 90 語のシズルワードを用いる.

Analysis of Japanese-Chinese Translation Expression of Sizzle Words for Providing Food and Beverage Information in Chinese

Akihiro Ishibashi[†] Mai Miyabe Hirabayashi[‡] Takashi Yoshino[†]

[†]Faculty of Systems Engineering, Wakayama University

[‡]Graduate School of Medicine, The University of Tokyo

¹⁾シズルワードはその表現内容によって, 食感系, 味覚系, 情報系の 3 系統に分かれている.

²⁾同率を含むため, 情報系のシズルワードは 31 語である.

3.2 分析の手順

分析対象であるシズルワードの既存対訳表現を, 以下の日中辞書 Web サイト³⁾から取得し, 傾向を分析する.

- 中国語辞書 - Weblib 日中中日辞典⁴⁾
- 中日辞書 北辞郎⁵⁾
- BitEx 中国語⁶⁾

まず, 各 Web サイト内で, 分析対象の各シズルワードをクエリとして検索し, 提示された中国語の語彙の中から, 併記されている解説が食品とは関係ないと考えられるものを除外した. 次に, 中国語が母語である 1 名の大学生が各語彙を確認し, 食品のおいしさを表す語彙として不適切なものを除外した.

3.3 分析結果と考察

3.2 節の手順により, 今回は計 66 語 (食感系 23 語, 味覚系 19 語, 情報系 24 語) のシズルワードの既存対訳表現を取得できた. 表 1 に取得結果を示す.

取得結果から, 以下の傾向が確認できた.

- 食感系は 30 語中 23 語 (約 77%), 味覚系は 29 語中 19 語 (約 66%), 情報系は 31 語中 24 語 (約 77%) についての既存対訳表現が取得されており, 系統による日中辞書からの既存対訳表現の取得可否への影響は見られなかった.
- 「さくさく」は 18 種類の既存対訳表現が取得できたが, 似た表現である「サクッと」や「さっくり」はそれぞれ 2 種類のみであり, 意味の似たシズルワードであっても, 取得できる既存対訳表現の数には差がある.

4 シズルワードの類似翻訳表現の分析

4.1 分析対象とする中国語表記のレシピテキスト

シズルワードの日中対訳表現での類似翻訳表現を分析するために, 本稿では, 中国人向けのレシピサイトである「下厨房」⁷⁾からテキストを取得する. このテキストには, 各レシピに関するレシピ名, レシピの紹介文, 料理手順, 料理を行う上でのヒントが含まれている.

取得するテキストの内容が偏らないように, 各カテゴリ内のレシピ一覧から 2 ページ分ずつ, 重複がないように取得した⁸⁾. 結果として, 20,591 件のレシピテキストを取得した.

4.2 分析の手順

分析の手順を以下に示す.

- (1) 取得した 20,591 件のレシピテキストから, 1,000 件をランダムサンプリングする¹⁰⁾.
- (2) 抽出した各レシピテキストを, 1 文ごとに分割する. 分割結果は, 12,081 文である.
- (3) 各文に対して, 前処理¹¹⁾を行う.
- (4) シズルワードの既存対訳表現について, 以下の手順で類義語を取得する.

³⁾これらは, 検索エンジンである Google での「日中辞書 サイト」を検索クエリとした際, 上位 3 ページ分の検索結果に含まれる日中辞書 Web サイトである.

⁴⁾中国語辞書 - Weblib 日中中日辞典: <https://cjjc.weblib.jp/>

⁵⁾中日辞書 北辞郎: <http://www.ctrans.org/>

⁶⁾BitEx 中国語: <https://bitex-cn.com/>

⁷⁾下厨房: <https://www.xiachufang.com/>

⁸⁾「下厨房」は会員数が 2 億を超える中国のレシピサイトで, レシピ投稿数は 110 万を超える. そのため, 類似翻訳表現を取得するには十分な量のテキストデータを取得できると考えられる.

⁹⁾下厨房のカテゴリは, 味の特徴や使用されている材料などで分かれている. 各カテゴリからレシピを取得することにより, 語彙に偏りが発生することを防ぐ.

¹⁰⁾分析の際の膨大な計算時間を短縮するため.

¹¹⁾記号の削除や, 丸囲み文字, 半角文字の正規化を行った.

