

シズルワードに着目した Twitter 上のおいしさの表現の分析

加藤 大介[†] 宮部 真衣^{††} 荒牧 英治^{††} 灘本 明代[†]

[†] 甲南大学 知能情報学部 〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1

^{††} 京都大学 学際融合教育研究推進センター 〒 606-8507 京都府京都市左京区聖護院川原町 53

E-mail: [†]si071040@center.konan-u.ac.jp, ^{††}{mai.miyabe,eiji.aramaki}@gmail.com,

^{†††}nadamoto@konan-u.ac.jp

あらまし 近年, Twitter に代表されるマイクロブログサービスの普及により情報発信が容易になり, 食品の感想や紹介, 宣伝を気軽に発信することが可能になっている. そこで本論文では, Twitter には多くの人が食品の率直な感想を投稿していることに着目し, 食品のおいしさを表す「さくさく」や「芳醇」といったようなシズルワードを用いて, ツイートを分析する. さらに, 一般の Web で用いられているシズルワードとツイート上で用いられているシズルワードを比較する. これにより, ある食品に対する人々のおいしさの表現方法を抽出することが可能になり, 今後シズルワードを用いた食品検索等への応用が期待される.

キーワード Twitter, 食品, シズルワード

1. はじめに

「さくさく」や「新鮮」等, 食に関する表現は膨大な数があり, 食品の味や食感を伝える, 宣伝を行う等の場合に多種多様の表現が食品に対して使われている. これらは, 食品の味のイメージや判断要素としても重要であり, 特に食べたことのない食品の味のイメージや自分の食べたい物のイメージと実際の食品の味の比較の為に要素としてもこれら食に関する表現は非常に重要な意味合いを持つ. さらに, 近年では Web 検索でこれらの抽象的な表現や食品名をクエリとして検索を行い食品情報を取得することも一般的になっている. しかしながら, Web 上の情報は膨大でありユーザ自身で情報の取捨選択を行うことは大きな負担がかかる. さらにこのような味や食感をイメージする言葉から食品を検索する場合, 食品の情報が偏ってしまう. また, 食品名で検索した場合でも食品ごとに多様な表現が使われており, 味のイメージが掴みづらいといった問題が起こり, ユーザが求める情報が得られない場合が多い.

一方, 近年 Twitter に代表されるマイクロブログの普及により情報発信が容易になり, 人々の食品に対する感想や紹介, 宣伝を気軽に発信することが可能になっている. 特に Twitter ではユーザが様々な食品を食べた感想やユーザの食品に対するイメージ等を容易に発信することができ, このような情報が Twitter 上には豊富に存在している.

そこで本論文では, 「あっさり」「さくさく」「季節限定」といった, “おいしさ”を連想させる食に関する言葉である「シズルワード」[1] に注目し, Twitter 上でのシズルワードと食品の関連について分析を行う. さらに, 一般的な Web 上でのシズルワードと食品との関連も分析し, Twitter と一般の Web でのシズルワードと食品との関連を比較する. これらの分析を行うことにより, 食品の持つ味のイメージも分析する. 食品とシズルワードの関係を分析することで, 将来的に食品推薦や検索に役立てることが期待される.

また, 本論文では「”おいしさ”を連想させる食に関する言葉」に着目し, 抽象的な食べたい物のイメージから連想される食品について分析を行うため, 「まずさ」に関しては考慮しないものとする.

以下第 2 章には関連研究を, 第 3 章にはシズルワードの定義を, 第 4 章には Twitter 上のシズルワードの分析について, 第 5 章には Twitter と一般の Web のシズルワードの比較分析について, 第 6 章にはまとめと今後の課題を述べる.

2. 関連研究

食品に対する表現の研究は盛んに行われている. 大橋ら [1] は, 「”おいしさ”を感じさせる言葉」を「シズルワード」と定義し, アンケート調査を行いシズルワードから連想される食品の分析を行っている. 本論文では, 大橋らの分析結果のシズルワードの一部を用いる. また, 本論文ではアンケートではなく Twitter, Web サイトの情報をを用いてシズルワードから連想される食品について分析を行っている点, 及び食品から共起されるシズルワードについての分析を行っている点が異なる.

