

テキストマイニングを用いた味覚表現の
意図理解に関する研究

2023年 3月

立花 伸一

テキストマイニングを用いた味覚表現の
意図理解に関する研究

立花 伸一

理工情報生命学術院

システム情報工学研究群

筑波大学

2023年 3月

論文概要

本論文は、一般生活者の味覚によって判断され、用いられる、明確に定義ができない味覚表現の共通的な意図理解を目的とした、定量的評価手法の提案を行ったものである。

近年、消費者をとりまく経済・社会的環境の変化に伴い、食の外部化、簡便化、高級化など食生活の変化が進み、消費者は自らのライフスタイルに合わせた食生活を志向する傾向にあるとされ、食へのニーズはますます多様化、高度化しているといわれる。そうした環境下、飲食業界にとっては、消費者へのPR戦略の一つとして、商品の魅力を、「激甘」「ほろ苦」などといった直接的な味に関する表現の他、「サクサク」「とろり」などといったオノマトペと呼ばれる擬音語や擬態語、「もちもち」「ジューシー」などといったおいしさを感じさせる言葉といわれるシズルワードなどを用いて表現することで、商品の差別化や販売促進が図られ、様々なPR効果が発揮されているといわれる。加えて、消費者間のSNSや口コミ投稿サイトなどによるコミュニケーションの重要性の高まりを背景に、PRされた商品のおいしさが、消費者にとってどのように感じられたのかといったことを、SNS投稿文や口コミ情報の分析・評価により、商品開発に繋げる動きがみられ、おいしさを表す表現に注目が集まっている。

一般的に、味覚は、味、香り、食感、色など味わう対象から五感を通じて感じられるものや、食の経験や味わう際の環境、各種媒体などから収集される食に関する情報など様々な要因が複雑に作用して、形成されると考えられている。そのため、人間の持つ感性の一つとなる味覚には個人差があるとされる。その一方で、多くの一般生活者にとっては、飲食物は、個人または家族や友人、職場の同僚などの身近な存在と味わう機会が主であり、その飲食物から感じられた具体的、詳細なおいしさを他者に積極的に伝える必要性がなく、おいしいやうまい、まずいといった一般的な表現を用いて表現することが多い。そ

のため、飲食物から感じたおいしさを具体的、詳細に表現する機会は少なく、実際に、飲食物の味を表現する場合、その味をおいしいやまずいといった一般的な表現以外を用いて表現することが難しいといった事態が生じる。例えば、日本酒は、その銘柄を選ぶ場合、多くは甘口や辛口、味の濃淡などを示す指標や、その味を説明したものを参考にすることが多い。一般的に、日本酒は糖分が多く含まれたものが甘口、糖分が少ないものが辛口といわれる。しかし、甘口は甘みを感じるものとして容易に他人と共有できる一方で、辛口は辛いと感じる成分が含まれているわけではなく、その味を何の味であるか表現しづらいものといえる。さらには、説明された味の情報として多く用いられる、キレやフルーティ、コク、まろやかなどといった表現は、明確な定義がないものである。また、たとえおいしいやまずいといった一般的な表現以外を用いて、おいしさを表現することがあったとしても、同一の飲食物の味を表現する際には、異なる味覚に関する表現が用いられるといった事態が生じることがある。例えば、味覚に関して「甘み」や「辛み」「コク」などといった表現を用いる。この中で、「甘み」が砂糖の味であり「辛み」が唐辛子の味であるように容易に他人と共有できる言葉である一方、その逆に「コク」のように何の味であるかが定義しづらい言葉もある。これはコクに対して明確な定義がないためである。しかし、コクに類似する味覚表現として、「深みのある味」や「濃厚な味」、「複雑な味」などを想起することができる。また、コクのある食品として連想されるものとしては、煮込み料理、ラーメン、カレーやビールなど広範囲にわたり、これらは、「深みのある味」や「濃厚な味」、「複雑な味」としても連想が可能といえる。コクのニュアンスや感覚などは、曖昧であるにもかかわらず、おおよそ我々の多くに共有されているものといえる。

こうした背景から、飲食業界などでは自社商品の味覚的な特徴を、多くの消費者へ共有させる必要があるが、一般生活者各人が各々の味覚によって判断して用いられる味覚表現は、曖昧性を含んでいるために、客観的な評価を困難にするという問題が生じる。そこで、本研究では、一般生活者各人の味覚によって判断された味覚に関する表現が記載された、感性の情報源といえる口コミ情報を対象に、テキストマイニングによって、各人の味覚によって判断され用

いられる、明確に定義ができない味覚表現の共通的な意図理解を目的とした、一連の定量的評価手法の提案を行った。

第3章では、日本酒を味わった際の明確に定義ができない味覚表現の共通的な意図理解を目的とした定量的評価手法の提案を行った。重回帰分析を用いた日本酒の味覚表現の検証の結果、辛口とされる日本酒の場合、辛口の代わりに、キレやサッパリといった用語で表現される可能性があることが確認された。また、甘口とされる日本酒の場合、なめらかやフルーティ、やわらかといった表現が用いられることが確認された。その他、日本酒の味に濃醇さが感じられる場合、甘味やコク、まろやかと表現される傾向にあることが示唆された。一連のテキストマイニングによる分析手法によって、「日本酒度」「酸度」「甘辛度」「濃淡度」と、日本酒に関する各種味覚表現における関係性について、定量的な評価を可能にした。

第4章では、曖昧な味覚表現の例としてコクを取り上げ、明確に定義ができない味覚表現に対する、共通的な意図解明を目的とした定量的評価手法の提案を行った。本実験では、コクと対の表現になることが確認された、サッパリ表現との比較を行う形で一連の分析を進めた。まず、独自の変更でBERTに適用させたLIMEを用いて、テキスト分類に寄与する特徴語を可視化することで、コクに関係するとされる特定の食材の存在を確認した。次に、それら特定の食材から感じられるとした、コク表現以外の定義された味覚やテクスチャー用語を抽出し、それらの共起頻度から算出した使用率を用いて、t検定による比較を実施した。そして、最終的にコクの特性として、油脂の濃厚感が大きく寄与することを確認した。また、一般生活者は、クリーミー、こったり、まったり、濃厚といった表現を用いて、明確に定義のできない味覚表現であるコクと同等の感覚を表現している可能性が示唆された。一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現の共通的な意図理解のための定量的評価を可能にした。

第5章では、明確に定義ができない味覚表現が、どのように共有され使用されるのかといった合意形成に至る点に関して評価を行うため、料理レシピ共有サイトの口コミ情報を対象に、明確に定義ができない味覚表現は、他者情報

からの影響を受けて、一般生活者の間で合意形成されていくことの傾向有無を示す、定量的評価手法を提案した。前章と同様に曖昧な味覚表現の例としてコクを取り上げ、ロジスティック回帰分析および材料の類似度を用いた比較分析を行った。分析結果、他者情報としてコクに関する情報がレシピ上に用いられる場合、該当レシピのロコミ上において、コクという味覚表現が有意に用いられる傾向にあることが確認された。一連のテキストマイニングによる分析手法によって、明確に定義ができない味覚表現は、他者の影響を受けて一般生活者の間で合意形成されていくことを示す可能性があることを示す定量的評価を可能にした。

これまで生化学分野を中心に、味覚を明らかにすることは試みられてきたが、これらは化学的な分析による手法であった。本論文で示した一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現の共通的な意図理解に向けた定量的評価を可能にしたといえる。

目次

第 1 章 緒論	1
第 2 章 テキストマイニングを用いた味覚表現の解明における課題	4
2.1 はじめに.....	4
2.2 味覚表現の分類タスクに関する自然言語処理	4
2.2.1 自然言語処理の流れ.....	4
2.2.2 BERT の概念.....	16
2.2.3 LIME の概念	18
2.3 世界中で利用される料理レシピ共有サイト.....	20
2.4 テキストマイニングと料理・味覚表現に関する先行研究	22
2.4.1 生化学分野におけるコクに関する研究	22
2.4.2 日本酒における味覚に関する研究	26
2.4.3 認知言語学・身体化認知に関する研究	26
2.4.4 食に関わる言葉に関する研究	27
2.4.5 料理レシピを対象にした自然言語処理に関する研究	29
2.5 テキストマイニングと料理・味覚表現における研究分野の課題	33
2.6 おわりに	35
第 3 章 飲食物への味覚表現と味覚関連指標の関係性に関する定量的評価 ..	36
3.1 はじめに.....	36
3.2 日本酒の味覚表現の検証.....	37
3.2.1 使用データ	37
3.2.2 分析方法	38
3.2.3 実験環境	39
3.2.4 データセットの作成.....	39
3.3 重回帰分析による検証結果.....	41
3.3.1 各変数間の相関分析.....	41
3.3.2 影響要因の重回帰分析	43
3.3.3 実験結果の考察	48
3.4 まとめ	49

第 4 章 明確に定義ができない味覚表現の意図理解に向けた定量的評価	50
4.1 はじめに	50
4.2 曖昧な味覚表現における特徴語の分析	50
4.2.1 使用データ	50
4.2.2 分析方法	53
4.2.3 実験環境	53
4.2.4 データセットの作成.....	53
4.2.5 BERT によるテキスト分類	55
4.2.6 LIME による特徴語の抽出.....	56
4.2.7 実験結果の考察	59
4.3 特徴語の味覚要素とテクスチャー特性の解明	59
4.3.1 使用データ	59
4.3.2 分析方法	60
4.3.3 実験環境	61
4.3.4 アソシエーション分析	61
4.3.5 共起用語の抽出	63
4.3.6 使用率の算出.....	66
4.3.7 使用率の検定.....	68
4.3.8 コクの特徴語として確認された食品成分の分析	69
4.3.9 実験結果の考察	73
4.4 おわりに	74
第 5 章 他者情報が味覚表現に与える影響の定量的評価手法	75
5.1 はじめに	75
5.2 他者情報が味覚表現に与える影響の検証	76
5.2.1 使用データ	76
5.2.2 分析方法	76
5.2.3 実験環境	77
5.2.4 データセットの作成.....	77
5.3 ロジスティック回帰分析による検証結果	80

5.3.1	料理カテゴリーの属性比較	80
5.3.2	味覚表現の属性比較.....	81
5.3.3	ロジスティック回帰分析の実行.....	82
5.4	料理レシピ材料の類似度を用いた比較分析	83
5.4.1	データセットの作成.....	83
5.4.2	比較分析の実行	86
5.4.3	実験結果の考察	86
5.5	おわりに	87
第 6 章	結論	88
	謝辞.....	92
	参考文献.....	93
	関連業績.....	110

図表目次

図 2-1	word2vec における CBOW と Skip-gram の概要 ([Mikolov 2013a]より抜粋)	7
図 2-2	RNN のイメージ図	10
図 2-3	LSTM の構造	11
図 2-4	Seq2seq の概要	13
図 2-5	Attention の概要 ([Luong 2015]より抜粋)	14
図 2-6	Transformer の概要 ([Vaswani 2017]を一部編集)	15
図 2-7	BERT における事前学習とファインチューニング	17
図 2-8	LIME のイメージ図 ([Ribeiro 2016]を一部編集)	18
図 2-9	LIME 出力イメージ図 ([Ribeiro 2016]を一部編集)	20
図 3-1	日本酒度の概要	36
図 3-2	楽天市場における口コミ情報例	37
図 4-1	クックパッドの口コミ情報例	51
図 4-2	コクラベルの出力結果例	57
図 4-3	サッパリラベルの出力結果例	57
図 5-1	クックパッドの投稿レシピ画面の概要	79
表 2-1	料理レシピサイトの世界ランキング	21
表 3-1	楽天市場の日本酒に関するデータ概要	38
表 3-2	目的変数の概要	38
表 3-3	抽出された味覚表現	40
表 3-4	本データセットの基礎統計量	41
表 3-5	各目的変数と説明変数間の相関係数	42
表 3-6	各説明変数間の相関係数	43
表 3-7	日本酒度を目的変数とした重回帰分析	44
表 3-8	酸度を目的変数とした重回帰分析	45
表 3-9	甘辛度を目的変数とした重回帰分析	46
表 3-10	濃淡度を目的変数とした重回帰分析	47

表 4-1	クックパッドのデータ概要	51
表 4-2	コク表現と逆接の助詞および接続詞と共起する単語	52
表 4-3	BERT によるテキスト分類の実験結果	55
表 4-4	LIME による出力結果	58
表 4-5	アソシエーション分析結果の例	62
表 4-6	味覚に関する抽出用語	63
表 4-7	力学的特性に関する抽出用語	64
表 4-8	幾何学的特性に関する抽出用語	65
表 4-9	その他の特性に関する抽出用語	66
表 4-10	コクに寄与する食材と共起する用語の使用率	67
表 4-11	サッパリに寄与する食材と共起する用語の使用率	67
表 4-12	t 検定結果	69
表 4-13	実験対象となる食材	70
表 4-14	コクに寄与する食材の成分含有量 (100g あたり)	71
表 4-15	サッパリに寄与する食材の成分含有量 (100g あたり)	72
表 4-16	t 検定結果	72
表 5-1	クックパッドのデータ概要	76
表 5-2	抽出された味覚表現	78
表 5-3	コク情報有無別の料理カテゴリーの出現状況	81
表 5-4	味覚表現に関する基礎統計量	82
表 5-5	ロジスティック回帰分析の結果	83
表 5-6	抽出した 28 組のレシピ名・材料および Jaccard 係数	85
表 5-7	t 検定結果	86

第1章 緒論

近年、消費者をとりまく経済・社会的環境の変化に伴い、食の外部化、簡便化、高級化など食生活の変化が進み、消費者は自らのライフスタイルに合わせた食生活を志向する傾向にあるとされ、食へのニーズはますます多様化、高度化しているといわれる。そうした環境下、飲食業界にとっては、消費者への PR 戦略の一つとして、商品の魅力を、「激甘」「ほろ苦」などといった直接的な味に関する表現の他、「サクサク」「とろり」などといったオノマトペと呼ばれる擬音語や擬態語、「もちもち」「ジューシー」などといったおいしさを感じさせる言葉といわれるシズルワードなどを用いて表現することで、商品の差別化や販売促進が図られ、様々な PR 効果が発揮されているといわれる。加えて、消費者間の SNS や口コミ投稿サイトなどによるコミュニケーションの重要性の高まりを背景に、PR された商品のおいしさが、消費者にとってどのように感じられたのかといったことを、SNS 投稿文や口コミ情報の分析・評価により、商品開発に繋げる動きがみられ、おいしさを表す表現に注目が集まっている。

一般的に、味覚は、味、香り、食感、色など味わう対象から五感を通じて感じられるものや、食の経験や味わう際の環境、各種媒体などから収集される食に関する情報など様々な要因が複雑に作用して、形成されると考えられている。そのため、人間の持つ感性の一つとなる味覚には個人差があるとされる。その一方で、多くの一般生活者にとっては、飲食物は、個人または家族や友人、職場の同僚などの身近な存在と味わう機会が主であり、その飲食物から感じられた具体的、詳細なおいしさを他者に積極的に伝える必要性がなく、おいしいやうまい、まずいといった一般的な表現を用いて表現することが多い。そのため、飲食物から感じたおいしさを具体的、詳細に表現する機会は少なく、実際に、飲食物の味を表現する場合、その味をおいしいやまずいといった一般的な表現以外を用いて表現することが難しいといった事態が生じる[Yamamoto 2005]。例えば、日本酒は、その銘柄を選ぶ場合、多くは甘口や辛口、味の濃淡などを示す指標や、その味を説明したものを参考にすることが多い。一般的に、日本酒は糖分が多く含まれたものが甘口、糖分が少ないものが辛口といわれる。しかし、甘口は甘みを感じるものとして容易に他人と共有できる一方で、辛口は辛いと感じる成分が含まれているわけではなく、その味を何の味であるか表現しづらいものと

いえる。さらには、説明された味の情報として多く用いられる、キレやフルーティ、コク、まろやかなどといった表現は、明確な定義がないものである。また、たとえおいしいやまずいといった一般的な表現以外を用いて、おいしさを表現することがあったとしても、同一の飲食物の味を表現する際には、異なる味覚に関する表現が用いられるといった事態が生じることがある。例えば、味覚に関して「甘み」や「辛み」「コク」などといった表現を用いる。この中で、「甘み」が砂糖の味であり「辛み」が唐辛子の味であるように容易に他人と共有できる言葉である一方、その逆に「コク」のように何の味であるかが定義しづらい言葉もある。これはコクに対して明確な定義がないためである。しかし、コクに類似する味覚表現として、「深みのある味」や「濃厚な味」、「複雑な味」などを想起することができる。また、コクのある食品として連想されるものとしては、煮込み料理、ラーメン、カレーやビールなど広範囲にわたり、これらは、「深みのある味」や「濃厚な味」、「複雑な味」としても連想が可能といえる。コクのニュアンスや感覚などは、曖昧であるにもかかわらず、おおよそ我々の多くに共有されているものといえる。

こうした背景から、飲食業界などでは自社商品の味覚的な特徴を、多くの消費者へ共有させる必要があるが、一般生活者各人が各々の味覚によって判断して用いられる味覚表現は、曖昧性を含んでいるために、客観的な評価を困難にするという問題が生じる。そこで、本研究では、各人の味覚によって判断された味覚に関する表現が記載された、感性の情報源といえる口コミ情報を用いて、各人の味覚によって判断され用いられる、明確に定義ができない味覚表現の共通的な意図理解を目的とした、定量的評価手法の提案を行う。

本研究の構成は以下のとおりである。第2章では、まず、本研究の分析手法となる自然言語処理の流れを整理し、本研究で用いたBERTおよびLIMEに関する概念や手法を説明する。次に、本研究での分析対象となる口コミ情報が投稿される料理レシピ共有サイトについて説明する。そして、最後に本研究に関連の深い、生化学分野におけるコク、日本酒に関する味覚、認知言語学・身体化認知、食に関わることば、料理レシピを対象にした自然言語処理、に関するそれぞれの先行研究を概観し、課題を整理する。第3章では、日本酒を取り上げ、ECサイト上の日本酒に関するユーザの口コミ情報を用いて、一般生活者が日本酒を飲んだ際に用いられる明確に定義ができない味覚表現の共通的な意図理解を目的とした定量的評価手法の提案を行う。日本酒ラ