表 1: 各シズルワードにおける既存対訳表現の取得結果

既存対訳表現の取得の可否	系統	シズルワード
取得可能 (66 語)	食感系 (23 語)	さくさく (18), ふわふわ (12), なめらか (11), しっとり (11), ふっくら (10), もちもち (7), ふんわり (6), かりかり (6), ほかほか (6), とろける (5), しゅきしゅき (4), トロトロ (3), ふわっと (3), ぷりぷり (3), サクッと (2), ぱりっと (2), さっくり (2), しゅきしゅき (2), ジューシー (1), コシのある (1), あつあつ (1), とろっと (1), からっと (1)
	味覚系 (19 語)	甘い (15), 濃厚 (11), さわやか (11), さっぱり (10), まろやか (9), やみつき (8), 芳醇 (7), 香ばしい (4), あっさり (4), 味わい深い (3), クリーミー (3), リッチ (3), うまみ (2), コクがある (2), くせになる (2), こんがり (2), ビリ辛 (2), スパイシー (1), 後を引く (1)
	情報系 (24 語)	贅沢 (10), 新鮮 (8), 匂 (7), 本格的 (7), 絶品 (6), 完璧 (6), 手作り (6), 本場 (6), 自家製 (5), 自然 (5), 天然 (4), 産地地消 (3), 秘伝 (3), こだわり (3), 熟成 (3), ヘルシー (3), 焼きたて (2), 炊きたて (2), 出来たて (2), 無農薬 (2), 揚げたて (1), 無添加 (1), 食べごろ (1), フレッシュ (1)
取得不可能 (24 語)	食感系 (7 語)	もっちり, ホクホク, とろーり, 口溶け, カリッと, 舌ざわりのよい, じゅわー
	味覚系 (10 語)	風味豊か, 深みのある, 飽きのこない, 後味がよい, 後味すっきり, スイート, マイルド, 脂の乗った, デリシヤス, すっきり
	情報系 (7 語)	季節限定, 採れたて, 厳選素材, 産地限定, 体にやさしい, プレミアム, 昔ながらの味

※シズルワードの後ろの丸括弧内の数字は、各シズルワードについて取得した既存対訳表現の数を表す。

表 2: 取得できた類似翻訳表現

シズルワード	既存対訳表現	類似翻訳表現
くせになる	上癮:「癖になる」,「やみつき」	可口:「口に合う」,「おいしい」
やみつき		
新鮮	鮮:「新鮮である」	嫩:「若い」,「柔らかい」
ふわふわ	松:「ふわふわと柔らかい」,「緩い」	松软:「柔らかい」
濃厚	濃:「濃い」,「濃厚である」	稠:「濃い」,「どろどろ」
		厚:「味が濃い」,「濃厚である」

※鍵括弧内に各語彙を日中辞書 Web サイト⁴⁾⁵⁾⁶⁾内でクエリとして検索し、取得できた主な日本語訳を示す。なお、「松软」のみ日本語訳が取得できなかったため、百度 汉语¹⁵⁾での検索結果を Google 翻訳で日本語化したものを主な日本語訳として示している。

- I. 中国語の形態素解析エンジンである jieba¹²⁾を用いて、各文を形態素 (語彙) に分割する。
- II. 分割した語彙をトークンとして追加した BERT の “bert-base-chinese” モデル¹³⁾の Fine-Tuning を、1 文ごとに分割した各レシピテキストを用いて行い、各語彙についての分散表現を集計する。
- III. 集計した各語彙の分散表現から、シズルワードの各既存対訳表現とのコサイン類似度を計算する。
- IV. 各既存対訳表現について、コサイン類似度が高い上位 10 個の語彙を、その既存対訳表現の類義語として取得する¹⁴⁾。
- (5) 取得した類義語が、類似翻訳表現として相応しいか検証する。
 - I. 中国語辞書 Web サイトである「百度 汉语¹⁵⁾」を用いて、取得した類義語をクエリとして検索する。
 - II. 各類義語についての解説ページがあり、かつ解説が既存対訳表現と似た意味を表している可能性がある¹⁶⁾と判断できれば、その類義語を類似翻訳表現の候補として取得する。
 - III. 候補となった類義語について、中国語を母語とする 1 名の大学生が確認し、既存対訳表現と意味が似ている語彙のみを類似翻訳表現とする。