早川ら [2] は, 食品のテキストチャーに着目し, テク単語を 220 語選択しスチャー用語を辞書類, 専門書, 食に関するものの報告の文献, テクスチャー研究者へのアンケートから日本語テクスチャー用語を収集し, 出現頻度と妥当性, 表現の新旧についてまとめている. 本論文では, テクスチャーに関する表現以外にも着目する点, 表現だけでなく表現と食品の関連についても分析を行っている点が異なる.

Kanwipa ら [3] [4] は, 料理レシピ投稿サイトからオノマトペを用いてレシピを推薦するシステムを構築し, 食材, 調理法とオノマトペの関連について分析を行っている. 本論文では, オノマトペ以外の表現にも着目する点, 食材, 調理法ではなく食品そのものと表現の関連について分析する点, レシピではなく食品そのものを対象とする点が異なる.

清野ら [5] は, 食品のクチコミにオノマトペを使用すること

についての有用性を、特定の食品に対してのアンケート調査によって分析している。本論文では、オノマトペ以外の表現にも着目する点、複数の表現と複数の食品についての関連性について分析する点が異なる。

3. シズルワードの定義

大橋ら [1] は「『おいしい』感覚と言葉」において「おいしそう」「食べたい」「飲みたい」を感じる言葉を「シズルワード」と定義し、シズルワードとして単語を 220 語選択し、それらを味覚、食感、情報という 3 つの領域に分類し分析を行っている。さらに領域、年齢、男女ごとに美味しさを感じる言葉のランキング作成も行っている。本論文では「シズルワード」を大橋らの定義する「“おいしい”を連想させる食に関する言葉」とし、「『おいしい』感覚と言葉」中の 3 つの領域別シズルワードランキングからそれぞれ上位 30 語ずつを参照した。表 1 に本論文で用いたシズルワードを示す。分析を行うに当たって有用ではなかった味覚領域の「美味」を除き、味覚領域 29 語、食感領域 30 語、情報領域 31 語 (同率が含まれる為) の計 90 語のシズルワードを対象にして分析を行う。

4. Twitter 上のシズルワード表現分析

本論文では、シズルワードと食品の関連の分析に当たって 1025 品目 (内 254 品目は Wikipedia の「菓子の商品名」カテゴリ^(注1)に属する商品) からなる食品リスト、90 語からなるシズルワードリストを作成し、それらについて分析を行う。食品リストは単品の料理、菓子類、飲料からなり、調味料、食材は含まない。例えば「ステーキ」「味噌ラーメン」は食品リストに含むが「牛肉」「味噌」は食品リストに含まない。

4.1 分析の手法

データセット

2013 年 12 月 11 日～2013 年 12 月 12 日のツイートを対象として Twitter 上で各シズルワードをクエリとして検索を行い、各シズルワードを含むツイートをシズルワードごとに 1000 件と、2013 年 11 月 1 日のシズルワードを含まないツイート 90000 件を無作為に取得し、それらを分析に用いた。

分析手順

まず、上記データの中から各食品名が含まれるツイート数を X 、各シズルワードが含まれるツイート数を Y として、シズルワードと食品の単語共起頻度 J 、Simpson 係数 S 、Dice 係数 D をそれぞれシズルワード、食品ごとに求め、各シズルワードから見た食品のランキングを「シズルワード基準」、各食品から見たシズルワードのランキングを「食品基準」として分析を行う。

$$J = \frac{|X \cap Y|}{|X| + |Y|} \quad (1)$$

$$S = \frac{|X \cap Y|}{\min(|X|, |Y|)} \quad (2)$$

$$D = 2 \times \frac{|X \cap Y|}{|X| + |Y|} \quad (3)$$

上記 3 つの式全ての場合において $Y = 1,000$ また、Simpson 係数の性質上 $X = 0$ の場合正確な計算ができず $|X \cap Y|$ の値が極端に低い場合に逆にスコア係数が高くなってしまいうという問題があるため、 $X = 0$ の場合および $|X \cap Y| < 3$ の場合は Simpson 係数の計算は行わず、ランキングにも含めない。