ベル上に表示される「日本酒度」「酸度」などの指標を目的変数，口コミ情報で確認された各種味覚表現に関する頻度を用いた指標を説明変数とし，重回帰分析により評価し，日本酒に関する各種指標と味覚表現における関係性について確認する。第4章では，明確に定義ができない味覚表現に対する，共通的な意図理解を目的とした定量的評価手法を提案する。まず，BERTによる予測結果をLIMEに適用する実験により，明確に定義ができない味覚表現における特徴語の解明を行う。次に，特徴語として確認された食材に対して，共起する味覚やテクスチャー用語をアソシエーション分析により抽出，集計した結果をもとに，明確に定義ができない味覚表現が持つ，味覚やテクスチャー特性の解明に向けた実験を行う。そして，各結果について考察し，一連の実験結果から，本研究で提案した手法の有効性を確認する。第5章では，明確に定義ができない味覚表現が，他者情報からの影響を受けて，一般生活者の間で合意形成されていくことの傾向有無を定量的に評価する手法を提案する。コクなどは明確に定義ができない味覚表現であるにもかかわらず，我々の多くは，コクという表現を多く目や耳にし，広範囲にわたる飲食物のおいしさを表す際に用いられる傾向にあることから，コクなどの明確に定義ができない味覚表現は，他者が用いた表現の影響を受けている可能性が考えられる。そこで，料理レシピ共有サイトに投稿されたレシピ情報上における明確に定義ができない味覚表現の有無と，投稿されたレシピに対する口コミ情報上における各種味覚表現との関係性をロジスティック回帰分析により評価し，明確に定義ができない味覚表現が，他者情報からの影響を受けて，一般生活者の間で合意形成されていくことの傾向有無を確認する。第6章は結論として，本研究を総括する。

第2章 テキストマイニングを用いた味覚表現 の解明における課題

2.1 はじめに

本章では、最初に、本研究の分析手法となる自然言語処理の流れを整理し、本研究で用いた BERT および LIME に関する概念や手法を説明する。次に、本研究での分析対象となる口コミ情報が投稿される料理レシピ共有サイトについて説明する。そして、最後に本研究に関連の深い、生化学分野におけるコク、日本酒に関する味覚、認知言語学・身体化認知、食に関わることば、料理レシピを対象にした自然言語処理、に関するそれぞれの先行研究を概観し、課題を整理し、本研究の意義および目的について述べる。

2.2 味覚表現の分類タスクに関する自然言語処理

本研究では、情報分類や情報抽出などの自然言語処理タスクでブレークスルーを起こしたといわれる、ニューラルネットワークを利用した言語モデルである BERT を用いた口コミ情報の分類を行っている。そして、深層学習をはじめとする高い識別性能をもつ機械学習モデルがどのように判断してラベリングを行なったのかを人間でも解釈可能な形で提示をする手法である LIME を用いて、分類に効いた特徴語の可視化を行っている。本節では、まず、この BERT の位置づけを確認するため、自然言語処理の流れを整理する。そして、本研究で用いた BERT および LIME に関する概念や手法を説明する。

2.2.1 自然言語処理の流れ

自然言語処理とは我々が日常的に用いる自然言語を，機械上で扱う技術全般を指す。自然言語処理は，大きく「機械に言語を理解させる」「機械上で言語を処理する」の二つの立場に分けることができるが，自然言語処理の研究が始まった当初から，自然言語処理は，前者の立場に立って研究されてきたといわれる [Okumura 2010] , [Kurohashi 2016]。本項でも，前者に関する技術を中心に扱う。

Bag of Words

Bag of Words(BoW)は，最もシンプルな特徴抽出手法の一つであり，古典的な方法に分類される [Harris 1954]。BoW は，文書中で出現する特徴の順序を維持しないため，場合によっては重要な情報が削除される可能性がある。たとえば，「is this a nice day」と「this is a nice day」は，文脈を考慮せずにテキストデータを解析すると，等価とみなされる。また，文書から得られるすべての特徴を含めるとモデルサイズが大きくなり，スパース性や数値的不安定性が生じる。そのため，有用な BoW モデルを構築するには，多くの場合，文書テキストを慎重に前処理する必要がある。

TF-IDF

BoW のようにモデル固有の制限がない，代替モデルはいくつか存在するが，中でも，TF-IDF [Jones 1972], [Salton 1983] は，しばしば使用されるモデルといえる。TF-IDF は，文書内の単語の重要度（重み）を示す手法の一つである。TF-IDF では，BoW のような単純な出現回数だけでなく，その単語がどの程度希少なものであったかという指標を組み合わせることで単語を定量的に表現している。また，その算出方法から，文書におけるある単語の重要度として解釈される。TF-IDF のそれぞれの指標の求め方を次式 2-1, 2-2, 2-3 に示す。

$$tf(t, d) = \frac{n_{t,d}}{\sum_{s \in d} n_{s,d}} \quad (2-1)$$

$$idf(t) = \log \frac{N}{df(t)} + 1 \quad (2-2)$$

$$\begin{aligned} tfidf(t, d) &= tf(t, d) \cdot idf(t) \\ &= \frac{n_{t,d}}{\sum_{s \in d} n_{s,d}} \left(\log \frac{N}{df(t)} + 1 \right) \end{aligned} \quad (2-3)$$

先ず TF と IDF は、それぞれ Term Frequency（出現頻度）と Inverse Document Frequency（逆文書頻度）の略であり、最終的な TF-IDF 値はこの二つの指標の積で求められる。TF の式では、ある文書 d の中のある単語 t の出現頻度を求めている。ここで $n_{t,d}$ は文書 d における単語 t の出現回数であり、これを文書 d における全ての単語の出現回数を足した値である $\sum_{s \in d} n_{s,d}$ で割っている。TF にて出現回数が多ければ多い程、その単語がその文書において重要であると判断できる。一方の IDF の式では、全文書数 N をある単語 t が出現する文書数 $df(t)$ で割ったものの対数を求めている。またこの計算だけでは、全ての文書に登場する単語の場合は 0 となるため、一律で 1 を加えるのが一般的である。IDF では、どの文書にも出現する単語は、一般的に使われる単語であり、特定の文書の持つ特徴的な意味を支える単語とは判断されないことを意味する。

TF と IDF を掛け合わせた TF-IDF 値は、ある文書内ではよく出現するものの、文書を渡っては出現しない希少な単語の方の値が大きくなり、文書クラスタリングに大きな影響を与える可能性があると判断できる。

Word2Vec

TF-IDF は現在でも標準的に使われている例もあり、一定の精度で文書理解を行うことが期待できる手法といえるが、コンピューターの性能向上に伴い、より複雑な処理を経て自然言語を理解させる方法が考案されるようになってきている。今日までにニューラルネットワークモデルを採用したアプローチは多数存在するが、ニューラルネットワークが自然言語処理に活用された成功例として、word embedding がある。word embedding とは、ニューラルネットワークを用いて、大規模なコーパスから語の意味のベクトル表現（分散表現）を学習したものである。Mikolov らによって 2013

年に発表された Word2Vec は、最も有名な word embedding のアルゴリズムである [Mikolov 2013a], [Mikolov 2013b]。Word2Vec の最大の特徴は、それまで困難とされていた単語間の関係性も表現出来る点にある。代表的な例としては、「王 - 男性 + 女性」で得られるベクトルと、「女王」のベクトルがほとんど一致する（次式 2-4）。

$$v_{king} - v_{man} + v_{woman} \cong v_{queen} \quad (2-4)$$

ここで v_{king} は「king」という単語の、 v_{man} は「man」という単語の、 v_{woman} は「woman」という単語、 v_{queen} は「queen」という単語のベクトル表現である。そのほかにも、似たような意味をもつ単語の分散表現のコサイン距離が近くなるというような関係も成り立つ。このような特徴から、単語のベクトル表現は、自然言語処理分野の様々なタスクで利用されており、Word2Vec では、主として CBOW もしくは skip-gram により実現されている（図 2-1）。

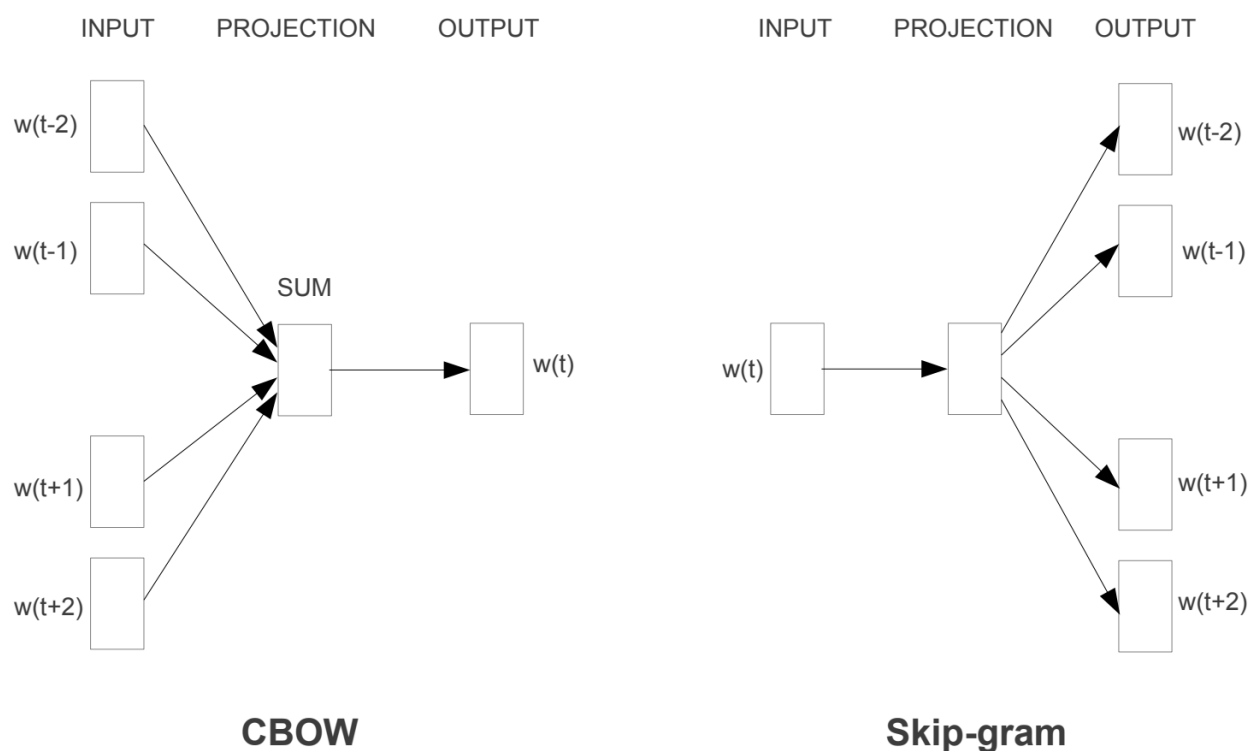


図 2-1 word2vec における CBOW と Skip-gram の概要 ([Mikolov 2013a]より抜粋)

CBOW では 2 つの入力層から中間層を経て出力がされ、この中間層は各入力層の全結合による変換後の値が平均されたものになる。単純な理解としては、一つ目の入力 w_1 に対して、二つ目の入力層にて w_2 として変換されたとすると、中間層は $\frac{1}{2}(w_1 + w_2)$ となる。入力層から中間層への変換は、全結合層によって行われ、この重みが単語の分散表現として解釈される。

CBOW モデルでは出力層にて各単語スコアが出力されるが、そのスコアに対して Softmax 関数を適応することで、総和が 1 となる 0 から 1 の間で値を得る。この値は、前後の単語を与えた時にその中央にどの単語が出現するかといった確率を表すと解釈することができる。仮に「今日 XXX を食べます」という文書があり、「今日」「を」という文脈から「カレーライス」を予想したいとする。CBOW の学習では正解ラベルとニューラルネットワークが出力した確率の交差エントロピー誤差を求め、それを損失としてその損失を少なくしていく方向に学習を進める。ここで学習用の文書(学習用コーパスと呼ばれることが一般的)を単語列 $W(w_1, w_2, \dots, w_T)$ とし、このコーパスに含まれる単語数を T 、またある時点 t における単語を w_t とする。すなわち、次式 2-5 で示すように、ある単語 w_t を前後から予測する条件付き確率関数 P とし、対数尤度関数 L を最大化するように CBOW では学習が進められる [Saito 2018]。

$$L = -\frac{1}{T} \sum_{t=1}^T \log P(w_t | w_{t+1}, w_{t-1}) \quad (2-5)$$

対する skip-gram では、直感的には CBOW と逆のタスクを行うことで学習を進める。すなわち、中央の単語から前後の複数のコンテキストを予測することになる。先の例と比較すれば、先ず「カレーライス」という単語を置き、その前後に「今日」「を」が来るタスクとして学習を進める。skip-gram の入力層はひとつで出力層はコンテキストの数だけ存在し、それぞれの出力層で個別に損失を求め、それらを足し合わせたものを最終的な損失としている (式 2-6)。

$$L = -\frac{1}{T} \sum_{t=1}^T \log P(w_{t-1} | w_t) + \log P(w_{t+1} | w_t) \quad (2-6)$$

一般的に学習データがあまり大きくない場合は **skip-gram**，大きい場合には **CBOW** の方が分散表現は良いとされている。また，類義語と対義語の区別が苦手なことや出現頻度が下がると正しいペアを作ることが難しくなることが指摘されている [Nishio 2014]。

RNN

Word2Vec は単語の分散表現を獲得するために考案する手法であり，文脈の単語の並びを意識できないため，言語モデルとして活用することはできない。こうした課題に対応するために，再帰型ニューラルネットワーク (**RNN:recurrent neural network**) [Elman 1990] が用いられる。

RNN は，音声，動画，文章などの時系列データを扱うことのできるニューラルネットワークである。時系列データとは，要素一つ一つに順序関係があり，その並びに意味が隠れているようなデータを指す。例えば，時系列データには，株価のデータを挙げることができる。時間軸に従ってデータが並んでおり，ある一つのデータは，その 1 時点前のデータから影響を受け，また，その後のデータに影響を与えており，並びに意味があると考えられることができるためである。また，テキストデータも時系列データとして挙げることができる。テキストには，並び替えると意味の破綻する順序関係が存在しており，その並びによって意味を構成すると考えることができるためである。時系列データを扱う推定問題として，時点 t までのデータを元に，時点 $t+1$ のデータを予測するという問題を考えることができる。例えば，株価の例であれば今日までの株価を元に，明日の株価を予測する問題がこれに相当する。テキストデータの例であれば，文章が途中まで与えられており，その次の単語を予測する問題がこれに相当する。従来のニューラルネットワークは入出力で扱えるデータの大きさが決まっている。また，複数の連続するデータの関係性を考慮することができず，時系列データを扱うことができない。一方，**RNN** は入力値，出力値に可変長のデータを扱うことができる。**RNN** は中間層の情報を保持して，次の入力時に保存した中間層の情報を利用できる仕組みを持つ。文章を単語に分けた場合，時系列に順々に表れた複数の単語から次に出てくる単語の推測が可能となる。**RNN** は中間層を保存することで，次の入力時に保存した中間層の情報を統合して出力結果を出すことができる。

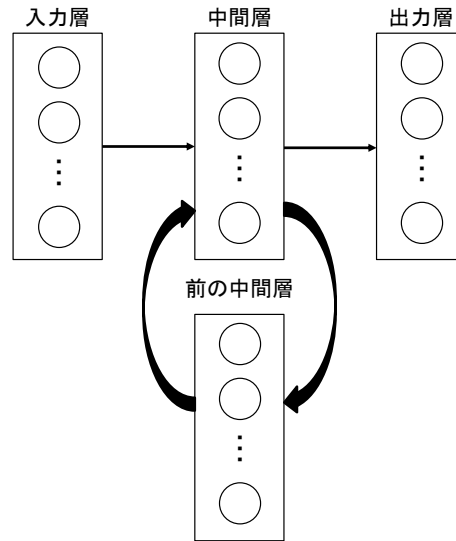


図 2-2 RNN のイメージ図

RNN は中間層の部分が従来のニューラルネットワークとは違い、前回の中間層の状態を保持する仕組みを持つ（図 2-7）。保持した中間層の情報は次回の入力層からの状態に統合することができる。RNN の入力値を次式 2-7 に示す。

$$s_t = f(U_{x_t} + W_{s_{t-1}}) \quad (2-7)$$

s_t : 中間層の値

f : 非線形関数

U, W : 重み

x_t : t ステップ時の入力値

時刻 t における入力値と時刻 t までの中間層の値を統合している。RNN では時刻 t に入力されたデータは時刻 $t-1$ のデータと統合し、中間層を経由した後出力される。また時刻 t の中間層の状態も保存され、次回の入力と統合する。以上のように、前回の中間層の情報と入力層からの情報を合わせることで、情報に前後関係を持たせることができる。

例えば、「明日のランチ」の次にくる単語は「が」「を」「は」が候補として挙げられる。従来のニューラルネットワークでは直前の単語の「ランチ」に続く単語しか予測

できず、「明日の」は考えることはできない。しかし、RNN の場合は「明日の」まで再帰して考えることができる。

ニューラルネットワークは重みとバイアスを修正する際に、誤差情報を出力層から入力層に向かって伝播する。誤差情報は最初の層まで遡るが、誤差が遡る層が多いと不必要な誤差情報が増える。誤差伝播の際は誤差情報に層ごとの重みをかけ伝播する。かける重みの値が 0 に近いと誤差が急速に小さくなり消失，あるいは急速に大きくなって発散し，学習ができなくなる勾配消失問題が生じる。理論上，RNN は長期的なデータも利用できるが，実際は 10 ステップ以上遡ることが難しいとされる。

LSTM

勾配消失問題の解決策として Long Short Term Memory(LSTM) が提案されている [Hochreiter 1995], [Hochreiter 1997]。LSTM は情報の削除やセル状態に追加する能力を持ち，ゲートと呼ばれる機能を中間層に持つ。LSTM の中間層の構造を図 2-3 に示す。