4.3 分析結果と考察

表 2 に分析結果を示す。今回は類似翻訳表現として、計 5 語の語彙を取得できた。既存対訳表現との比較により、類似翻訳表現には以下の傾向が見られた。

- シズルワードの対訳表現ではないものの、意味が近い表現である
“上癮”に対する“可口”や、“鮮”に対する“嫩”は表現する内容は一致しないが、文中で共起しやすく、評価者も抽象的な意味では似ていると判断したと考えられる。

¹²⁾ <https://github.com/fxsjy/jieba>

¹³⁾ <https://huggingface.co/bert-base-chinese>

¹⁴⁾ 既存対訳表現の類義語として取得する語彙は、動詞、形容詞、副詞に限る。また、同じシズルワードへの別の既存対訳表現を類義語として取得した場合は、その語彙を類義語から除外する。

¹⁵⁾ 百度 汉语-更懂汉语, 更懂你: <https://dict.baidu.com/>

¹⁶⁾ 中国語で表記されている解説を Google 翻訳を用いて日本語化し、その内容から既存対訳表現と似た意味を表す可能性があると考えられるかどうかを、著者が判断する形で行った。

- シズルワードの対訳表現であるが、日中辞書から取得することができない

“松软”は“松”と同様に「柔らかい」を表現するが、日中辞書に解説ページが掲載されていない。また、“厚”は“濃”と同様に「濃厚である」を表現するが、「濃厚」をクエリとした日中辞書の検索結果には表示されない。

5 おわりに

本研究では、日中辞書サイトから取得したシズルワードの既存対訳表現と、各語彙の分散表現からコサイン類似度を計算することで取得した類似翻訳表現について分析を行った。分析の結果、計 66 語の既存対訳表現と、計 5 語の類似翻訳表現が取得された。また、取得した表現には以下の傾向があることを確認した。

- シズルワードの系統は、日中辞書からの既存対訳表現の取得可否に影響を与えない。
- 意味の似たシズルワードでも、取得できる既存対訳表現の数には差がある。
- 類似翻訳表現には、シズルワードの対訳表現ではないが、意味が近い表現である場合がある。また、シズルワードの対訳表現であるが、日中辞書から取得することができないシズルワードの対訳表現が存在する。

今後の課題として、より多くの類似翻訳表現を取得し、中国語での飲食情報提供の際の表現の幅を広げることが挙げられる。

謝辞

本研究の一部は、JSPS 科研費 18K11555 および 21K12138 の助成による。

参考文献

- [1] 日本政府観光局 (JNTO): 国籍/目的別 訪日外客数 (2004 年~2021 年), 日本政府観光局 (JNTO) (2022).
- [2] 国土交通省 観光庁: 訪日外国人の消費動向 訪日外国人消費動向調査結果及び分析 2019 年 年次報告書, 国土交通省 観光庁 (2020).
- [3] B・M・FT ことばラボ: Sizzle Word シズルワードの現在 2018 改訂「おいしいを感じる言葉」調査報告書, 株式会社 B・M・FT 出版部 (2018).
- [4] Inose, H.: Translating Japanese onomatopoeia and mimetic words, Translation and Re-search Project, pp.97-116 (2008).
- [5] 進藤綺乃, 吉野孝, 宮部真衣: 単語の分散表現を用いたおいしさを表す言葉の日英対訳の抽出, 研究報告グループウェアとネットワークサービス (GN), 2020-GN-109, No.37, pp.1-8 (2020).
- [6] 早川文代, 陳舜勝, 王錫昌, 李再貴, 齋藤昌義, 馬場康維, 横山雅仁: 中国語テキスト表現の収集と分類, 日本食品科学工学会誌, 2004 年 51 巻 3 号, pp.131-141 (2004).
- [7] Kato, D., Miyabe, M., Aramaki, E., Nadamoto, A.: Comparative Analysis of Sizzle Words on the Internet, iiWAS'14, pp.440-444 (2014).
- [8] 大橋正房, シズル研究会: 「おいしい」感覚と言葉食感の世代, 株式会社 B・M・FT 出版部 (2010).