4.2 結果と考察

シズルワード基準の単語共起度ランキングの結果を表 2 に、Simpson 係数の結果を表 3 に、Dice 係数の結果を表 4 に、食品基準の単語共起度ランキングの結果 (シズルワード基準のもの) を表 5 に、Simpson 係数の結果を表 6 に、Dice 係数の結果を表 7 示す。また、ランキング上位 5 件まで求められなかった場合、求められた値までを記載し、それ以下の順位は「-」を記載する。例えば、ある食品に対してシズルワードが 3 つしか共起しなかった場合、4 位以下は「-」と記載する。

「スープがかつおだしでさっぱり うまうま」の様に Twitter 上の食に関するツイートの多くはユーザが実際に食した時の感想と見られるものが多く、味に関しての評価は全体的に良好である。しかしシズルワードと食品名両方を含むツイートは全体的に少なく「ラーメンが美味しかった」「かつ丼おいしい」など「おいしい」という感想だけで味に対して言及していないツイートも多く見られる。これは Twitter が 140 文字以下の短文を投稿するサービスであり、ユーザも基本的に簡素な文を投稿する傾向にあるためと考えられる。更に取得できるツイートの中に含まれる食品数は時間帯によって差があり、一日の間で昼と夕方が比較的多くの食に関するツイートが投稿されているこいとがわかる。しかし、食品、シズルワードの時間帯ごとの規則性や特徴を見出すことはできなかった。また、季節により差が見られる場合がある。例えば、7 月～8 月の間のツイート^(注2)では、シズルワード「ぶりぶり」が含まれるツイートの中で、最も多く含まれる食品はどの時間帯であっても「エビ」だったが、11 月～12 月^(注3)の間のツイートではどの時間帯であっても「牡蠣」が最も多く含まれていた。理由としては、Twitter はリアルタイムな投稿が取得できるため、時期によっては旬の食品が上位に来る可能性があると考えられる。今回明確な差が取得できたのはこの「エビ」と「牡蠣」の場合のみだったが、さらなる分析を行う必要がある。

また、シズルワード基準で見た単語共起頻度、Simpson 係数、Dice 係数のランキングの結果を比較すると、Simpson 係数の結果にはシズルワードから連想されるものとして妥当でないものが他の 2 つの場合よりも多く含まれ、閾値を下回る場合が多い為、Simpson 係数を用いることは有用でないと考えられる。単語共起頻度と Dice 係数の結果は概ね類似しており、どちらもランキングとして妥当な結果になっている。この 2 つを比較した場合、ランキングとしての精度は Dice 係数の方が全体的に多少高い結果となっている。更にこれらのランキングではメ

(注2): 7 月, 8 月に取得したシズルワードを含むツイート 100 件ずつ

(注3): 11 月, 12 月に取得したシズルワードを含むツイート 100 件ずつ

(注1): <http://ja.wikipedia.org/wiki/Category:菓子の商品名>

表 1 使用したシズルワード一覧

領域	シズルワード
味覚	うまみ コクがある 香ばしい 風味豊か クセになる やみつき 濃厚 味わい深い 深みのある まろやか スパイシー 飽きのこない クリーミー こんがり 後味がよい あっさり 後味すっきり ピリ辛 スイート さっぱり マイルド 脂の乗った リッチな 芳醇 デリシャス さわやか 後を引く 甘い すっきり
食感	ジューシー もちもち もっちり とろける サクサク ホクホク とろーり シャキシャキ コシのある 口溶け サクッと ふっくら ふわふわ ふんわり カリッと なめらか トロトロ じゅわー あつあつ プリプリ ふわっと パリッと さっくり ホカホカ とろっと シャキッと しっとり 舌触りのよい カリカリ カラッと
情報	焼きたて 季節限定 揚げたて 新鮮 炊きたて 出来たて 取れたて 旬 産地直送 絶品 天然の 厳選素材 秘伝の こだわりの 産地限定 完熟 贅沢な 無添加 手作り 本格的 体にやさしい 本場の プレミアム 食べごろ 昔ながらの味 熟成 自家製 ヘルシー 自然の 無農薬 フレッシュ

表 2 Twitter 上における単語共起頻度のランキング例 (シズルワード基準)

シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度
揚げたて	ポテト	0.044	シャキシャキ	レタス	0.014	とろける	チーズ	0.020
	チキン	0.036		キャベツ	0.014		グラタン	0.010
	ラム	0.032		大根	0.005		チョコ	0.008
	ベーコン	0.032		ラーメン	0.005		トマト	0.007
	コロッケ	0.005		サラダ	0.005		ケーキ	0.006

表 3 Twitter 上における Simpson 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数
揚げたて	唐揚げ	0.833	シャキシャキ	レタス	2.000	とろける	グラタン	4.250
	ソーダ	0.556		水菜	0.750		トマト	0.596
	ジンジャー	0.500		大根	0.391		チーズ	0.453
	天ぷら	0.462		キャベツ	0.318		せんべい	0.273
	ポテト	0.306		玉ねぎ	0.278		チョコレート	0.150

表 4 Twitter 上における Dice 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数
揚げたて	ポテト	0.307	シャキシャキ	レタス	0.2	とろける	グラタン	0.175
	チキン	0.272		キャベツ	0.182		チーズ	0.117
	ベーコン	0.267		大根	0.071		トマト	0.086
	ラム	0.266		サラダ	0.069		チョコレート	0.061
	カツ	0.063		キムチ	0.058		ハム	0.056

表 5 Twitter 上における単語共起頻度のランキング例 (食品基準)

食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度
餃子	焼きたて	0.017	ドーナツ	揚げたて	0.006	エビ	カラッと	0.005
	パリッと	0.004		サクサク	0.004		揚げたて	0.005
	カリッと	0.003		やみつき	0.003		サクサク	0.004
	ピリ辛	0.003		しっとり	0.003		プリプリ	0.003
	ジューシー	0.003		出来たて	0.002		ピリ辛	0.002

ジャーな食品が上位を占めており、マイナーな食品や、より詳細な食品名のは上位に含まなかった。メジャーな食品が上位にあるのは、ツイートはユーザの感想であることが多い為、広く一般的に食べられているものほど多くのユーザが感想を投稿している為と考えられる。また、詳細な食品名が上位に含まれていないのは、Twitter には余り長文を投稿せず簡潔な文で投稿する傾向にある為、例えば「ミートソーススパゲティ」を「スパゲティ」や「パスタ」のように省略して投稿するなど、詳細な料理名を書かない場合も多い為と考えられる。

さらに、食品基準で見た単語共起頻度、Simpson 係数、Dice 係数のランキングの結果を比較すると、シズルワード基準でみた場合と同様、Simpson 係数の結果は計算できたものについて

は概ね良好な結果になる。しかしながら、閾値を下回っている値について計算を行わない為、他 2 つの場合に比べてあまり良い結果が出ない場合があるという問題がある。食品基準でみた場合も単語共起頻度と Dice 係数の結果は概ね類似しているが、全体的に Dice 係数の方が単語共起頻度の場合よりも情報系のシズルワードが上位に来る傾向にある。食品基準でみた場合のランキングは、情報系のシズルワードがあまり上位に含まれないという特徴が明確にある。これは情報系のシズルワードは宣伝や紹介の場面で多く使われるが、Twitter 上には宣伝目的の投稿の方が少数であるためであると考えられる。

表 6 Twitter 上における Simpson 係数のランキング例 (食品基準)

食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数
餃子	焼きたて	0.229	ドーナツ	揚げたて	0.3	エビ	揚げたて	0.1351
	パリッと	0.060		サクサク	0.2		カラッと	0.1351
	カリッと	0.036		しっとり	0.15		サクサク	0.1081
	ジューシー	0.036		やみつき	0.15		プリプリ	0.0811
	ピリ辛	0.036		-	-		-	-

表 7 Twitter 上における Dice 係数のランキング例 (食品基準)

食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数
餃子	焼きたて	0.021	ドーナツ	揚げたて	0.013	エビ	カラッと	0.013
	パリッと	0.009		サクサク	0.004		ピリ辛	0.006
	手作り	0.005		ヘルシー	0.004		自家製	0.004
	香ばしい	0.005		やみつき	0.004		揚げたて	0.002
	絶品	0.004		季節限定	0.002		贅沢な	0.002