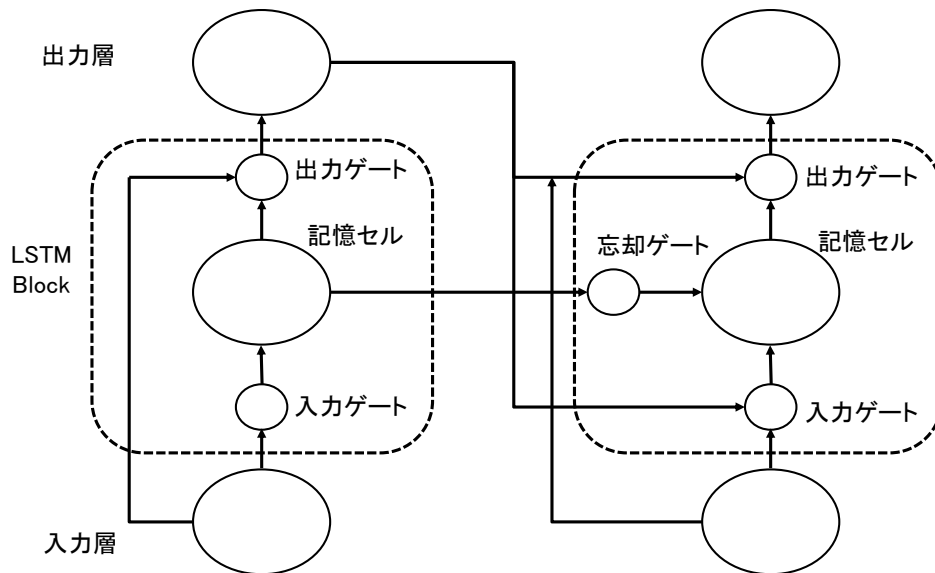


図 2-3 LSTM の構造

LSTM は RNN の中間層が LSTM Block と呼ばれるメモリに置き換えられたものである。LSTM Block は入力，出力，忘却ゲート，記憶セルで構成される。記憶セルは

データ全般を記憶する場所で、ゲートと組み合わせることで不要な誤差を除き、必要な誤差だけを通すことができる。ゲートは選択的に情報を取り出せる手段である。入力ゲートは前回の出力層の重みにより誤差を受け取るか判断し、記憶セルに通す情報を制御する。一方、出力ゲートは記憶セル内の誤差を次の層に必要な情報だけが通るように制御する。入力、出力ゲートでは必要な誤差だけを通すことで、誤差が不安定になるのを防止する。忘却ゲートでは中間層に記憶する必要がなくなった情報を除去することができる。忘却ゲートがあることで不要な情報が増えるのを防ぐことができる。LSTMでは記憶セルとゲートにより誤差を制御することにより1000ステップ以上の長期的な学習が可能である。

LSTMは、RNN機能が追加された拡張版という位置付けのため、RNNに含まれる。RNNやLSTMを複数組み合わせることで、入力文書と同義の別の文書を出力することも可能となる。入力で使われた言語を別言語で出力するような翻訳タスクが代表的な用途である。

Seq2seq

Sequence-to-sequence framework(Seq2seq) [Sutskever 2014] は、encoder-decoder モデルと呼ばれるニューラルネットワークの一種で、入力された時系列データからそれに対応する時系列データを出力するモデルである。入力系列をエンコード（圧縮）するための Encoder とエンコードされた系列をデコード（解釈）する Decoder から構成される（図 2-4）。

Seq2seq は RNN を 2 つ利用する。Encoder と呼ばれる RNN に入力時系列データを先頭から一つ一つ入力し、最後の隠れ層 h_N を求める。この h_N を Decoder と呼ばれるもう一つの RNN の隠れ層の初期値とし、ある時系列の単語から次の単語を推論することを繰り返すことで時系列データを出力する。例えば、「とても」「美味しい」「日本酒」という時系列データを入力した後の隠れ層 h_3 を利用して、「EOS」という単語列の区切りを表す特殊トークンから「いいね」を、「いいね」から「居酒屋」を、と単語を推論していくことで、「とても美味しい日本酒」という文に対応する「いいね居酒屋行く？」という時系列データを出力する。

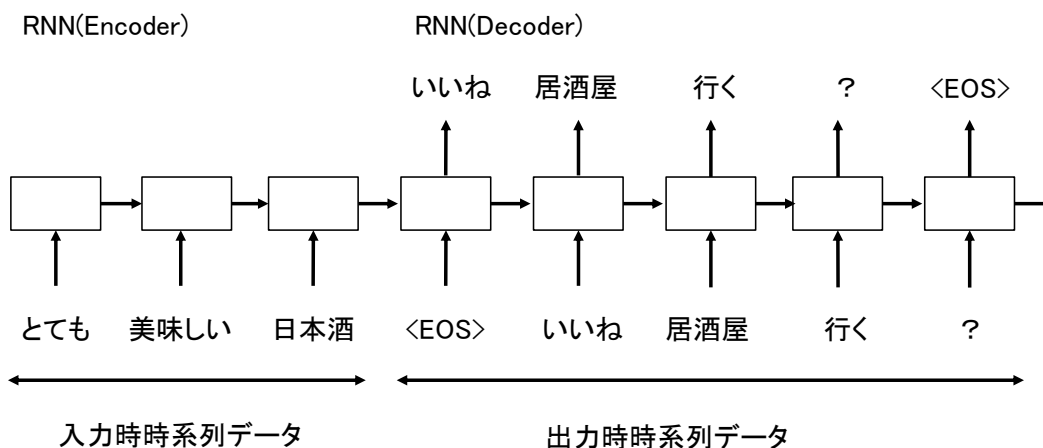


図 2-4 Seq2seq の概要

学習は以下の式で表される入力時系列データ $X = \{x_1, x_2, \dots, x_N\}$ に対する出力時系列データ $Y = \{y_1, y_2, \dots, y_{N'}\}$ の尤度を最大化する (式 2-8)。

$$p(y_1, y_2, \dots, y_{N'} | x_1, x_2, \dots, x_N) = \prod_{t=1}^{T'} p(y_t | x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_{t-1}) \quad (2-8)$$

なお、 $p(y_t | x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_{t-1})$ は softmax 関数で求めた Seq2seq が扱う語彙の確率分布である。推論時には、Encoder を用いて計算した隠れ層 h_{N1} を利用し、

「EOS」を Decoder の先頭に入れることで、出力時系列データ先頭の語彙の確率分布を推論する。Decoder に推論した単語を入力し続け、上式を最大化する単語列を推論結果とする。なお、探索はすべての語彙に対し行くと時間が膨大にかかるため、ビームサーチで行う。

Seq2Seq は、RNN を利用しているため時系列データに強く、入力と出力が時系列データであるようなタスクに用いられる手法となる。例えば、機械翻訳や音声認識の他、チャットボットのように対話を行うアプリケーションや、コンパイラのようにソースコードを機械語に変換するようなタスクなどにも用いることが可能となる。

Attention

Seq2seq は異なる時系列データの変換という点で大きな成果を挙げたといえるが、①入力する文章の長さにかかわらず、固定長ベクトルに圧縮していることで必要な情報までが捨象されてしまうこと、そして②単語や文章同士の照応関係を利用することができない、といった問題点がある。これら Seq2seq の問題を解決したのが、Attention 付き Seq2seq となる。Attention[Bahdanau 2014], [Luong 2015]はそれまで Encoder 部分から作られる固定長ベクトルが最後の部分しか利用されていなかった点に着目し、各単語が入力される際に出力される固定長ベクトルをすべて利用することで①単語の数と同じ数だけの固定長ベクトルを獲得することができ、②そのことで各単語間の照応関係を獲得することを可能にしている (図 2-5)。

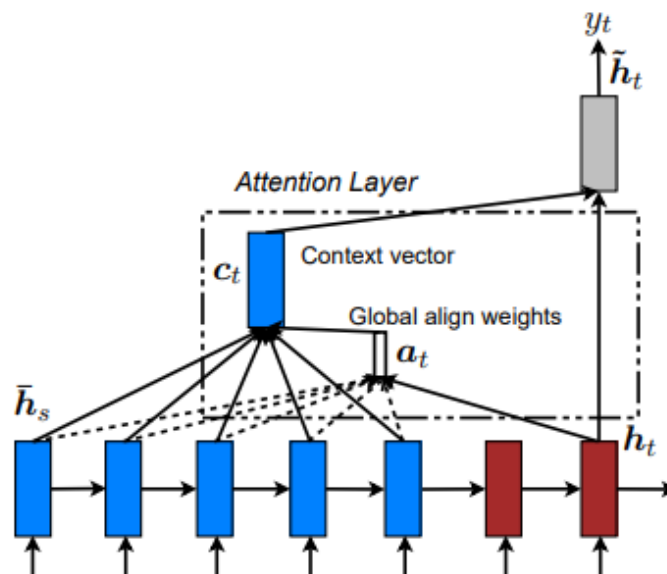


図 2-5 Attention の概要 ([Luong 2015]より抜粋)

Attention は単語間の照応関係を一対一対応で考えるのではなく Softmax 関数などを用いてこれらを確率的なものであると考える。例えば、機械翻訳タスクにおいて、dinner は「夕食 : 0.6」「食事 : 0.2」「食べ物 : 0.01」という形で照応するとされる。これらの数値は単語間の関係性を示しているため、深層学習の問題とされていたブラックボックス化をさげ、どのような単語と紐づいているかを可視化することができるという利点がある。

Transformer

Transformer は、「Attention is all you need.」 [Vaswani 2017] という論文タイトルにて 2017 年の国際会議で発表され、さらに翻訳性能が向上した。このモデルが現在、ニューラル機械翻訳で主流となっている。

Attention を利用することでより精度が高まった一方で、時系列データを逐次的に処理するため、依然として RNN などを併用することで生じている問題である、データを並列処理ができず、計算の高速化が困難という問題があった。こうした問題を解決したのが、RNN などを利用せず、Attention 層だけを用いて構築するモデルの Transformer である。

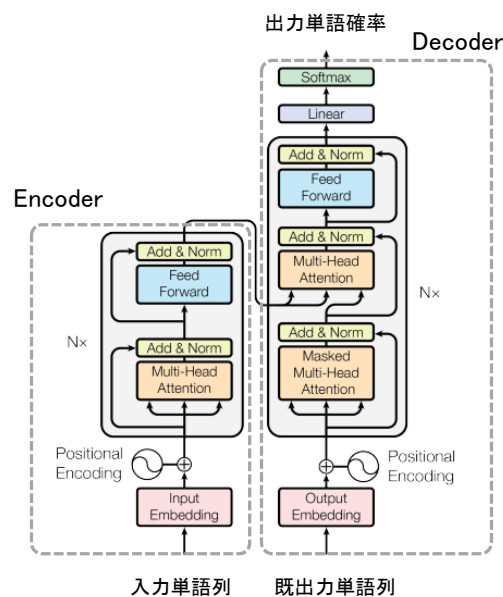


図 2-6 Transformer の概要 ([Vaswani 2017] を一部編集)

それまでの RNN を用いた系列変換モデルでは、系列データ中で離れた位置の単語の情報を組み合わせるには再帰構造を通して系列を伝搬させていた。その際、単語間の距離が離れていると伝搬中に情報が減衰してしまい、距離が離れた文脈情報を扱にくく、さらに順次再帰の値を計算するため並列化が難しかった。Transformer モデルではこれらの問題を解決し、距離が遠い単語の情報の扱いに優れており、処理を並列化して計算を高速化できるという特徴を備えている。

主な用途である機械翻訳タスクでは、翻訳対象とする 2 つの言語の正解対訳文を入力とし、その入力文に対する出力として学習することで、学習したデータと同じような入力文に対して、学習した訳文と同様な翻訳ができるようになる。例えば日英翻訳の場合、「これはバナナです。: This is a banana.」, 「私はスプーンを持っている。: I have a spoon.」などの対訳文を数十万から数百万文対程度学習することで、機械翻訳のモデルを構築する。

図 2-6 は Transformer のモデル構造である。図 2-6 左半分は、翻訳元言語の単語列を入力して、分散表現の系列を出力する Encoder である。図 2-6 右半分は、この分散表現の系列をもとに翻訳先言語の単語列を生成する Decoder になっている。Decoder は、先に求めた分散表現の系列とこれまでに訳出された単語列から次の翻訳先言語の単語を出力する構成をとっており、最初は文頭を意味する特別な単語を入力して最初の単語を出力し、その後は出力した単語を入力する度に一単語ずつ逐次単語を出力していく。

2.2.2 BERT の概念

Transformer に基づくニューラルネットワークの Encoder・Decoder モデルの Encoder 部分を利用し、MLM (Masked Language Model) と呼ばれるタスクで学習するモデルとなる BERT (Bidirectional Encoder Representations from Transformers) が、2018 年に Google から発表[Devlin2019] された。BERT は、自然言語処理のさまざまなタスクに適用可能なニューラル言語モデルであり、発表時のさまざまなタスクにおいて最先端の手法を大きく上回る精度を出している。

MLM は、文中のランダムにマスクされた部分の単語を予測するタスクであり、図 2-7 (左) の例では、マスクされた「強風」と「遅延」を周辺の単語から推測する問題を学習することになる。MLM タスクでは、通常のテキストから教師データを自動作成し、学習することができる。このような仕組みは自己教師あり学習 (Self-Supervised Learning) と呼ばれ、大規模なテキストさえ用意すれば、その規模に応じた教師データで学習することができるため、Transformer の利用とともに BERT の性能が高くなる要因の一つとなっている。

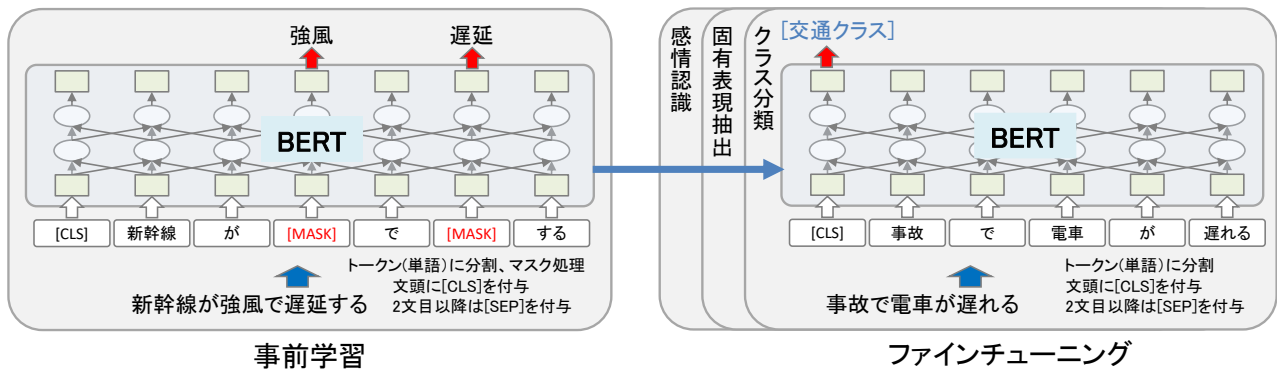


図 2-7 BERT における事前学習とファインチューニング

MLM などのタスクであらかじめ学習された言語モデル（事前学習済み言語モデル）は、ファインチューニングと呼ばれる処理で、自然言語処理のさまざまなタスクに最適化することができる。図 2-7（右）の例では、事前学習したモデルをもとに「政治」「経済」「天気」「交通」などのあらかじめ定めたクラスのいずれかに振り分けるクラス分類タスクでファインチューニングを実施している。そのほか、入力された文章の内容を表す情報（ラベル）を複数種類割り当てるタスク（マルチラベル分類）や、テキストから人名や組織名などの表現を特定するタスク（固有表現抽出）などの情報分類や情報抽出のさまざまなタスクで事前学習済み言語モデルをファインチューニングすることで、それぞれのタスクに対応した高精度なモデルを作成することができる。

BERT が発表されて以降、事前学習済み言語モデルに関する研究は急増しており、より大規模なコーパスを用いて事前学習を実施した RoBERTa[Liu 2019]、学習時の消費メモリを減らし、学習時間を削減可能な ALBERT[Lan 2020]、知識蒸留のメカニズムを用いて性能を維持しつつモデルのサイズを削減し動作速度を向上させた DistilBERT[Sanh 2019] など、BERT を改善したより高精度な事前学習済み言語モデルが提案されている。日本語版の BERT についても、東北大学[Tohoku University 2020]、京都大学[Shibata 2019]、情報通信研究機構（NICT）[NICT 2020]が Wikipedia や Web などのコーパスを用いて事前学習したモデルを公開している。このような事前学習済み言語モデルを用いることにより、情報分類や情報抽出タスク、さらには情報検索タスクにおいても著しく性能が向上している。

2.2.3 LIME の概念

これまでみてきたように、近年、深層学習をはじめとする高い識別性能をもつ機械学習モデルが様々な分野に応用されている。一方で、高性能な機械学習モデルは大量の教師データから学習した複雑なルールを用いて判定するため、それらのモデルの多くはブラックボックスであり、ユーザがモデルの挙動や予測の原因について知ることが困難になっている。そこで、こうした問題を解決するために、モデルや予測結果への解釈性に関する研究が活発に行われている。こうした説明可能な機械学習アルゴリズムを用いたシステムは、一般的に「説明可能な AI(XAI: Explainable AI)」と呼ばれ、機械学習を活用するあらゆる分野で関心が高まっている。

機械学習モデルがどのように判断してラベリングを行なったのかを人間でも解釈可能な形で提示をする手法として、LIME (Local Interpretable Model-Agnostic Explanations) [Ribeiro 2016] がある。LIME は、任意の機械学習モデルに対する入力を説明変数、出力を目的変数として解釈可能なモデルである線形回帰モデルに学習させ、各特徴の重要度を求めることによって解釈を行う手法である。複雑な推論アルゴリズムに対応するために、特徴空間内の予測対象のベクトルの近傍を学習することによって、局所的に正しい説明を生成することができる。LIME は推論結果の説明として、出力に対する各入力特徴の重要度を数値として捉えることができ、重要度の高い特徴を提示することが可能となる。LIME のイメージ図を図 2-8 に示す。