5. Twitter と一般の Web のシズルワードの比較分析

Twitter は一般の Web と異なり、ユーザの率直な感想が多く述べられていると考え、一般の Web でのシズルワードと食品との関係と Twitter のそれとを比較することを行う。

5.1 一般の Web 上のシズルワード表現分析

第 4 章で用いた食品リストとシズルワードリストを用いて分析を行う。また一般的な Web の分析対象として Google による Web 検索結果上位 100 件を用いる。

5.1.1 分析の手法

一般的な Web 上でのシズルワードと食品の関係の分析を行うにあたり、はじめに「”シズルワード” AND ”食品名”」をクエリとして Google による検索結果数を取得し、シズルワードごとの比較を行った。その結果、例えば「こってり、ラーメン」といったように、シズルワードに対して検索結果数上位に含まれる食品は同じようなものばかりであり、シズルワードごとに特徴を見いだすことは出来なかった。また、例えば「さくさく、酒」のように各シズルワードに対し妥当とは言い難い食品ばかりが上位を占める結果となった。上位に含まれる食品は一般的なものが殆どであり、Web ページの量そのものが大きく影響していると考えられる。

そこで、食品リストの「食品名」をクエリとして Google 検索を行い、食品ごとに 100 件のスニペット、シズルワードごとに 100 件のスニペットを取得し、各スニペット中にシズルワードが含まれているかを正規表現によるマッチングを行い検証する。各シズルワードの含まれるスニペット件数を X 、各食品名の含まれるスニペット件数を Y として共起頻度 (1) 式、Simpson 係数 (2) 式、Dice 係数 (3) 式を用い、それぞれシズルワード、食品ごとに求め、各シズルワードから見た食品のランキングを「シズルワード基準」、各食品から見たシズルワードのランキングを「食品基準」として分析を行う。

5.1.2 結果と考察

食品名で検索した場合のシズルワード基準のランキング結果として、表 8 に一般の Web 上における単語共起頻度の結果

を、表 9 に一般の Web 上における Simpson 係数の結果を、表 10 に一般の Web 上における Dice 係数の結果を示す。食品名で検索した場合の食品基準のランキング結果として、表 11 に一般の Web 上における単語共起頻度の結果を、表 12 に一般の Web 上における Simpson 係数の結果を、表 13 に一般の Web 上における Dice 係数の結果を示す。さらに、シズルワードで検索した場合のシズルワード基準のランキング結果として、表 14 に一般の Web 上における単語共起頻度の結果を、表 15 に一般の Web 上における Simpson 係数の結果を、表 16 に一般の Web 上における Dice 係数の結果を示す。シズルワードで検索した場合の食品基準のランキング結果として、表 17 に一般の Web 上における単語共起頻度の結果を、表 18 に一般の Web 上における Simpson 係数の結果を、表 19 に一般の Web 上における Dice 係数の結果を示す。ランキング上位 5 件まで求められなかった場合、求められた値までを記載し、それ以下の順位は「-」を記載する。

食品名で検索した場合の結果をみると、シズルワード基準でみた場合、共起頻度と Dice 係数の結果は全体的に似通っており、ランキングの相違点を見比べても結果の良し悪しに大きな差は見られない。Simpson 係数の結果は共起頻度、Dice 係数のランキングと順序が異なっている場合が多く、上位を占める食品の妥当性について Simpson 係数は共起頻度、Dice 係数と比べ悪い結果であった。更にどの場合も「カラッと、ゴーヤチャンプルー」の様にランキング上位にマイナーな食品、非常に具体的な食品名が多く含まれている。これは一般的な食品の名称ではあるが、種類が多岐に渡るようなクエリ、例えば「パスタ」、「パン」などは多くの食品を含むため、味の表現が大きくばらつき、食品の味に言及していない場合が多い為、そういった食品よりも共起頻度、Simpson 係数、Dice 係数のランキングではクエリが具体的な食品の方が上位になりやすいと考える。さらに、「ボンゴレビアンコスパゲッティ」の様にクエリとなる食品名がより具体的であればあるほどその食品についての説明、紹介をするページが多く含まれるようになり、その説明の中にシズルワードが用いられることが多くなる為であると考えられる。食品基準でみた場合、Simpson 係数の計算をする

表 8 食品名で検索した場合の Web 上における単語共起頻度のランキング例 (シズルワード基準)

シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度
揚げたて	サターアUNDERギー	0.214	シャキシャキ	ハムタマゴロール	0.167	とろける	湯豆腐	0.075
	天ぷら	0.212		水菜	0.143		オニオングラタンスープ	0.065
	串カツ	0.071		照焼き	0.125		ネクター	0.060
	フィッシュアンドチップス	0.069		みょうが	0.085		エビクリームコロッケ	0.057
	ガンモドキ	0.061		レンコン	0.067		カスタード	0.047

表 9 食品名で検索した場合 Web 上における Simpson 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数
揚げたて	ガンモドキ	0.286	シャキシャキ	水菜	0.448	とろける	湯豆腐	0.200
	天ぷら	0.269		チンジャオロース	0.273		ライム	0.200
	串カツ	0.250		三つ葉	0.250		サーモン	0.200
	サターアUNDERギー	0.214		ミョウガ	0.235		ハムソテー	0.167
	フィッシュアンドチップス	0.154		緑豆	0.222		サーロイン	0.167

表 10 食品名で検索した場合 Web 上における Dice 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数
揚げたて	天ぷら	0.224	シャキシャキ	水菜	0.109	とろける	湯豆腐	0.048
	串カツ	0.075		みょうが	0.061		オニオングラタンスープ	0.040
	サターアUNDERギー	0.067		ごぼう	0.043		ネクター	0.039
	ガンモドキ	0.056		レンコン	0.039		カスタード	0.031
	オニオンリングフライ	0.042		ミョウガ	0.037		ゴーダチーズ	0.024

表 11 食品名で検索した場合 Web 上における単語共起頻度のランキング例 (食品基準)

食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度
餃子	手作り	0.009	ドーナツ	手作り	0.007	エビ	ブリブリ	0.002
	ジューシー	0.005		こだわりの	0.004		自然の	0.002
	パリッと	0.002		秘伝の	0.003		旬	0.002
	厳選素材	0.002		しっとり	0.003		濃厚	0.001
	絶品	0.001		ヘルシー	0.003		ふんわり	0.001

表 12 食品名で検索した場合 Web 上における Simpson 係数のランキング例 (食品基準)

食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数
餃子	ジューシー	1	ドーナツ	手作り	1	エビ	-	-
	手作り	0.909		しっとり	1		-	-
	-	-		こだわりの	1		-	-
	-	-		とろける	1		-	-
	-	-		無添加	1		-	-

際多くの場合で閾値を下回ってしまい、参考になるような結果が得られない。共起頻度と Dice 係数の結果には殆ど差が出ないものの、どちらの結果も良好である。また、情報系のシズルワードが比較的上位に入りやすい傾向にある。そして、取得したスニペットは Web ページは個人の料理のレシピサイト、飲食店のサイト、食品販売を行う店舗のサイト、通販サイト、料理の口コミサイト、個人の料理に関する感想のブログなどが主であり、特に店舗等販売目的の宣伝のページとレシピサイトが比較的多く取得できた。これらのサイトはいずれも食品に関して詳細な紹介、説明、宣伝を行うことが目的である為、シズルワードのように「おいしさを感じさせる」表現が多く使用さ

れている。特に店舗、通販サイトなどの宣伝目的のサイトのスニペットには「新鮮」「季節限定」などの情報系のシズルワードが非常に多く含まれた。

一方、シズルワードで検索した場合の結果をみると、こちらもシズルワード基準でみた場合、共起頻度と Dice 係数の結果は殆ど変わらない。Simpson 係数の結果にはシズルワードから連想されるものとして妥当でないものが他の 2 つの場合よりも多く含まれる。そしてどの場合を見ても、全体的に食品で検索した場合とは逆に一般的な食品ほど上位に含まれる結果になり、具体的な料理名、マイナーな料理は殆ど上位に含まれなかった。これはシズルワードで検索した場合、シズルワードと共起し易