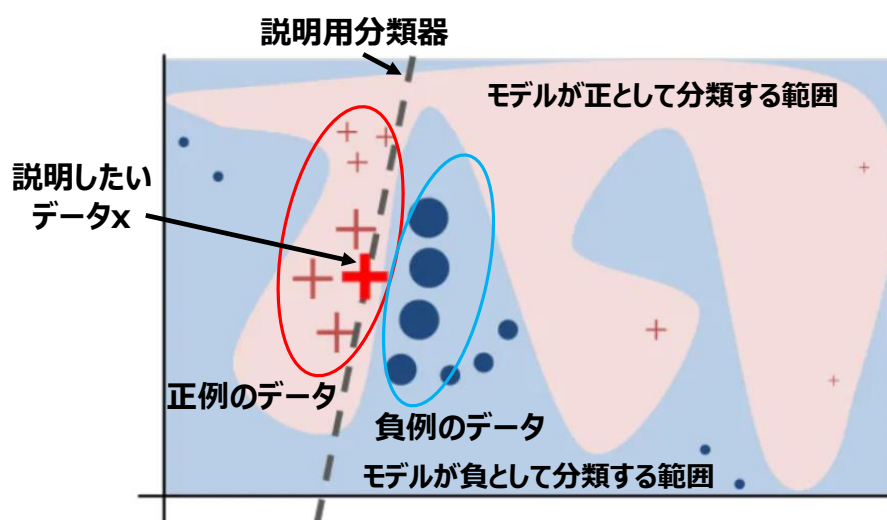


図 2-8 LIME のイメージ図 ([Ribeiro 2016] を一部編集)

LIME のアルゴリズムを以下に示す。

LIMEによる説明生成のアルゴリズム	
入力:	説明対象のモデル f , 入力ベクトル x
入力:	サンプル数 N , 類似度カーネル関数 π_x
入力:	説明に用いる特徴料の数 K
出力:	各特徴量の重要度 w
1.	$Z \leftarrow \{\}$
2.	for $i \in \{1, 2, 3, \dots, N\}$ do
3.	$z_i \leftarrow \text{sample_around}(x)$
4.	$Z \leftarrow Z \cup \langle z_i, f(z_i), \pi_x(z_i) \rangle$
5.	end for
6.	$w \leftarrow K - \text{Lasso}(Z, K)$
7.	return w

入力には、説明対象のモデル $f \in \mathbb{R}^d \rightarrow \mathbb{R}$ と、入力ベクトル $x \in \mathbb{R}^d$ を与える。また、パラメータとして、サンプリングするベクトルの数 N と説明に用いる特徴量の数 K 、そして、 x との類似度を算出する類似度カーネル関数 π_x を与える。 π_x は、局所性をコントロールする関数であり、次式 2-9 で表される。 σ はカーネル幅、 D はコサイン距離やユークリッド距離といった距離関数を表す。

$$\pi_x(z) = \exp\left(-\frac{D(x, z)^2}{\sigma^2}\right) \quad (2-9)$$

まず、推論結果の説明を行う入力ベクトル x を中心として、ベクトルをサンプリングする(3行目)。サンプリングした入力ベクトル z_i と説明対象のモデルによる出力 $f(z_i)$ 、そして入力ベクトル x と z_i の類似度 $\pi_x(z_i)$ を集合 Z に追加する(4行目)。3, 4行目の操作を N 回分行う。そして、 z_i を説明変数、 $f(z_i)$ を目的変数として、 K 個の特徴量のみを用いた複数の線形回帰モデルを学習し、説明に最適なモデルの選定を行う(6行目)。最適な線形回帰モデルを $\xi(x)$ 、候補の線形回帰モデル $g \in G$ とすると、 $\xi(x)$ の選定は次式 2-10 に基づいて行われる。 $\xi(x)$ には、 x の局所性 π_x を考慮した、 f の g による近似の損失関数 L を用いる。

$$\xi(x) = wx = \underset{g \in G}{\operatorname{argmin}} L(f, g, \pi_x) + \Omega(g) \quad (2-10)$$

L の計算式を次式 2-11 に示す。また、関数 Ω は、モデル g の説明の複雑度を表し、 g の重みのベクトルのうち非ゼロの要素数を返す関数が一般的に用いられる。

$$L(f, g, \pi_x) = \sum_{z \in Z} \pi_x(z) (f(z) - g(z))^2 \quad (2-11)$$

上述のアルゴリズムによって出力された、重み w を用いて説明のために可視化した一例を図 2-9に示す。

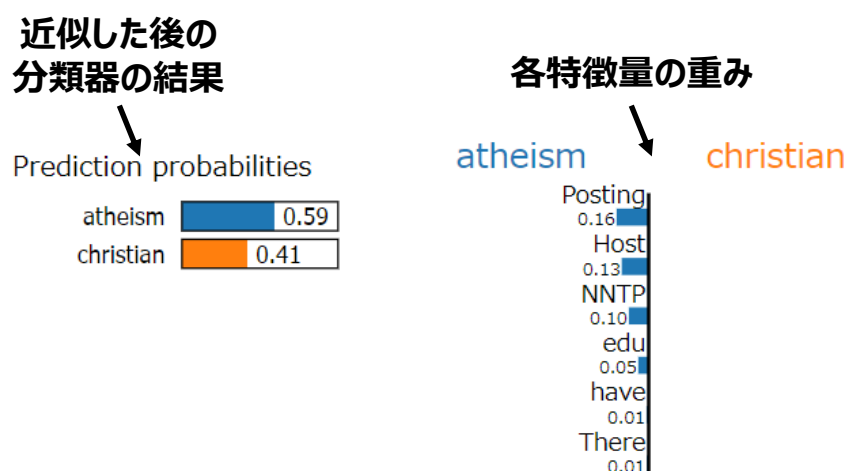


図 2-9 LIME 出カイメージ図 ([Ribeiro 2016]を一部編集)

図 5 は、無神論者(atheism)とキリスト教信者(christian)の投稿をランダムフォレストにより分類し LIME に適用したものの出力結果の一つである。本投稿では、近似後の分類器の結果では、無神論者の投稿である確率が 0.59 と予測されている。また、各特徴量の重みとは、近似後の分類器の偏回帰係数であり、予測結果にどの特徴量がどの程度寄与しているのかを表す。Posting が 0.16、Host が 0.13、NNTP が 0.10 などとなっており、近似後の分類器の予測結果に大きな影響を与えた特徴量を表していると解釈できる。

2.3 世界中で利用される料理レシピ共有サイト

本節では、本研究での分析対象となる口コミ情報が投稿される料理レシピ共有サイ

トについて説明する。

料理レシピサイトは世界中で利用されている人気のコンテンツである。Web サイト解析ツール SimilarWeb のカテゴリー「FoodandDrink>CookingandRecipes」における料理レシピサイトの世界ランキングを表 2-1 に示す。

表 2-1 料理レシピサイトの世界ランキング

Rank	Website address	Average visit duration [mm:ss]	Pages/Visitors	Country
1	cookpad.com	02:03	2.78	Japan
2	allrecipes.com	01:49	2.40	USA
3	giallozafferano.it	02:58	1.99	Italy
4	foodnetwork.com	01:44	1.89	USA
5	delishkitchen.tv	01:05	1.47	Japan
6	nefisyemektarifleri.com	01:54	2.39	Turkey
7	journaldesfemmes.fr	08:21	2.40	France
8	chefkoch.de	03:00	3.33	Germany
9	kurashiru.com	01:09	1.54	Japan
10	tudogostoso.com.br	00:53	1.41	Brazil

2022 年 5 月 9 日現在

本ランキングは、トラフィック数によるものである。これによると、2022 年 4 月のランキング 1 位に位置するのは日本のクックパッドであり、上位 10 件には日本、アメリカ、ドイツ、イタリア、フランス、ブラジルと、6 か国のサイトがランクインしている。それぞれの国が自国の食文化を持ち、自国の言語で書かれたレシピを検索し、利用している様子がうかがえる。ランキング 1 位のクックパッドは、1998 年に開設された一般ユーザが作成したレシピを Web 上に掲載するユーザ投稿型のレシピが大半を占めるレシピ共有サイトで、2022 年 4 月現在、クックパッドにはおよそ 360 万件的レシピが掲載されている。一方で、ランキング 2 位の Allrecipes は、1997 年に米国で開設された、クックパッドと同じユーザ投稿型のレシピ共有サイトであるが、WEB 上で検索可能なレシピは、2022 年 4 月現在、およそ 5 万 5 千件ほどであり、レシピ投稿数に大きな差がある。表 2-1 で示したランキング外ではあるが、同じくユーザ投稿型のレシピ共有サイトで、日本の楽天レシピで掲載されているレシピ数も、2022 年 4 月現在およそ 249 万件であり、他国と比べて、日本は Web レシピ資源が極めて豊富な国といわれている[Yamakata 2019]。そうしたレシピの多様性が多くのユーザのニーズを満たす一方で、そのレシピ数の多さから、より詳細な条件による検索

を可能とすることや、レシピ全体としての集合知を取り出す仕組みを構築することが、重要になっている。

2.4 テキストマイニングと料理・味覚表現に関する先行研究

本節では、本研究テーマと関連の深い、生化学分野におけるコク、日本酒に関する味覚、認知言語学・身体化認知、食に関わることば、料理レシピを対象にした自然言語処理、に関するそれぞれの先行研究を概観し、今回取り組む研究の意義について整理する。

2.4.1 生化学分野におけるコクに関する研究

コクの定義を試みる代表的な研究として「食品のコクは、味、香り、食感による多くの刺激（複雑さ（深み））のバランスで形成されるものであり、それらの刺激に広がりや持続性が感じられる味わいである」とした定義を提案したものがある[Nishimura 2018a], [Nishimura 2021b]。コクは、食べ物に含まれる物質やその構造など、食べ物の素材から感じる刺激によってもたらされることから、食べ物のおいしさを決定する要因の一つであり、客観的評価が可能であるとし、コクの要素を「複雑さ」「広がり」「持続性」の3つとしている。例として、熟成した食肉は、調理すると複雑さ、広がりや持続性が強くなり、これは、熟成により肉の呈味成分や香気成分が増えることで、さらに複雑さ、広がり、持続性が強くなるためであるとしている[Nishimura 2019]。熟成処理中には、うま味物質も増加するため、コクの要素である味わいの広がりや持続性も増強されるということである。また、食肉製品であるソーセージは、うま味物質や脂肪が添加されているため、香りの強さ、風味の複雑さ、味の持続性が、有意に強くなることが確認されている[Nishimura 2018b]。その他、増粘剤で「とろみ」をつけた溶液を用いて、コクの要素である広がりや持続性を調べた結果、「とろみ」は、香気成分の放出を抑制し、「香り（口中香）の持続性」を強めることが明らかとなったという報告[Nishimura 2021a]もある。さらには、食品にコクを付与する物質を「コク味」(kokumi)物質と呼び、kokumi物質の受容機構を明らかにする目的で実施した官能特性評価による研究などがみられる[Kuroda 2018] [Kuroda 2021]。

ビールの特徴を表す「コク」と「キレ」という表現は、古くから広告などで盛んに使われてきたこともあり、消費者にもなじみのある表現となっている一方、その意味については、曖昧であり、個人の捉え方も異なっていると考えられる。また、酒税法の改正によって、ビール以外にも発泡酒、新ジャンルおよびビール風味のノンアルコール飲料といった区分のビール様飲料が浸透してきているが、ビールとそれ以外のビール様飲料では、コク、キレを含む香味の差が大きいといわれ、香味と香味表現の研究は絶えず行われている。ビールのコクは、官能評価に関する国際基準に従い「ビールを飲んだ時の風味（味、香り）や食感の豊かさ（body）」という定義[ASBC 1992]に近いと考えられており、日本では、ビール酒造組合・国際技術委員会 BCOJ が 2002 年に策定したフレーバーホイール[Furusho 2002]の 14 の中分類の中に「ボディ」

（香味の厚み、インパクト）という項目がある。また、ビールのコクに強く関連する味の強さ、厚みは、まず、エキス分の多少によって判断できるとされ、エキス分（揮発性固形分）は糖（オリゴ糖～多糖）、タンパク質および分解物、ポリフェノール、メラノイジンなどが含まれており、これらの量はビールの原材料、製法（条件）によって大きく異なってくるとされる[Yoshida 1976]。その他、ビールのコクについては、風味(味、香り)や食感成分が関与し、ビール中のアルコール、残存炭水化物、苦味質およびポリフェノール類などの成分が、総合的にビールのコクを形成している[Tanimura 2002] とされるものがある。さらには、上記の定義に従い、ビールを飲んでいる最中にビール成分の口腔内細胞への相互作用（刺激）が大きく長時間続くほどコクがあると評価すると仮定し、センサーを用いて官能評価によりコクとの相関関係をみた研究がある[Kaneda 2014]。

ワインにおいては、ボディ感がコクに近いものと考えられるとされている[Goto 2020]。ボディ感は、ワインの味わいの成分が豊かで、バランスが取れており、厚みとして感じられる場合に用いられる言葉ということであるが、ワインの品質に大きく寄与する性質でありながら、ボディ感の指標は評価者によって微妙に異なることもあり、その実態は明らかになっていないということである。例えば、赤ワインの裏ラベルに表示されるライトボディ・ミディアムボディ・フルボディとは、この場合のボディ感はタンニン（縮合タンニン、プロアントシアニジン）による渋味が主体と考えられているが、渋ければボディがある、というものではなく、渋味の性質も重要と考えられているということである。ワイン業界では経験的に種子のタンニンは渋みが荒いが、

果皮のタンニンは心地良い渋みを示すと言われてきており、種子と果皮のタンニンは組成が異なり、それが両者の官能特性の違いをもたらしているという報告[Vidal 2003]がある。さらに、ワインが熟成するとタンニンが重合し、渋味が丸くなる、と言われていたが、ワインの酸性条件ではタンニンは徐々に分解して鎖長が短くなり、総量と同じでも鎖長が短いほど渋味が穏やかになるということが報告されている。白ワインについては、さわやかな酸味とブドウの品種によって様々に異なる香が大きな魅力といわれるが、このようなフレッシュさとは別に、ボディ感も特に高級な白ワインに求められる特性の一つに挙げられている。貴腐ワインのような極甘口ワインは豊かなボディ感を持つワインだが、単に甘いだけではなく、酸味やかすかな苦味とのバランス、グリセロールによるテクスチャー、さらに香味の複雑さがボディ感に寄与していると考えられている。一方、辛口の白ワインにもボディ感という概念はあり、成分的な解明はなかなかされてこなかったということであるが、[Gawel 2014]では、果汁に含まれるフェノール化合物のカftarリック酸が酸化され、同じく果汁に含まれるグルタチオンと結合した化合物が、白ワインの Oily な口当たりに寄与することを報告している。また、ブドウや酵母から抽出される多糖類の影響を示唆する報告もあるが、pH やアルコール分の影響の方が大きいとも報告されており、まだまだ不明の点が多く残されている。

コーヒーの味わいの特徴として、苦味、酸味、香りに加えて、独特のコクによるとしており、この中で、コクの持続性は、焙煎により生成される苦味物質により生じるとした報告がある[Nishimura 2022]。焙煎時の加熱温度と加熱時間や、挽いた粉の大きさが、コーヒーの味わいやコクに影響を与えるということである。コーヒーの苦味は、カフェインに加えて、焙煎処理で生じるクロロゲン酸ラクトンやビニルカテコール・オリゴマーも苦味の強さやコクに重要であるとされ、これらは、焙煎温度が高くなる深煎りの豆に多くなり、その豆から抽出した深煎りコーヒーは苦味とコクの持続性が強くなるということである。また、焙煎したコーヒー豆を細挽きすると、粉の表面積が増えるため、豆から抽出される酸味物質、苦味物質、香り物質の量が、粗挽きの場合よりも多くなり、コクの複雑さ、広がり、持続性が強くなるとしている。

緑茶については、緑茶による和食のコク増強効果に関する研究[Itoen 2019]がある。濃厚感、広がり、持続性に、うま味の強さを加えてコクを定義付けし、パネル評価や時間軸を考慮して多角的にコクの増強効果を検証している。この研究では①緑茶

と和食を交互に摂取すると和食のコクが増強する，②緑茶の渋味がコクの増強効果を抑制する，③コクの増強効果を最大限生かすには，渋味が少なく，アミノ酸豊富な緑茶は望ましい，といったことを明らかにしている。

チョコレートに関しては，嗜好性の高いチョコレートの香味表現は多種多様であり，一般的な五味（甘味・塩味・苦味・酸味・旨味）と，香り（ナッティ・フルーティ・フローラル・カカオなど）があり，特に高品質のチョコレートについては，後者の香りが重要視されているという[Sakiyama 2022]。チョコレートの味を表現するにあたり，乳製品を使用しない，カカオと砂糖から作られるダークチョコレートを例にとれば，まず砂糖由来の甘味，カカオ由来の苦味が表現として最も使用され，ついでカカオ豆由来の酸味と続くとされ，チョコレートにおいて旨味という表現が使用されることは稀であるという。一方，コクのあるチョコレートという表現は，旨味よりも使用されることが多く，これは，チョコレートの芳醇な香りや，喫食時において油脂が口中で融解し，経時的に乳化状態が変化していくことや，カカオ本来が持っている呈味成分などが複雑に影響して表現されていると考えられるという。科学的には，発酵中にカカオ豆のタンパク質が，酵素分解によってペプチド・遊離アミノ酸へ変化，およびその後のロースト工程における糖とのメイラード反応によって，苦味・香氣成分へと変化しチョコレートのコクに寄与しているとされる。

市販カレーの製品開発においては，カレーのおいしさにはコクを始めとして，様々な構成要素が複雑に関与していると考えられており，おいしさを視覚化し，味の分解，組み立てを容易にするために，風味波形図というものが用いられることがあるという[Shimizu 2019]。風味波形図とは，食品を口に含んでから感じられる風味の強さと時間変化を表したグラフである。風味波形図において，コクは，風味の強さ軸の縦の高さと，時間経過軸の横の広がりとして描かれるもので，カレーの製品設計をする際のコクの考え方について，風味波形図をもちいて，整理を行うということである。カレーのコクにつながる品質要素としては，基本五味(甘味，塩味，酸味，苦味，うま味)，辛味などが考えられており，これらの品質要素を，時間軸(トップ，ミドル，ラスト，ベース)に沿って割り当て，トップの立ち上がり，ミドルの呈味感や，ラストの伸びについて複雑に味を重ね合わせることにより，コクのあるカレーを提供することを可能にしている，とされている。