表 13 食品名で検索した場合 Web 上における Dice 係数のランキング例 (食品基準)

食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数
餃子	手作り	0.018	ドーナツ	手作り	0.015	エビ	ブリブリ	0.004
	ジューシー	0.011		こだわりの	0.007		自然の	0.004
	パリッと	0.004		秘伝の	0.006		旬	0.004
	厳選素材	0.004		しっとり	0.006		濃厚	0.002
	絶品	0.002		ヘルシー	0.006		ふんわり	0.002

表 14 シズルワードで検索した場合の Web 上における単語共起頻度のランキング例 (シズルワード基準)

シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度	シズルワード	食品名	共起頻度
揚げたて	天ぷら	0.194	とろける	チーズ	0.114	シャキシャキ	サラダ	0.097
	ポテト	0.140		クリーム	0.079		もやし	0.050
	ポテトチップ	0.130		生クリーム	0.050		レタス	0.040
	カルピ	0.099		牛乳	0.040		ごぼう	0.040
	ドーナツ	0.050		コーヒー	0.030		ベーコン	0.030

表 15 シズルワードで検索した場合 Web 上における Simpson 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数	シズルワード	食品名	Simpson 係数
揚げたて	天丼	1.000	とろける	牛乳	1.000	シャキシャキ	レタス	1.000
	ポテトチップ	1.000		コーヒー	1.000		ごぼう	1.000
	ポテト	1.000		クリーム	0.889		もやし	0.833
	ドーナツ	1.000		生クリーム	0.833		サラダ	0.769
	スコーン	1.000		チーズ	0.706		ベーコン	0.750

表 16 シズルワードで検索した場合 Web 上における Dice 係数のランキング例 (シズルワード基準)

シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数	シズルワード	食品名	Dice 係数
揚げたて	天ぷら	0.374	とろける	チーズ	0.291	シャキシャキ	サラダ	0.230
	ポテト	0.246		クリーム	0.158		もやし	0.113
	ポテトチップ	0.230		生クリーム	0.113		キャベツ	0.113
	カルピ	0.198		ケーキ	0.113		レタス	0.077
	ドーナツ	0.095		牛乳	0.077		ベーコン	0.077

表 17 シズルワードで検索した場合 Web 上における単語共起頻度のランキング例 (食品基準)

食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度	食品名	シズルワード	共起頻度
餃子	もちっと	0.006	ドーナツ	揚げたて	0.005	エビ	ブリブリ	0.005
	パリッと	0.004		出来たて	0.003		じゅわー	0.003
	カリッと	0.003		クリーミー	0.003		揚げたて	0.003
	厳選素材	0.003		体にやさしい	0.002		カリカリ	0.002
	やみつ	0.002		モチモチ	0.001		風味豊か	0.002

い食品は話題に挙がりやすい一般的な食品の方が多く取り上げられているからだと考えられる。更に食品基準でみた場合、こちらは Simpson 係数の値が多くの食品について取得できない為、結果として殆ど参考になれない。共起頻度と Dice 係数の場合は良好だが、食品で検索した場合と変わらず結果に殆ど差がつかない。そして食品で検索した場合と比べ、情報系のシズルワードはあまり上位結果になる。これは食品で検索した場合よりも個人のブログ、レシピサイトなどが多く取得できた為、宣伝目的の情報系のシズルワードの割合が減った為であると考えられる。

5.2 Twitter と Web 上のシズルワードの比較

「もちり」というシズルワードに対して、Twitter では「焼きとり」という食品が取得でき、一般の Web では「イングリッシュマフィン」という食品が取得できた。つまり、Twitter を用いて取得した食品の多くは一般的なものであり、具体的な食品名ほど取得できなかったのに対し、Google 検索で取得した Web ページのスニペットでは具体的な食品名ほどシズルワードと共起していることがわかった。Twitter への投稿は簡潔な短文が多く、ユーザも多くの場合食品を詳細に説明、紹介することが目的でないため食品に対してシズルワードのような語を用いた食品の説明を行うことは少ない。さらに多くのユーザが食

表 18 シズルワードで検索した場合 Web 上における Simpson 係数のランキング例 (食品基準)