2.4.2 日本酒における味覚に関する研究

生化学分野において、日本酒の味や風味を分かりやすく表現することを目的とした複数の研究がある。日本酒の官能評価分析における試験において標準用語として使用することを目的に、清酒のフレーバーホイールを作成したものがあある[Utsunomiya 2006], [Utsunomiya 2007]。フレーバーホイールは、清酒中に個々に確認することのできる香味特性を表す 86 の用語を 16 のクラスに分類し、さらにクラスの中を 2 つに階層化して定義されており、においや味について複数の項目が存在している。その他、日本酒の味や風味を分かりやすく表現するために、味覚センサーやガスクロマトグラフ質量分析計を用いた機器分析と官能評価とを統計学的に組み合わせることで、「日本酒味わいマップ」を開発したものがあある[Kanno 2018], [Kanno 2019], [Kanno 2020]。本マップでは、日本酒の味を先味後味軸と淡醇軸で区分し、キレ、コク、すっきり、ふくよか、おだやかなの 5 つのエリアを定義している。

テキストマイニング手法による、日本酒の味表現に関する研究も複数みられる。雑誌や書籍から日本酒やワインの味わい表現を抜き出し、各表現コーパスを作成し、音象徴語の共起関係に注目した分析を行ったものがあある[Fukushima 2015], [Fukushima 2016]。味覚における音特徴語の使用原理として、味わいの状態よりも変化を表現するために用いられる傾向があることなどが確認され、暗黙的な身体感覚の言語化方略としての音特徴語を明らかにしている。また、オノマトペを使った日本酒の味の表現方法について研究したものがあある[Otsuka 2015]。個人的な味の感じ方を言語化するプロセスについてオノマトペに注目し、その表現から味の感じ方の違いについて共通性と個別性について分析している。その他、日本酒の味の表現方法について、蔵元から提供される日本酒の説明文データを用いて、どのような味についてどのような比喩的表現が用いられているのかの分析や、味表現の類似性を計算することにより、同様の味を持つものを探索することを試みた研究[Fukumoto 2016], [Fukumoto 2017], [Fukumoto 2018]があある。

2.4.3 認知言語学・身体化認知に関する研究

認知言語学的なアプローチにより、様々な米菓を表すオノマトペ表現は、実際に米

菓を食べている時はパリパリ・サクサクと表現するのに対し、思い描くときはバリバリとなる、という食べている場面と記憶の比較によって、そのズレを観察した研究[Uno 2017]がある。記憶における食文化と食体験の拮抗が、このズレの背後にあると考えられると論じている。関連したものとして、米菓の外装パッケージの「カリッとかるーい歯ざわり・香ばしさ」の「カリッ」というオノマトペを食べる前に見せると、食べた後に見せる時より有意に「カリッと」感じる回答が多いことを、実験を通じて観察した研究[Funakubo 2016]がある。

また、他者の発言が共食者の主観的な「おいしさ」および電気味覚閾値に与える影響を検討した研究[Sakauchi 2017]がある。他者の発言は共食者の「おいしさ」「食感」および「好み」といった情動評定に有意に影響を与えたが、電気味覚閾値には影響を与えないことが確認され、他者の発言は、生体よりも情動への影響が強いことが推察されている。加えて、他者の発言が「おいしさ」に影響を与える場合、被験者の共感性と他者の発言が「おいしさ」に与える影響についても検討されたが、本実験では共感的配慮尺度と主観的な「おいしさ」の間には有意な関係が認められず、共食者の発言によって、おいしく感じるか否かは、その個人の共感性とは関係がないことが推察されている。

その他、身体感覚や身体運動が、意思決定や対人認知などに影響を与えるという身体化認知に関する研究がある。五感の中でも味覚に関する検討は比較的少ないとされているが、水、甘いドリンク、もしくは不快感を伴う苦いドリンクのいずれかを飲ませた後に、非道徳なエピソードを提示し、それに対する態度を測定した研究[Eskin 2011]がある。実験結果、苦いドリンクを飲用したことによって生じた不快感が、非道徳的なエピソードに対する不快感を高めることが示されている。また、社会的判断が味覚に及ぼす影響に関する研究として、ロマンチックな(sweet)経験を書き出すことで、チョコレートがより甘く感じることを示したもの[Chan 2013]がある。

2.4.4 食に関わる言葉に関する研究

食に関わる言葉を多面的に分析し、その構造的特徴を明らかにすることを目的にするなどした、食べ物の性質やおいしさを表す言葉についての研究は複数ある。日本語の味覚に関する表現の収集として代表的なものとしては、「味ことば」と称された味を

表現する言葉を漫画や雑誌書籍から事例収集し、味ことばの分類を行った一連の研究 [Seto 2003], [Seto 2005]がある。甘いや苦いなどの味覚を直接表現する言葉は限定的であり、不足する表現の大部分はメタファーにより補われていることを指摘し、味覚に関する表現の階層的分類を、比喩と感覚領域の観点から網羅的に示している。その一方、収集された用語のうち、「コクのある」という表現については、表現としては単独であっても意味的には複合的であり、コクとは何かということが、表現の立場からも興味深いテーマであるとされている。その他、日本語における食感覚の擬音語・擬態語を収集、性別・年齢階層別におけるその認知度を分析し [Hayakawa 1999], [Hayakawa 2005], テクスチャー用語として、ISO11036 Texture Profile¹の3要素と対応した「力学的特性 (mechanical attribute)」「幾何学的特性 (geometrical attribute)」「その他の特性 (other attribute)」の3つの大分類と、中分類15、小分類64に分類が行われた研究 [Hayakawa 2013a], [Hayakawa 2013b]がある。テクスチャー用語445語のうち、いずれのテクスチャー特性にも対応すると判定されなかったため、解析から除くとされたテクスチャー用語があり、そのうちの一つに「コクがある」が含まれている。また、「もちもち」「ジューシー」「サクサク」などといったおいしさを感じさせる言葉をシズルワードと定義し、「味覚系」「食感系」「情報系」の3つの領域に分類し、アンケート調査などにより、それらの時系列変化を追うなどした分析が行われている [Ohashi 2016], [B. M. FT Kotoba Lab 2016]。コクを含む表現はシズルワードに含まれており、概念的な説明や用語の活用場面や想起される料理の整理はされているが、明確な定義を目指したものではない。

その他、[Sakurai 2000] は、嗅覚に関するおいしさが、「甘い (味覚) ~」「やわらかい (触覚) ~」「ふんわりとした (触覚)」のように他感覚関連用語 (味覚・触覚など) を用いて表現される、共感覚的比喩表現が多いことを示している。また、[Kusumi 2004] では、味覚固有のことばの数は、味覚表現の多様さに比べて少ないことを指摘しており、他の感覚に固有のことばを転用する共感覚表現 (例: やわらかな (触覚) + 甘さ (味覚)) や、比喩を用いて味覚を表現していることを示したものなどがあ

¹ <https://www.iso.org/standard/76668.html>

る。

2.4.5 料理レシピを対象にした自然言語処理に関する研究

近年、料理レシピを対象にしたさまざまな言語処理研究が増加している。料理手順に関する自然言語の理解に着目し、その要素技術である単語分割、料理レシピの要約やフローグラフの推定、料理レシピの特徴を用いたレシピ分類、ユーザ補助のための料理レシピ説明文生成、料理オントロジーの構築に関する研究などがその一例である。

日本語を対象とする自然言語処理では、まず文中の単語を同定する単語分割が課題となる。単語分割は、各文字間に単語境界の有無を判定する課題である。レシピテキストの単語分割には、「たまねぎ」、「玉ねぎ」、「たま葱」などのように全部または一部が平仮名で書かれる食材名などが多いことなどから、一般分野に比べて難しいといわれている。そこで、辞書やコーパス（例文集）の整備などにより、単語分割精度を引き上げた研究[Maeta 2015]がある。また、クックパッドにおいては、「しょうゆ」には 100 種類以上のバリエーション（「しょう油」「醤油」「おしょうゆ」など）があるということである。そこで、クックパッドでは Encoder-Decoder モデルで材料名を正規化されている [Harashima 2018]。レシピ中の材料名（例しょう油）を Encoder に入力し、正規化後の材料名（例しょうゆ）を Decoder から出力される。管理栄養士が正規化した材料名でモデルを学習し、管理栄養士が正規化していない材料名にこのモデルを適用しているということである。

料理レシピの要約については、複数テキスト要約の技術を用いて、レシピサイトに投稿されたある料理に関する複数のレシピから、典型的な材料と手順を出力する手法を提案したものがあある [Nanba 2013]。要約にあたり考慮すべき点として、表記揺れに関しては概念辞書および同義語辞書の構築、省略に関しては述語項構造解析、多対多の同定に関しては格フレーム表現を用いて対処する手法が提案されている。

フローグラフの推定としては、手続き文章をフローグラフとして表現する形式を、レシピテキストに応用してコーパスの作成が試みられている [Mori 2013]。コーパスを作成する際に、単語の出現位置情報の保持や、固有表現タグを用いているため、前者は単語分割器、後者は固有表現認識の研究に活用できることや、述語項構造解析に

も応用することを述べている。また、そのフローグラフからレシピ文の自動生成も行われている[Yamazaki 2014]。形態素解析を行った後に付与されるレシピ用のタグの並びに関するテンプレートを作成し、どのテンプレートに属するかを計算してレシピ文を自動生成している。他にも、Webでレシピ検索した結果から得られた上位数件のレシピから典型的な調理手順と個々のレシピの特徴を発見する手法も提案されている[Yamakata 2013], [Yamakata 2016]。その際、レシピを事前にフローグラフ化することを試みており、編集距離によって算出された数値を基に要約する2つのレシピを選択し、挿入、削除、置換の3つの操作を行い要約が行われている。

ユーザ投稿型レシピサイトにおける料理名では、例えば「簡単！カルボナーラ」や「子供が喜ぶオムライス」のように、「修飾表現+料理」で表現されることが多いとされる。同一の修飾表現のレシピでも、ある材料が用いられていない場合や、別の材料や調理器具に代替されている場合など、レシピのネーミングコンセプト（修飾表現の根拠を示すレシピの特徴）はさまざまである。例えば、「子供が喜ぶカルボナーラ」のように同じ料理名で同じ修飾表現を含むレシピでも、「子供が喜ぶ」の根拠はレシピごとに異なりそれぞれ異なるネーミングコンセプトをもつといえる。また、「子供が喜ぶカルボナーラ」と「子供が喜ぶハンバーグ」のように異なる料理であっても同じネーミングコンセプトをもつ場合がある。このようなネーミングコンセプトを抽出するために、料理名と修飾表現に着目してレシピを分析した研究がある[Tachibana 2013], [Wakamiya 2014]。修飾表現の根拠をネーミングコンセプトと定義し、材料と調理器具に着目して典型的要素との差異を抽出し、確信度を用いて代替関係を判定することでレシピの差異要素を分類し、特徴パターンを用いてレシピを分類することで、ネーミングコンセプトを抽出している。

クックパッドや楽天レシピなどの投稿型レシピサイトでは、膨大なレシピデータの中からユーザが目的のものを探するために、キーワード検索を行うか、階層的なカテゴリーをたどる必要がある。このカテゴリーは、材料（肉・野菜・魚介など）や料理の種類（和・洋・中）に基づいたものが中心であるため、例えば「鉄分補給」や「子供の運動会」などの目的に応じたレシピを探すのに十分ではない。そこで、目的別の料理レシピカテゴリーを構築し、そのカテゴリーにレシピを自動分類する手法を提案した研究がある[Kanauchi 2016]。レビュー情報の特徴抽出を用いてカテゴリーの分類を行い、タグの自動付与に基づいてカテゴリー分類を行っている。そして、類似す

るカテゴリー名に関してグルーピングを行い、手作業でディレクトリの整理を行っている。

クックパッドや楽天レシピなどのユーザ投稿型レシピ検索サイトの利用者は、レシピ投稿者もレシピを利用するユーザも、その調理技術レベルは、上から下まで幅広く分布していると考えられるため、技術の高い調理者が投稿したレシピは、場合によっては、料理初心者にとって配慮に欠けたものになっている可能性がある。しかしながら、レシピ中に料理のコツやポイントなどのアドバイスがたくさん記述されていれば、料理初心者でも、技術の高い調理者の投稿レシピに挑戦できる可能性がある。そこで、ユーザ投稿型レシピ検索サイトにおいて、投稿ユーザが生成した調理手順から利用ユーザの調理を補助する役割をもつ料理アドバイスを抽出し、類似するレシピに補完することで、初心者にやさしいレシピを提示できるシステムの提案を行った研究がある。各料理手順がアドバイスであるか否かの分類をする必要があるが、各手順に含まれる単語を素性とした深層学習に基づく手法を提案し、類似レシピの手順中に記載された料理アドバイスの自動検出を提案している[Furumoto 2016]。また、当該レシピへの料理アドバイスの挿入として、レシピ A に対して、類似レシピ B 中のアドバイスを挿入するといったタスクが必要となる。2 文の内容がどの程度似ているのかを測る方法にはさまざまなものが存在するが、各文を文中の内容語（名詞・動詞・形容詞）を用いてベクトルとして表現し、コサイン類似度で文間の類似度を測るというベクトル空間型モデルを用いた手法では、対応付けの精度が低く、複数のレシピ間で類似手順を対応付けるというタスクにおいて、**paragraph vector** という手法を用いて各手順を分散表現で表すことで、ベクトル空間型モデルを用いる場合よりも高い精度で類似手順の対応付けが可能であることを示している[Shigeta 2018]。また、関連したものとして、調理手順そのものではない付加的説明文を自動生成する手法を検討した研究がある[Kagawa 2022]。研究料理レシピのような創作や操作手順の説明文書では、なぜそれをするのかといった、手順にとどまらない付加的な情報が文書の有用性を向上させるが、そのような文書を書く経験に乏しい者にとっては何を書くべきか考えることすら難しいといった課題がある。そこで、料理レシピの読者に役に立つ付加的情報が「代替品・アレンジ・調理工程の理由・適した材料・注意点・盛り付け」に類別されることを実験的に明らかにし、ユーザが入力した文に対し、クックパッドデータセットで学習した GPT-2 モデルにより文を生成し、それらに BERT に基づく分類モ

デルを適用して上記の 6 種類に属する文を選出し、提案する仕組みを構築している。

また、料理レシピを対象にした言語処理を目的とした料理オントロジーがいくつか公開されている。代表的なものとして、日本語の料理レシピおよび特許データベースを情報源とし、表層パターンおよび統計処理を用いて上位・下位関係、同義語、属性、部分・全体関係の用語対の候補を自動収集した後、それらをすべて人手で確認することで、料理オントロジーを構築したものがある[Nanba 2014]。他にも、材料および調理動作に関する代表語と同義語、クックパッドに投稿された料理レシピデータ、クラウドソーシングで収集した料理の印象データを情報源とし、材料に加え、料理名や調理法(料理名と調理動作と材料の関係)を含む料理基本知識ベースを構築したもの[Kiyomaru 2018] がある。

その他、料理レシピを対象にした研究としては、レシピの完成画像の色情報に基づいてレシピの可視化を行うことで、ユーザが検索結果の全体像を効率的に把握可能とする手法を提案し、色情報を利用することで、ユーザに似た印象を与える写真を可視化空間上で近くに配置可能となり、ユーザが興味があるレシピの効率的な選択を可能としたもの[Hirakawa 2017]や、料理動画から調理動作の並列性、動作の頻度、カメラワークによる編集意図を難易度要素として抽出し、料理動画の時間特性に基づき難易度を算出することで、料理レシピの難易度を判定する手法を提案したもの[Akiguchi 2018]などがある。

料理手順を対象にした研究以外では、「ふわふわ」や「さくさく」といったオノマトペといわれる表現に注目した研究が複数ある。オノマトペとは、音の響きや対象物体の形状などを感覚的に表現したものであり、擬音語や擬態語ともよばれ、近年、漠然とした料理のイメージに基づき効率的にレシピを検索する一手段などとして研究が行われている。その一例としてオノマトペロリシステム[Watanabe 2015] がある。本システムでは、クエリとして入力されたオノマトペを関連度の高い食材のキーワードに置き換えて検索する手法や、レシピに含まれる各用語のオノマトペとの関連度合計を計算することで、クエリであるオノマトペと各レシピとのランキングを再現する手法が提案されている。例として「さくさく」で検索した場合、料理レシピの記述に「さくさく」を含まずに「さくさく」の食感である料理レシピの検索を可能としている。このシステムでは、オノマトペと強く共起する特定の食材の有無がランキングに大きな影響を与えるためである。例えば、” さくさく ” に強く共起する食材は ” 薄力粉 ”

または ”小麦粉 ” が多いことが確認されており，これらの食材が含まれているまたは，強く共起する食材を多く含んでいた場合，食感がさくさくと判定される。また，飲食店サイトにおいて，口コミにオノマトペが多く登場することに着目し，オノマトペと食品カテゴリーを，トピックモデル LDA を用いてクラスタリングすることにより，オノマトペを特徴量とした意外性の高い飲食店の推薦方法を提案した研究 [Kato 2015a] もある。ユーザ毎に重視する五感，もしくは食品カテゴリーで重視される五感に関連するオノマトペを軸にしたレコメンドをすることでユーザの嗜好にあったレコメンドが可能であることを示している。他にも，Twitter 上などのおいしさの表現分析として，オノマトペなどと食品の単語共起頻度や Simpson 係数，Dice 係数を，シズルワードと食品毎に求めた研究 [Kato 2015b] などもある。

2.5 テキストマイニングと料理・味覚表現における研究分野 の課題

生化学分野におけるコクに関する研究においては，コクの定義を試みることや，コクを生かした商品開発などを目的に官能評価を実施するなど様々なコクに関する研究開発が行われている。これらの研究で示されたコクの定義やコクを付与する可能性のある物質を特定する試みは，コク研究の更なる発展に向けて有用なものである。しかし，提案されているコクの定義はやや曖昧さが残るものであり，定義されたものが一般生活者にとって広く当てはまるものであるかどうかの検討は十分であるとはいえない。また，多くの官能評価は，訓練されたパネリストによって実施されたものであり，一般的な生活者がコクを感じる際の評価と一致するものであるかどうかの検証は十分であるとはいえない。