食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数	食品名	シズルワード	Simpson 係数
餃子	-	-	ドーナツ	-	-	エビ	ブリブリ	2.500
	-	-		-	-		じゅわー	1.000
	-	-		-	-		揚げたて	1.000
	-	-		-	-		-	-
	-	-		-	-		-	-

表 19 シズルワードで検索した場合 Web 上における Dice 係数のランキング例 (食品基準)

食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数	食品名	シズルワード	Dice 係数
餃子	もちっと	0.012	ドーナツ	揚げたて	0.010	0.006	ブリブリ	0.010
	パリッと	0.008		出来たて	0.006		じゅわー	0.006
	カリッと	0.006		クリーミー	エビ		揚げたて	0.006
	厳選素材	0.006		体にやさしい	0.004		カリカリ	0.004
	やみつき	0.004		モチモチ	0.002		風味豊か	0.004

べているものほど投稿が多くなるため、一般的な食品が多く取得できたと考えられる。

反対に Web サイトでは食品の感想、調理法、紹介、宣伝が目的であることが多く、より詳細、具体的な食品の説明や「おいしいそう」と思わせるような説明が必要とされる場合が多く、Twitter よりも豊富な表現が食品に対して使われている。更に一般的でない食品であっても情報が必ず存在するため、食品のメジャー度や、食べられている量に関わらず食品の情報が取得できた。しかし具体的な食品名ほどシズルワードと共起する結果となり、シズルワードからの連想と言う点を考えると多くの人がすぐに思いつかないような食品が共起した為、この点に関しては Twitter の方が良質な情報が取得できたと考えられる。更に全体的に宣伝文句としてよく使われるような味に直接関係のない情報系のシズルワードが Twitter に比べて非常に多く使われている。

また、シズルワードと食品の共起については大きな特徴の差が出たものの、どちらも妥当性は良好なものであった

6. まとめと今後の課題

本論文では、「おいしいさ」を連想させる食に関する言葉を「シズルワード」と定義し、Twitter 上と一般的な Web 上の情報に対し、食品の表現に用いられるシズルワードについて分析を行った。シズルワードから連想される食品としては Twitter、一般的な Web 共に比較的良好な結果が取得できたが、それぞれで大きく特徴が異なる結果となった。シズルワードのような食品に関する表現と食品の関連性を分析することで、シズルワードを食品推薦などに利用することが期待できるようになる。

今後の課題は以下の通りである。

- 今回の結果を元に、実際に食品推薦、検索システムを構築する。
- Twitter はリアルタイムな投稿がされるサービスであるが、本論文では時間や季節による違いを明確化することができなかった為、これらに着目した分析を今後行っていく必要がある。
- 一つのツイートの中に複数の食品、シズルワードが存在

するとそれら全てが共起してしまうので、対策を講じる必要がある。

- 分析対象として本論文では Twitter と Google の検索結果を用いたが、その他の有用な分析対象を検討する必要がある。

文 献

- [1] 大橋正房, 武藤彩加, 山本真人, 爲国正子, 汲田亜紀子, 洪澤文明, 小川裕子, 「おいしい」感覚と言葉 食感の世代, BMFT 出版部, 2010
- [2] 早川文代, 井奥加奈, 阿久澤わゆり, 齋藤昌義, 西成勝好, 山野善正, 神山かおる, 「日本語テクスチャー用語の収集」, 日本食品科学工学会誌, No.8, pp.337-346, 2005
- [3] Kanwipa Lertsumruaypun, 渡辺知恵美, 中村聡史, オノマトベロリ:オノマトベを利用した料理推薦システム, 情報処理学会研究報告, No.6, pp1-7, 2009
- [4] Kanwipa Lertsumruaypun, 渡辺知恵美, 中村聡史, レシピ検索システムにおけるオノマトベとレシピ用語集合の関連付け, 情報処理学会研究報告. データベース・システム研究会報告, No.15, pp.1-8, 2010
- [5] 清野誠喜, 玉置怜, 滝口沙也加, 食品のクチコミにおけるオノマトベの効果, 農林業問題研究, pp.249-254, 2011