日本酒に関する味覚に関する研究においては，アルコール飲料業界では，酒の味を分かりやすく表現することは重要な課題の一つとされ，関連する多くの研究開発が行われており，消費者にとって自分の好みの酒を選択する際などに有用な情報が増えてきている。しかし，提案された味や香りに関する表現は，やや曖昧さが残るものであるといえる。また，共通理解のために一定の訓練が必要となったり，訓練されたパネ

リストによって実施された官能評価によるものであるため、一般的な生活者が日本酒の味を感じる際の評価と一致するものであるかどうかの検証は十分であるとはいえない。また、テキストマイニング手法による日本酒の味表現に関する研究では、これらの研究対象とするデータは、専門家により執筆された雑誌などの著作物や、ごく少数の者による日本酒を味わった際の表現、特定の蔵元による日本酒に関する説明文を対象にしているものであり、一般的な生活者の声が広く反映されたものであるとはいえない。そのため、一般的な生活者の評価と一致するものであるかどうかの検証は十分であるとはいえない

認知言語学・身体化認知に関する分野においては、他者情報や身体感覚が味覚に与える影響を観察する先行研究が確認され、今後の関連研究の更なる発展が期待できるものである。しかし、コクなどの明確に定義ができない味覚に関する、他者情報の影響を検証しようとする研究は十分ではない。

食に関わる言葉研究においては、コクを含む表現は、明確な分類を困難なものにしているといえ、明確に定義ができない味覚に関する表現の意図解明を目指した研究はみられない。また、研究対象とするデータは、書籍や漫画などの著作物が多く、一般的な生活者の声が反映された料理レシピサイトの口コミ情報を対象にした研究は少ない。

料理レシピを対象にした自然言語処理に関する研究分野では、料理レシピや料理に関連するオノマトペを対象にした研究の発展が著しく、本研究分野の更なる発展が期待されるものといえる。しかし、コクなどの明確に定義ができない味覚に関する表現の共通的な意図解明を目指すとした、料理レシピサイトの口コミ情報を対象にした研究はみられない。

以上から、料理レシピサイトの口コミ情報を対象に、テキストマイニング手法を用いて、一般的な生活者が使用する、明確に定義ができない味覚表現の特徴を明らかにしようとする研究は見当たらず、本研究の目的である、明確に定義ができない味覚表現に対する、共通的な意図解明に向けた定量的評価手法の提案に関する研究は十分とはいえない。

2.6 おわりに

本章では、最初に、本研究の分析手法となる自然言語処理の流れを整理し、本研究で用いた BERT および LIME に関する概念や手法を説明した。次に、本研究での分析対象となる口コミ情報が投稿される料理レシピ共有サイトについて説明した。そして、最後に本研究に関連の深い、生化学分野におけるコク、日本酒に関する味覚、認知言語学・身体化認知、食に関わることば、料理レシピを対象にした自然言語処理、に関するそれぞれの先行研究を概観し、課題を整理した。その結果、本研究の目的である、各人の味覚によって判断され用いられる、明確に定義ができない味覚表現に対する共通的な意図理解に向けた、定量的評価手法の提案に関する研究は十分とはいええないことを確認した。

第3章 飲食物への味覚表現と味覚関連指標 の関係性に関する定量的評価

3.1 はじめに

本章では日本酒を取り上げる。日本酒は、その銘柄を選ぶ場合、多くは甘口や辛口、味の濃淡などを示す指標や、その味を説明したものを参考にすることが多い。一般的に、日本酒は糖分が多く含まれたものが甘口、糖分が少ないものが辛口といわれる。しかし、甘口は甘みを感じるものとして容易に他人と共有できる一方で、辛口は辛いと感じる成分が含まれているわけではなく、その味を何の味であるか表現しづらいものといえる。さらには、説明された味の情報として多く用いられる、キレやフルーティ、コク、まろやかなどといった表現は、明確な定義がないものであるが、それらのニュアンスや感覚などは、曖昧ではあるがおおよそ我々の多くに共有されているものでもある。

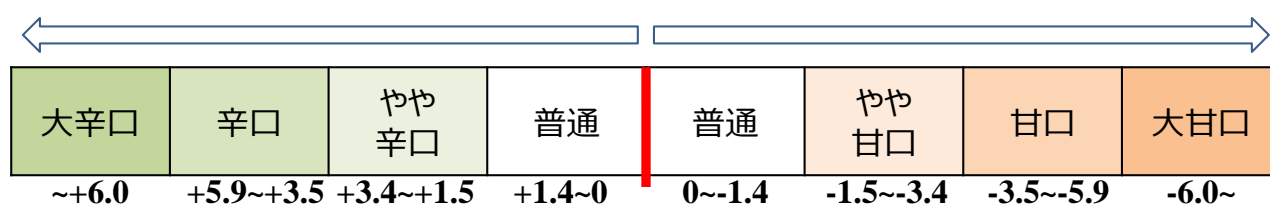


図 3-1 日本酒度の概要

以上の背景から、本研究では EC サイト上の日本酒に関するユーザの口コミ情報を用いて、一般生活者が日本酒を飲んだ際に用いられる明確に定義ができない味覚表現の共通的な意図理解を目的とした定量的評価手法の提案を行う。本実験では、味覚に関する表現と、「日本酒度」「酸度」などといった日本酒に用いられる各種指標との関係性を明らかにし、明確に定義ができない日本酒に関する味覚表現の定量的な評価を目的とする。具体的には、EC サイト上の日本酒に関するユーザの口コミ情報を用い

て、一般生活者が日本酒を飲んだ際に感じる味覚に関する特徴の解明を目指すものである。

3.2 日本酒の味覚表現の検証

本節では、分析手法および本実験に用いるデータセット作成に関して説明する。

3.2.1 使用データ

本実験では、国立情報学研究所より、研究利用目的に限定公開されている楽天市場のデータを使用する。

楽天市場は、日本の楽天グループが運営する EC サイトである。楽天市場データのうち、2015 年 1 月から 2019 年 12 月の間における商品レビュー情報の中から、商品ジャンルが日本酒に関する口コミ情報を分析対象とした。楽天市場における口コミ情報例を図 3-2 に示す。

★★★★★ 5

商品を使う人:自分局

まるでワインのようで、美味しい

獺祭(だっさい)純米大吟醸磨き45を購入いたしました。

冷やしませず、燗にもせず、そのままグラスに注いでいただきます。まずは香をかぐと安い日本酒にあるのツーンと感じは無く、ほのかに甘味が感じられる香り。続いて一口、まるやかでトロっとした舌触りかつフルーティな感じで、まるでワインを飲んでみたいですね。日本酒をあまり飲まない妻も、美味しいと言って一緒に飲んでおります。

★★★★★ 5

純米大吟醸酒 (旭酒造)

獺祭は、私が日本酒を好きになり始めた頃から、すぐにわたし好みのお酒と感じた銘柄です。

まだ日本酒を好きになって1年で少しづつではありますが、色々飲んで来ても、やはりこちらは好みのタイプでした。

山田錦を39%まで磨き上げた純米大吟醸酒なのに、びっくりする位安い価格。

企業努力でこの価格が実現しているのようですが、それだけでなく、本当に美味しいから素敵です。

獺祭の中で最も香りがきれいなお酒で喉越し軽快なスッキリとした味わいと酸元がおっしゃっている通り、高嶺過ぎず魅力あふれる美人な日本酒なイメージです。

でも磨き50も旨味も好きなので、次回注文する時も、多分両方頼んでしまうと思います。

- 原料 山田錦
- 精米歩合 39%
- アルコール度数 15~16度
- 日本酒度 6
- 酸度 1.1
- 蔵出荷年月 24年11月

出所：楽天市場

図 3-2 楽天市場における口コミ情報例

表 3-1 に対象としたデータの日本酒ジャンル別の総件数を示す。

表 3-1 楽天市場の日本酒に関するデータ概要

純米大吟醸酒	純米吟醸酒	純米酒	大吟醸酒	吟醸酒	本醸造酒	合計
17,378	8,881	12,694	5,291	2,188	4,001	50,433

3.2.2 分析方法

本実験では、日本酒ラベル上に表示される「日本酒度」「酸度」と、日本酒度と酸度を用いて算出される「甘辛度」「濃淡度」を目的変数、口コミ情報で確認された各種味覚表現に関する頻度を用いた指標を説明変数とし、重回帰分析を行う。重回帰分析では、回帰係数の数値の大きさを、その説明変数の目的変数への影響力の強さを統計的に比較できる。また、AIC などによるステップワイズ法によって、説明変数を削減し、説明力の高いモデルを作ることが可能となる。本実験では、口コミ情報で確認された各種味覚表現が日本酒に関する各種指標への関係性の検証を行うことから、本実験において、本分析手法を用いることは適当といえる。

各目的変数の概要[National Tax Agency 2018], [The National Research Institute of Brewing 2013] を表 3-2 に示す。

表 3-2 目的変数の概要

目的変数	内容
日本酒度	清酒の比重を示す指標で、甘口・辛口の目安になる。(+)になるほど糖分が少なく辛口、(-)になるほど糖分が多く甘口の傾向を示す。
酸度	乳酸、コハク、リンゴ酸などの有機酸の含有量の多少を示す指標。味の濃淡や甘辛に影響し、酸度が高いと味が濃く感じられる。また、日本酒度が同じ酒で比べた場合、甘味が隠れるため、辛口に感じる。
甘辛度	日本酒の甘口、辛口は糖分と酸のバランスで決まることから、日本酒度と酸度から計算して甘辛の指標としたもの。数値が高いほど甘口であり、低いほど辛口であることを示す。 $193593/(1443+\text{日本酒度})-1.16\times\text{酸度}-132.57$ で算出される。
濃淡度	日本酒の濃淡は糖分と酸のバランスで決まることから、日本酒度と酸度から計算して濃淡の指標としたもの。数値が高いほど濃醇であり、低いほど淡麗であることを示す。 $94545/(1443+\text{日本酒度})+1.88\times\text{酸度}-68.54$ で算出される。

3.2.3 実験環境

口コミ情報に関する前処理には Python を用い、形態素解析には MeCab と、新語・固有表現に強いとされる単語分かち書き辞書である mecab-ipadic-Neologd を用いた。重回帰分析には R を用い、変数選択には、パッケージ MASS の関数 stepAIC を用いた。

3.2.4 データセットの作成

本実験で用いるデータセットの作成手順およびデータセットの概要を示す。

(1) 口コミ情報データ抽出

日本酒は、同銘柄の場合でも特定名称の違いによって味が異なる。特定名称とは、原料や製造方法の違いによって、吟醸酒、純米酒、本醸造酒などと分類されたものである。そのため、表 3-1 で示した日本酒のうち、特定名称別の銘柄毎に、50 件以上の口コミ情報が確認されたものを対象とした。結果、37 種類・9529 件の口コミ情報を抽出した。

(2) 目的変数データの抽出

「日本酒度」「酸度」については、上記で抽出した日本酒の、蔵元 HP、楽天市場若しくは酒類販売店 HP の商品情報で確認できたものを抽出した。「甘辛度」「濃淡度」は表 3-2 によって示した算出式を用いた。

(3) 目的変数データの抽出

本実験では、日本酒の味覚表現に関する特徴語として表れやすいと考えられる用語に限定することとした。味覚に関する特徴語として、「直接的な味覚用語」「テクスチャー用語」「シズルワード」を選定する。味覚用語としては、味ことばとして収集された「塩味」「甘味」「酸味」「うま味」「苦味」「辛味」「渋味」の各味覚要素に関する用

語[Seto 2003]を用いる。テクスチャー用語としては、ISO11036 Texture Profile の3要素と対応する形で分類された「力学的特性」「幾何学的特性」「その他の特性」に関する用語[Hayakawa 2013a], [Hayakawa 2013b]を用いる。シズルワードは、「味覚系」「食感系」「情報系」の3つの領域に分類されたもの[Ohashi 2016], [B. M. FT Kotoba Lab 2016]を用いる。以上の中から、先で示した抽出された9529件の口コミ情報のうち、30件以上が確認された用語を抽出対象とした。なお、テクスチャー用語としては、さらっ、さらり、さらさらはなめらかさとすべりとして一つにカウントするなど、小分類の単位で一つとした。抽出した味覚表現を表3-3に示す。

表 3-3 抽出された味覚表現

抽出された味覚表現用語	口コミ上で用いられた味覚表現
辛味	辛口, 辛い, 辛味, 辛さ,
甘味	甘い, 甘口, 甘味, 甘さ, 甘すぎる
うま味	うまみ
コク	コクがある, コク
キレ	キレ, キレのよい
スッキリ	スッキリ, スッキリした
フルーティ	フルーティ
まろやか	まろやか
サッパリ	サッパリ
爽やか	爽やか
なめらかさ	口当り, 口当りがよい, 舌触り, するする, なめらか
流れやすさと濃厚感	濃厚, まったり, とろっ, とろける, とろみ, とろとろ, どろり, 口どけ, とろり, クリーミー
なめらかさとすべり	さらっ, さらり, さらさら, するする
やわらかさ	軽い, やわらかい
重さ	重い, どっしり

(4) 単語頻度を用いた値の算出

抽出された味覚表現に関する用語について、日本酒銘柄別に頻度を集計する。集計した頻度のうち、出現頻度が1回若しくは出現率が1%未満のものは、その日本酒の特徴を表したものではないと考えられるため、削除することとした。また、1種類の日本酒のみで確認された味覚表現に関する用語は、日本酒の味覚表現に関する特徴語として考えにくいいため、削除することとした。そして、本実験では、該当の日本酒自体のレビュー数によって味覚表現に関する用語の出現頻度が影響されてしまうことを考慮し、各用語の出現頻度を正規化したものを用いることとした。

(5) データセットの概要

結果、31種類計9010件の口コミ情報の日本酒と14種類の味覚表現に関する用語によるデータセットを作成した。本データセットの基礎統計量を表3-4に示す。

表 3-4 本データセットの基礎統計量

変数区分	変数	平均	標準偏差
目的変数	日本酒度	3.217	4.308
	酸度	1.420	0.245
	甘辛度	-0.354	0.374
	濃淡度	-0.496	0.573
説明変数	辛味	0.551	0.367
	甘味	0.496	0.326
	うま味	0.082	0.169
	コク	0.037	0.100
	キレ	0.041	0.130
	スッキリ	0.569	0.304
	フルーティ	0.472	0.450
	まろやか	0.052	0.117
	サッパリ	0.046	0.134
	なめらかさ	0.356	0.340
	流れやすさと濃厚感	0.038	0.130
	なめらかさとすべり	0.048	0.157
	やわらかさ	0.055	0.151
	重さ	0.027	0.100

3.3 重回帰分析による検証結果

本節では、3.2節で作成したデータセットを用いた相関分析および重回帰分析の結果を示す。

3.3.1 各変数間の相関分析

(1) 目的変数と説明変数間の相関分析

各目的変数と説明変数間の相関係数の算出結果を表 3-5 に示す。日本酒度と辛味は有意な正の相関、甘味は有意な負の相関が確認された。その他、日本酒度と、口当たりや舌触りなどのなめらかさや、濃厚やまったりなどの流れやすさと濃厚感といった味覚表現と有意な負の相関が確認された。酸度については、コクと有意な正の相関が確認された。甘辛度については、辛味は有意な負の相関、甘味は有意な正の相関が確認された。その他、甘辛度と、口当たりや舌触りという表現が含まれる、なめらかさは有意な正の相関が確認された。濃淡度については、甘味、コクと有意な正の相関が確認された。

表 3-5 各目的変数と説明変数間の相関係数

	日本酒度	酸度	甘辛度	濃淡度
辛味	0.504 **	-0.133	-0.434 *	-0.278
甘味	-0.542 **	0.237	0.396 *	0.374 *
うま味	-0.077	0.274	-0.127	0.246
コク	-0.192	0.430 *	-0.123	0.411 *
キレ	0.012	-0.066	0.036	-0.057
スッキリ	-0.231	-0.113	0.331	-0.013
フルーティ	-0.233	-0.090	0.315	0.006
まろやか	-0.215	0.219	0.061	0.249
サッパリ	-0.060	0.129	-0.035	0.124
なめらかさ	-0.474 **	0.176	0.370 *	0.303
流れやすさと濃厚感	-0.377 *	0.157	0.283	0.255
なめらかさとすべり	-0.084	0.100	0.013	0.109
やわらかさ	-0.039	-0.108	0.123	-0.074
重さ	-0.290	0.214	0.147	0.271

*p < .05, **p < .01, ***p < .001

(2) 説明変数間の相関分析

各説明変数間の相関係数の算出結果を表 3-6 に示す。

表 3-6 各説明変数間の相関係数

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 辛味	—													
2 甘味	-.300	—												
3 うま味	-.068	-.315	—											
4 コク	.000	.000	.534 **	—										
5 キレ	.142	.090	-.100	-.070	—									
6 スッキリ	-.175	.188	.016	-.357 *	.293	—								
7 フルーティ	-.630 ***	-.047	.201	.244	-.125	-.053	—							
8 まろやか	-.159	.146	-.131	-.161	.039	.106	-.052	—						
9 サッパリ	.054	.028	.328	.151	-.091	.271	.032	-.150	—					
10 なめらかさ	-.405 *	.081	.346	.222	-.220	.107	.390 *	.409 *	.134	—				
11 流れやすさと濃厚感	-.158	.300	-.141	.301	-.069	.046	.022	.000	-.076	.219	—			
12 なめらかさとすべり	-.301	-.062	-.145	-.111	-.073	.217	.298	-.135	.226	.105	-.057	—		
13 やわらかさ	.159	.051	-.017	-.134	.659 ***	.432 *	-.183	-.145	.522 **	-.152	-.065	.188	—	
14 重さ	-.057	.242	-.059	.374 *	-.009	.023	-.013	-.117	-.061	.037	.912 ***	-.039	-.063	—

*p<.05, **p<.01, ***p<.001

「うま味とコク」、「キレ、スッキリ、サッパリとやわらかさ」、「フルーティ、まろやかとなめらかさ」、「流れやすさと濃厚感と重さ」は有意な正の相関があることが確認された。また、「辛味とフルーティ、なめらかさ」、「コクとスッキリ」は有意な負の相関があることが確認された。

表 3-6 に示された独立変数間の相関係数は一部強い相関がみられ、多重共線性の問題が考えられるため、本データセットを用いて重回帰分析を行う際、変数選択による分析を試みた。

3.3.2 影響要因の重回帰分析

本実験では、ステップワイズ法を用いて、AIC が最も小さい説明変数の組み合わせを最も当てはまりがよい重回帰モデルとして選定し、選択された説明変数が、目的変数に及ぼす影響を確認する。

(1) 日本酒度に及ぼす影響

日本酒度を目的変数とした重回帰分析の結果を表 3-7 に示す。ステップワイズ法による変数選択を行った結果、辛味、甘味、なめらかさが選択された。そのうち、甘味となめらかさについて、有意な負の回帰係数であることが確認された。10 以上で多重共線性の影響があるといわれる VIF 値については、いずれも 10 未満であり、多重共線性の影響は無いと考えられる。

表 3-7 日本酒度を目的変数とした重回帰分析

変数名	重回帰係数	VIF
辛味	2.706	1.309
甘味	-5.884 **	1.101
なめらかさ	-4.373 *	1.199
(Intercept)	6.199 **	
Adjusted R-squared	0.465	
F値	9.413 ***	

*p<.05, **p<.01, ***p<.001

(2) 酸度に及ぼす影響

酸度を目的変数とした重回帰分析の結果を表 3-8 に示す。ステップワイズ法による説明変数の選択結果、甘味、うま味、コク、フルーティ、まろやか、なめらかさとすべりが選定された。そのうち、コク、なめらかさとすべりについて、有意な正の回帰係数であることが確認された。また、甘味とまろやかの回帰係数は、有意な傾向にある正の、フルーティの回帰係数は有意な傾向にある負の回帰係数であることが確認された。VIF 値については、いずれも 10 未満であり、多重共線性の影響は無いと考えられる。

表 3-8 酸度を目的変数とした重回帰分析

変数名	重回帰係数	VIF
甘味	0.215	1.194
うま味	0.409	1.684
コク	1.100 *	1.548
フルーティ	-0.180	1.223
まろやか	0.667	1.076
なめらかさとすべり	0.546 *	1.209
(Intercept)	1.264 ***	
Adjusted R-squared	0.335	
F値	3.429 *	

*p < .05, **p < .01, ***p < .001

(3) 甘辛度に及ぼす影響

甘辛度を目的変数とした重回帰分析の結果を表 3-9 に示す。ステップワイズ法による説明変数の選択結果、甘味、キレ、スッキリ、フルーティ、サッパリ、なめらかさ、なめらかさとすべり、やわらかさが選定された。そのうち、甘味、フルーティ、やわらかさについては、有意な正の回帰係数であることが確認された。キレ、サッパリについては、有意な負の回帰係数であることが確認された。VIF 値については、いずれも 10 未満であり、多重共線性の影響は無いと考えられる。

表 3-9 甘辛度を目的変数とした重回帰分析

変数名	重回帰係数	VIF
甘味	0.409 *	1.062
キレ	-1.876 *	3.722
スッキリ	0.354	1.371
フルーティ	0.358 *	1.435
サッパリ	-1.810 *	2.770
なめらかさ	0.277	1.309
なめらかさとすべり	-0.605	1.347
やわらかさ	2.263 *	5.409
(Intercept)	-0.962 ***	
Adjusted R-squared	0.411	
F値	3.527 **	

*p < .05, **p < .01, ***p < .001

(4) 濃淡度に及ぼす影響

濃淡度を目的変数とした重回帰分析の結果を表 3-10 に示す。ステップワイズ法による説明変数の選択結果、甘味、うま味、コク、フルーティ、まろやか、なめらかさとすべりが選定された。そのうち、甘味、コク、まろやかについて、有意な正の回帰係数であることが確認された。また、なめらかさとすべりの回帰係数は、有意な傾向にある正の回帰係数であることが確認された。VIF 値については、いずれも 10 未満であり、多重共線性の影響は無いと考えられる。

表 3-10 濃淡度を目的変数とした重回帰分析

変数名	重回帰係数	VIF
甘味	0.766 *	1.194
うま味	1.050	1.684
コク	2.203 *	1.548
フルーティ	-0.269	1.223
まろやか	1.582 *	1.076
なめらかさとすべり	1.204 .	1.209
(Intercept)	-1.055 ***	
Adjusted R-squared	0.386	
F値	4.033 **	

*p<.05, **p<.01, ***p<.001

3.3.3 実験結果の考察

(1) 日本酒度

日本酒度は、プラスになるほど辛口、マイナスの場合甘口とされている。日本酒度を目的変数とした重回帰分析結果、甘味については有意な負の回帰係数であることが確認された。口コミ情報を対象にした分析結果においても、日本酒度が示す指標と同様の傾向を示しており、甘口の日本酒は甘みを感じるものとして表現しやすい傾向にあることが示唆された。一方、辛味については、その回帰係数が正の影響ではあったが、その有意性を確認することができなかった。日本酒の場合の辛口は、糖分が少ないものを辛口といい、従来用いる辛味と異なり、辛いと感じる成分が含まれているわけではない。そのため、その味を辛口と直接的に表現することが難しい場合があることが考えられる。また、なめらかさは、有意な負の回帰係数であることが確認されており、甘口の日本酒を、口当たりや舌触りなどのなめらかさといった用語で表現される可能性があることが示唆された。

(2) 酸度

酸度は、酸度が高いと味が濃く、うまみを感じられるとされている。酸度を目的変数とした重回帰分析結果、コクについて、有意な正の回帰係数であることが確認された。有機酸による日本酒の味の濃さやうまみを感じる場合、コクと表現される傾向にあることが示唆された。

(3) 甘辛度

甘辛度は、その値が高いほど甘口、低いほど辛口を示すとされる。甘辛度を目的変数とした重回帰分析結果、甘味について、有意な正の回帰係数であることが確認された。日本酒度と同様に、口コミ情報を対象にした分析結果においても、甘辛度が示す指標と同様の傾向を示していることが確認された。また、フルーティや、やわらかさといった表現も、有意な正の回帰係数であることが確認されており、甘口の日本酒を、フルーティややわらかさといった用語で表現される可能性があることが示唆された。一方、日本酒度の場合と同様に、辛味については、その回帰係数が正の影響ではあったが、その有意性を確認することができなかった。しかし、キレやサッパリといった表現は、有意な負の回帰係数であることが確認された。辛口とされる日本酒の場合、辛口の代わりに、キレやサッパリといった用語で表現される可能性があることが示唆された。

(4) 濃淡度

濃淡度は、数値が高いほど濃醇であり、低いほど淡麗であることを示すとされる。濃淡度を目的変数とした重回帰分析結果、甘味、コク、まろやかについて、有意な正

の回帰係数であることが確認されており，日本酒の味の濃醇さを感じる場合，甘味やコク，まろやかと表現される傾向にあることが示唆された。

以上実験結果より，「日本酒度」「酸度」「甘辛度」「濃淡度」と，日本酒に関する各種味覚表現における関係性について，一定の定量的な評価を可能にしたといえる。

3.4 まとめ

本実験では，日本酒を味わった際の明確に定義ができない味覚表現の共通的な意図理解を目的とした定量的評価手法の提案を行った。重回帰分析を用いた日本酒の味覚表現の検証の結果，辛口とされる日本酒の場合，辛口の代わりに，キレやサッパリといった用語で表現される可能性があることが確認された。また，甘口とされる日本酒の場合，なめらかやフルーティ，やわらかといった表現が用いられることが確認された。その他，日本酒の味に濃醇さを感じられる場合，甘味やコク，まろやかと表現される傾向にあることが示唆された。一連のテキストマイニングによる分析手法によって，「日本酒度」「酸度」「甘辛度」「濃淡度」と，日本酒に関する各種味覚表現における関係性について，一定の定量的な評価を可能にしたといえる。

第4章 明確に定義ができない味覚表現の意 図理解に向けた定量的評価

4.1 はじめに

前章では、日本酒に用いられる、日本酒度や酸度、甘辛度、濃淡度といった各種指標と、一般生活者が明確な定義がなく使用している、日本酒に関する味覚表現との関係性を明らかにした。しかし、一般生活者が明確な定義がなく使用している、味覚表現自体の曖昧性は残っており、これら味覚表現の共通的な意図理解に向けた定量的評価が必要といえる。

本研究では、曖昧な味覚表現の例としてコクを取り上げ、明確に定義ができない味覚表現に対する、共通的な意図解明を目的とした定量的評価手法を提案する。手法提案に向けて、以下、4.2 章では、曖昧な味覚表現における特徴語の解明に向けた実験を、次の 4.3 章では、曖昧な味覚表現が持つ特性の解明に向けた実験を行い、各結果について考察する。そして、一連の実験結果から、本研究で提案した手法の有効性を確認する。

4.2 曖昧な味覚表現における特徴語の分析

本節では、分析手法および本実験に用いるデータセット作成に関して説明する。

4.2.1 使用データ

本研究では、国立情報学研究所より、研究利用目的に限定公開されているクックパッドデータを使用する。クックパッドは 1998 年に開設された一般ユーザが作成したレシピを Web 上に掲載するユーザ投稿型のレシピ共有サイトで、多くの一般ユーザに利用され、数百万件以上にわたる豊富なレシピ情報を有している。データには、2014

年 9 月 30 日までにクックパッドに掲載された 172 万品のレシピやレシピに関する材料、分量、手順のほか投稿されたレシピに対するユーザからの口コミ情報である「つくれぽ」などが収録されている。表 4-1 に主要なデータの総件数を示す。

表 4-1 クックパッドのデータ概要

(a)recipes			
カラム	内容	レコード数	容量
id	レシピのID	1,715,595	934.8MB
user_id	レシピ作者のID	1,715,595	
title	レシピのタイトル	1,715,595	
description	レシピの概要	1,715,590	
serving_for	レシピの分量	1,599,191	
advice	レシピのコツ・ポイント	1,588,170	
history	レシピの生い立ち	1,715,594	
published_at	レシピの公開日	1,715,595	

(b)ingredients			
カラム	内容	レコード数	容量
recipe_id	レシピのID	12,725,006	911.4MB
name	材料の名前	12,724,287	
quantity	材料の分量	12,680,050	

(c)steps			
カラム	内容	レコード数	容量
recipe_id	レシピのID	8,849,850	1,412.6MB
position	手順の位置	8,849,850	
memo	手順の内容	8,769,210	

(d)tsukurepos			
カラム	内容	レコード数	容量
user_id	レシピ作者のID	9,449,479	2,322.8MB
recipe_id	レシピのID	9,449,479	
message	つくれぽの内容	9,449,478	
comment	レシピ作者の内容	9,449,137	
entered_at	つくれぽの投稿日	9,449,479	

(e)search_categories			
カラム	内容	レコード数	容量
id	カテゴリのID	1,099	0.1MB
title	レシピのタイトル	1,099	
parent_id	親となるカテゴリのID	1,099	

(f)search_category_recipes			
カラム	内容	レコード数	容量
search_category_id	カテゴリのID	164,912	38.8MB
recipe_id	レシピのID	164,912	

クックパッドデータのうち、本研究では、口コミ情報である「つくれぽ」を分析の対象とする。口コミ情報を分析対象とすることで、一般生活者がコクを感じる際の特徴が確認されやすいと考える。クックパッドにおける口コミ情報例を図 4-1 に示す。



出所：クックパッド

図 4-1 クックパッドの口コミ情報例

(1) コク表現を含む口コミ情報の抽出

今回対象とする「コク」をキーワードとして、つくれぽの中にコクという表現が直接的に含まれるレシピ情報を抽出した。その結果、直接的にコクという表現が含まれる口コミ情報 49,342 件を得た。

(2) サッパリ表現を含む口コミ情報の抽出

味覚表現の属性を分類した研究[Muto 2003] では、味覚表現の一つとして「濃淡」という属性が分類されている。そこでは、「濃」として、コクに相当する「コックリ」が該当するとし、その他にも「コッテリ」などの表現が分類されている。一方、「淡」としては「サッパリ」「あっさり」などの表現が分類されている。そこで、コク表現がみられた 49,342 件のつくれぽデータのうち、コク表現と逆接の助詞および接続詞[Iori 2000] と共起する単語を確認した。その結果を表 4-2 に示す。

表 4-2 コク表現と逆接の助詞および接続詞と共起する単語

逆接の助詞, 接続詞	共起語	件数
けど, けれど, でも, けれども, のに, が, だが, ものの, しかし, にもかかわらず, ところが	サッパリ	1323
	あっさり	494

なお、抽出対象には一般名詞を除き、逆接の助詞および接続詞を挟んだコク表現と共起語の出現順は問わず、纏めて集計し、100 件以上が確認されたものである。

実際の使用例を次に示す。

- ・サッパリなのにコクがあつて美味しかったです！
- ・牛乳であっさりだけどチーズでコク旨♪

以上から、サッパリやあっさり表現は、コクと対に位置する属性と考えられ、特徴語の抽出を目的としたテキスト分類で高い精度を得るために、これら表現を選定することは適当といえる。本研究では、コク表現の抽出方法と同様に共起件数の多かった

「サッパリ」をキーワードとして、つくればの中にサッパリという表現が直接的に含まれる口コミ情報を抽出した。その結果、直接的にサッパリという表現が含まれる口コミ情報 198,156 件を得た。

4.2.2 分析方法

本実験では、まず、コク表現が含まれる口コミ情報とサッパリ表現が含まれる口コミ情報を抽出し、テキスト分類のためのデータセットを作成する。次に、データセットを用いてテキスト分類器を学習する。テキスト分類において高い分類精度が確認された場合、コクを感じる際の特徴語の存在を検証することが可能となる。さらに、分類に効いた特徴語を可視化することにより、具体的な特徴語の解明を目指す。

4.2.3 実験環境

本分析には Python を使い、形態素解析には MeCab と、新語・固有表現に強いとされる単語分かち書き辞書である mecab-ipadic-Neologd を用いた。

4.2.4 データセットの作成

コクとサッパリ表現を含む口コミ情報のデータセット作成の手順を以下に示す。

(1) 単語の抽出

本実験では、料理レシピに対する口コミ情報を分析対象としていることから、分類に寄与する特徴語として表れやすいと考えられる単語に限定することとした。食材や調味料などに関する名詞、調理に関連する動詞については、主に、料理レシピを対象にした言語処理を目的として公開されている料理オントロジー [Nanba 2014] を、料理名称については、クックパッドデータに収録されたレシピタイトルをそれぞれ利用した。また、味覚に関する表現としては、主に、先行研究で示した、味ことば [Seto 2003]、テクスチャー用語 [Hayakawa 2013a], [Hayakawa 2013b], シズルワード [Ohas

hi 2016], [B. M. FT Kotoba Lab 2016] を対象とした。また、「甘すぎない、油っぽくない」など否定形も一定の特徴語になると考えられるため、収集された動詞、形容詞、形容動詞の否定形も対象とすることとした。

(2) 単語の削除

特徴語になると考えにくいものとして、各データセットの上位 10%に共通してみられた「美味しい、うまい」などの一般的な評価表現として分類[Seto 2003] される表現の他、動詞や副詞、そして食材を除く一般名詞など以下に示す単語を削除した。

美味しい, 美味, 美味しさ, うまい, 旨, 良い, 最高, 好評, 嬉しい, やさしい, 作る, 入れる, 合う, たっぷり, 味, 簡単, 味付け, 手軽, 定番, プラス

さらに、コク表現を含むつくればデータのうち、コクと対に位置し、誤分類の原因になると考えられる、表 4-2 で示した「サッパリ、あっさり」の表現が含まれるつくればデータは除くこととした。また、抽出条件としたコクを含む表現はそれ自体が特徴語となるために除いた。

コク表現に関するデータセットを作成した際と同様に、サッパリ表現を含むつくればデータのうち、サッパリ表現と逆接の助詞および接続詞と共起し、100 件以上が確認された「コク、しっかり、濃厚、こってり、クリーミー」の表現が含まれるつくればデータを除くこととした。また、抽出条件としたサッパリを含む表現を除いた。手紙や報告書などと比べて SNS などでのツイートは一文あたりの長さが短く、十分に意味が捉えられないことが知られている [Hong 2010], [Pennacchiotti 2010], [Weng 2010]。そのため、以上で示した条件によって各つくればデータを形態素単位に分割した後、一文あたりの単語数が 4 単語以上となるものから、単語数が多い順に、データセットの対象とした。なお、コクとサッパリの各データセットを同数とするため、一文あたりの単語数 4 の場合、ランダム抽出する形とした。結果、各 6,500 件のデータセットを作成し、その中から 150 件をランダム抽出によりテストデータとし、残りを訓練データとした。

4.2.5 BERT によるテキスト分類

4.2.4 項で示したデータセットを用いて、コク表現とサッパリ表現を二値としたテキスト分類を行う。本実験では、様々な自然言語処理タスクで性能が向上することが知られている BERT[Devlin 2019] による分類を行った。なお、本実験では、BERT の日本語事前学習モデルとして、東北大学の乾・鈴木研究室から公開されている、日本語版 Wikipedia をコーパスに用いて訓練されたモデル²[Tohoku University 2020]を用いた。テキスト分類の実験結果を表 4-3 に示す。なお、BERT による分類精度との比較に、代表的な分類器として用いられる、ロジスティック回帰 (LR)、ランダムフォレスト (RF) [Breiman 2001]、Linear カーネルと RBF カーネルを用いたサポートベクターマシン (SVM_linear, SVM_rbf) [Vanik 1998]に、上述のデータセットにおける特徴量としてカウントベースと TF-IDF のそれぞれを用いた分類を実施した。

表 4-3 BERT によるテキスト分類の実験結果

		accuracy	precision	recall	f1-score
	BERT	0.883	0.884	0.883	0.883
count	LR	0.863	0.865	0.863	0.863
	RF	0.817	0.820	0.817	0.816
	SVM_linear	0.840	0.842	0.840	0.840
	SVM_rbf	0.750	0.777	0.750	0.744
tf-idf	LR_linear	0.857	0.857	0.857	0.857
	RF	0.823	0.824	0.823	0.823
	SVM_linear	0.837	0.837	0.837	0.837
	SVM_rbf	0.753	0.780	0.753	0.747

テキスト分類結果、各分類器と比較して BERT が最も高い分類精度を示すことが確認された。各指標において 90%に近い高い分類精度が確認され、コクを感じる際の特

² <https://github.com/cl-tohoku/bert-japanese>

徴語の存在が示唆されたといえる。この結果に基づき、次節で分類結果を LIME に適用し、コクの特徴語の存在を検証していく。

4.2.6 LIME による特徴語の抽出

機械学習モデルの結果を解釈する手法として知られている LIME[Ribeiro 2016] を用いて、BERT によるテキスト分類に寄与する単語を抽出し、コク表現に寄与する特徴語の存在を検証する。

(1) BERT 分類モデルへの LIME 適用

LIME でテキストデータを扱う場合、通常各単語を単位として、ランダムに単語を除いたトークン列を入力として線形回帰モデルを構築し、各単語の係数から分類結果に対する、特徴語の重要度となる、各単語の寄与率を求める。一方、BERT への入力は、サブワードと呼ばれる単語をさらに分割した部分文字列を単位とし、LIME と BERT ではテキストの分割単位が一致しない。例えば、LIME でテキストデータを扱う方法をそのまま BERT による分類器に適用すると、サブワードをランダムに除いたトークン列を LIME に入力し、その結果、サブワードごとの寄与率を得る。サブワード分割は単語の意味を考慮した分解を行わないため、LIME が出力した結果の解釈が難しい。そこで、本実験では、LIME への入力は各単語を単位としつつ、LIME から BERT を呼び出す直前で、ランダムに単語を除いたトークン列を連結し、それをサブワード分割したものを BERT に入力することで、分類結果に対する各単語の寄与率を求めることとした。

(2) LIME による特徴語の抽出結果

4.2.5 項での BERT による分類結果を、前述の変更を行った、LIME に適用した結果の出力例として、コクとサッパリの正解ラベルに対する LIME の出力結果を示す。図 4-2 では、正解ラベルがコクであるものに対して、97%の確率でコクであると予測し、その予測に寄与する特徴語として「マヨネーズ」や「絶妙」といった単語が重要

度と共に可視化されている。同様に図 4-3 では、正解ラベルがサッパリであるものに対して、99%の確率でサッパリであると予測し、その予測に寄与する特徴語として「梅」が重要度と共に可視化されていることが確認できる。

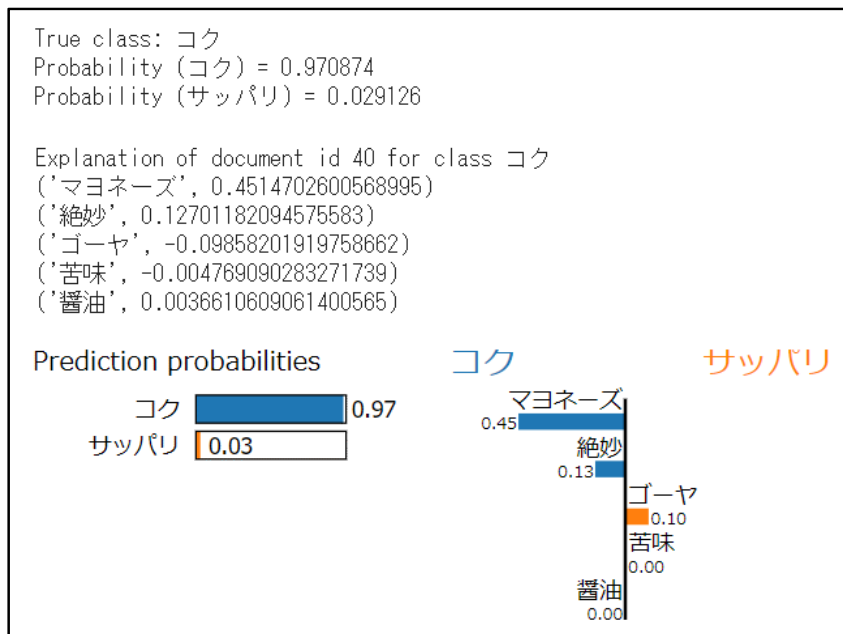


図 4-2 コクラベルの出力結果例

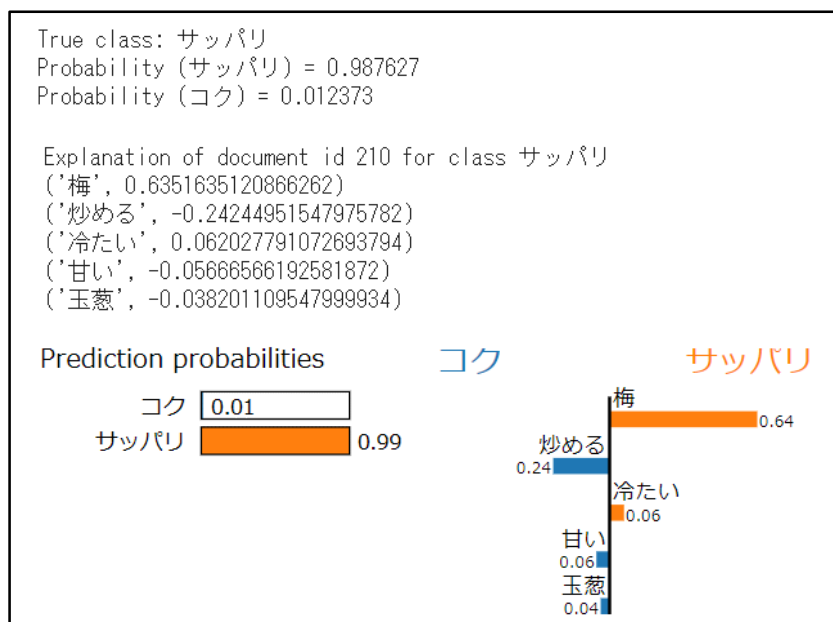


図 4-3 サッパリラベルの出力結果例

テスト用データ 300 件へ LIME を適用し、分類予測に寄与するとされた単語の集計を行った結果を表 4-4 に示す。なお、集計には、LIME の予測結果が正解ラベルと一致したものに絞り、重要度が上位の 3 単語を対象とした。また、「美味」や「よい」といった評価表現と、食材以外の一般名詞を除き、出現頻度が 5 件以上、両ラベルに出現する場合はいずれかに 80%以上に出現する条件により確認されたものとした。例えば、特徴語として確認された甘味は、コクに 13 件、サッパリに 1 件であり、コク特徴語としての出現率は 93%となるため、コクの特徴語として対象とした。

表 4-4 LIME による出力結果

単語	コク	サッパリ
マヨネーズ	30	
バター	23	
チーズ	14	
甘味	13	1
ごま油	9	
味噌	9	
まるやか	8	
牛乳	8	
たまらない	7	
うまみ	6	
カレー	5	1
米飯	5	1
オイスターソース	5	
やわらかい	5	
黒糖	5	
豚肉	5	
粉チーズ	5	

単語	コク	サッパリ
梅		36
シソ		17
酢		17
レモン		14
ポン酢		13
キュウリ		9
爽やか		9
ヘルシー		7
ワサビ		7
生姜		7
豆腐	1	6
和風		6
トマト		5
大根		5
柚子胡椒		5

表 4-4 に示すように、コク表現に寄与する特徴語として食材に関する用語（マヨネーズ、バター、チーズ、ごま油、味噌、牛乳、カレー、米飯、オイスターソース、黒糖、豚肉、粉チーズの計 12 種類）が多く確認される結果となった。また、味覚表現としては「甘味」や「うまみ」、その他「まるやか」や「やわらかい」といったテクスチャー用語も特徴語として確認された。一方、サッパリ表現に寄与する特徴語としても、食材に関する用語（梅、シソ、酢、レモン、ポン酢、キュウリ、ワサビ、生姜、豆腐、トマト、大根、柚子胡椒の計 12 種類）が多く確認された。

4.2.7 実験結果の考察

本実験では、コク表現とサッパリ表現を含むロコミ情報をデータセットとして、BERTによるテキスト分類の結果、一定の分類精度が確認され、コクには一定の特徴語が存在することが示唆された。そして、分類予測に寄与する単語集合を抽出することを目的に、BERTによる予測結果をLIMEに適用した結果、コクに寄与するものとして食材が多く確認された。コクの多くは食材によって感じたこととされ、その感じたことを明確な定義のある言葉で表現しきれていない可能性が考えられる。

4.3 特徴語の味覚要素とテクスチャー特性の解明

4.2節では、コクに関係する特徴語として、食材が多く確認された。本節の目的は、4.2節と同様にロコミ情報を用いた実験を通して、味覚要素やテクスチャー特性を解明することで、コクの特徴の定量的評価を目指すものである。実施した実験の手順と結果、そして考察を以下に示す。

4.3.1 使用データ

本実験では、4.2節と同じクックパッドデータを使用した。

(1) コク表現を含むロコミ情報の抽出

コクの特徴語である食材と共起する表現を収集するため、まず、コク表現が直接みられたレシピidをキーにして、つくればデータを抽出した。これにより、コクと表現されたつくれば以外に、コクの特徴語である食材に対してコク以外の味覚用語やテクスチャー用語で表現されたつくればが含まれる。次に、4.2.4項と同様の条件により抽出する単語を選定した。また、コクの対比表現と考えられる、サッパリ、あっさりを含むつくればデータは除く形とした。結果、583,272件がコクに関する共起用語の収集対象となった。

(2) サッパリ表現を含む口コミ情報の抽出

サッパリ表現に関しても同様の処理を行い、結果 473,962 件が対象 となった。

4.3.2 分析方法

4.2 節において、コクに関係する特徴語には食材が多いことが確認された。次に示すように、つくればでは、特徴語として確認された食材によってコクを感じたとされる表現が多くみられた。

- ・梅の酸味とマヨネーズのコクがたまりません！家族にも大好評でした★
- ・バターでコクが出てお店の味。モツ鍋好きの息子も「おいしい」連発！
- ・具たくさんで♪ごま油のコクが効いてますね！
- ・コクの決め手は味噌！納得の濃厚グラタン♪おいしくいただきました☆

その一方、次に示すように、同一レシピ上で同一食材を用いたつくればには、コクとコク以外の語を用いて表現するケースがみられた。

レシピタイトル	つくれば例
キャベツとハムの チーズスープ	ベーコンで作りました。チーズの <u>コク</u> があつて美味しかったです(^_^)ノ
	キャベツとハムとチーズの <u>旨味</u> で美味しいですね♪ランチに頂きます♪
	ベーコンで♪キャベツが甘い?チーズで <u>とろり濃厚</u> めっちゃ美味し！

特徴語として確認された食材に対して、コクと表現されるケースがある一方で、コク以外の味覚表現やテクスチャー用語で表現されるケースがあると考えられる。このことから、コクの特徴語として確認された食材と共起するコク以外の味覚表現やテクスチャー用語の特徴を評価することで、明確な定義のある言葉で表現しきれていないコクの特徴語である食材について、一定の定義を示すことが可能になると考える。

しかし、従来の共起分析を実施するだけでは分析対象となる文書上における共起関係の傾向把握に限られ、事例の一解析に留まる問題が生じる。そのため、本実験で

は、特徴語となる食材と共起する用語として「直接的な味覚用語」と「テクスチャー用語」を抽出対象にすることとした。味覚用語としては、味ことばとして収集された「塩味」「甘味」「酸味」「うま味」「苦味」「辛味」「渋味」の各味覚要素に関する用語 [Seto 2003] を用いる。テクスチャー用語としては、ISO11036TextureProfile の 3 要素と対応する形で分類された「力学的特性」「幾何学的特性」「その他の特性」に関する用語 [Hayakawa 2013a], [Hayakawa 2013b] を用いる。これらは定義のされた用語であり、特徴語となる食材と共起する用語の抽出対象として選定することが適切といえ、他の曖昧な味覚表現を分析する際にも本手法の展開が図れるものといえる。

本実験では、コク、サッパリ表現の特徴語となる食材に対し、味覚用語とテクスチャー用語が共起する頻度を集計し、両者の比較を行うことによって、曖昧な味覚表現であるコクに対する、定量的な評価を目指す。

4.3.3 実験環境

4.2 節で示した実験と同様に、本実験には Python を用い、形態素解析には MeCab と、新語・固有表現に強いとされる単語分かち書き辞書である mecab-ipadicNeologd を用いた。

4.3.4 アソシエーション分析

特徴的な食材と共起する用語を収集するために、アソシエーション分析 [Agrawal 1994] を用いた。アソシエーション分析の結果から、支持度上位 10% および lift 値 2.0 以上を対象にすることとした。結果、コクに関するデータセットは 6,837 件、サッパリに関するデータセットは 5,828 件となった。例として、コクとサッパリに寄与する特徴語の各上位 1 位であるマヨネーズと梅の結果を表 4-5 に示す。

マヨネーズと共起する用語として、コクがみられる他、「まるやか」や「こってり」、「マイルド」など、テクスチャー用語やシズルワードに相当する用語も確認できる。一方、梅と共起する用語としては、サッパリの他、「酸味」や「爽やか」などの味覚用語が確認できる。

以上の結果から、一定の味覚用語やテクスチャー用語が確認できたといえ、コクや

サッパリの特徴語として確認された食材と、コクやサッパリ以外の味覚用語やテクスチャー用語を抽出する手法として、アソシエーション分析は適当な手法といえる。

表 4-5 アソシエーション分析結果の例

(a) コクに寄与する特徴語：マヨネーズ

条件部(x)	結論部(y)	出現数(x∩y)	support	confidence	lift
マヨネーズ	合う	2347	0.00089	0.04976	2.57135
	コク	1344	0.00051	0.02849	3.66749
	醤油	757	0.00029	0.01605	6.87403
	組み合わせ	471	0.00018	0.00999	2.68583
	相性	449	0.00017	0.00952	2.40066
	まろやか	425	0.00016	0.00901	3.84370
	ボン酢	373	0.00014	0.00791	7.61136
	ごま	299	0.00011	0.00634	2.25455
	炒める	271	0.00010	0.00575	2.83041
	かける	251	0.00009	0.00532	2.03909
	コクがある	229	0.00009	0.00486	2.12772
	こってり	193	0.00007	0.00409	3.91551
	コク旨	187	0.00007	0.00396	3.01213
	マイルド	184	0.00007	0.00390	6.68442
	かつお節	167	0.00006	0.00354	2.86011

(b) サッパリに寄与する特徴語：梅

条件部(x)	結論部(y)	出現数(x∩y)	support	confidence	lift
梅	サッパリ	2892	0.00136	0.09106	2.92280
	合う	1168	0.00055	0.03678	2.20095
	マヨネーズ	901	0.00042	0.02837	6.76866
	シソ	877	0.00041	0.02761	4.90678
	酸味	699	0.00033	0.02201	5.66085
	かつお節	554	0.00026	0.01744	12.24417
	風味	493	0.00023	0.01552	2.80918
	組み合わせ	369	0.00017	0.01162	3.25658
	爽やか	327	0.00015	0.01030	2.05599
	ごま	297	0.00014	0.00935	2.81213
	相性	266	0.00012	0.00838	2.30628
	鰹	211	0.00010	0.00664	23.89989
	味噌	187	0.00009	0.00589	4.78786
	ワサビ	168	0.00008	0.00529	4.17544
	ごま油	160	0.00008	0.00504	2.35025

4.3.5 共起用語の抽出

コク表現，サッパリ表現の特徴語となる食材との共起が確認された用語について，味覚用語については要素毎に，テクスチャー用語については小分類毎に抽出した。結果を表 4-6，表 4-7，表 4-8，表 4-9 に示す。

味覚用語に関しては，コクに関係する食材と，酸味に関する用語の共起は僅かであった。また，渋味に関する用語はいずれにも確認されなかった（表 4-6）。

表 4-6 味覚に関する抽出用語

(a) コクに関係する食材と共起する用語

味覚	用語
塩味	しょっぱい，しょっぱさ，塩気，塩分，塩味
甘味	甘い，甘さ，甘辛い，甘辛さ，甘味
酸味	酸味
うま味	うまみ
辛味	スパイシー，ピリ辛，辛い，辛さ，辛味

(b) サッパリに関係する食材と共起する用語

味覚	用語
塩味	塩気，塩味
甘味	甘い，甘さ，甘酸っぱい，甘酸っぱさ，甘辛い，甘味
酸味	すっぱさ，酸っぱい，酸味
うま味	うまみ
苦味	苦味
辛味	ピリ辛，辛い，辛さ，辛味

テクスチャー用語に関しては、力学的特性に分類される用語が最も多いことから、多くの種類の共起がみられた。コクに関係する食材については、流れやすさと濃厚感に分類される用語との共起が多く確認された（表 4-7）。

表 4-7 力学的特性に関する抽出用語

(a) コクに関係する食材と共起する用語

中分類	小分類	用語
噛みごたえ	引き締まった感じの弾力	しっかり
	ほどよい噛みごたえのある弾力	もちもち
	かたさや強い噛みごたえ	かたい, 歯ごたえ
破碎	繰り返しの破碎	シャキシャキ
	切れやすさ	サクサク, サクッ
	破碎や折れやすさ	カリカリ, カリッ, パリパリ
凝集の小ささ	こわれやすい感じのもろさ	ほろほろ
	もろい感じのやわらかさ	ほくほく, ほっくり, ほっこり
	膨らんだ感じのやわらかさ	ふわっ, ふわふわ, ふんわり
変形しやすさ	やわらかさ	やわやか, やわらかさ
流動となめらかさ	流れやすさやすべり	さらさら
	流れやすさと濃厚感	クリーミー, とろける, とろっ, とろとろ, とろり, まったり, 濃厚

(b) サツパリに関係する食材と共起する用語

中分類	小分類	用語
噛みごたえ	ほどよい噛みごたえのある弾力	もちもち
	かたさや強い噛みごたえ	かたい, 歯ごたえ
	かたさと破碎	ポリポリ
破碎	繰り返しの破碎	シャキシャキ, シャキッ
	切れやすさ	サクサク
	破碎や折れやすさ	カリカリ, パリパリパリッ
凝集の小ささ	こわれやすい感じのもろさ	ほろほろ
	もろい感じのやわらかさ	ほくほく, ほっこり
	膨らんだ感じのやわらかさ	ふっくら, ふわっ, ふわふわ, ふんわり
変形しやすさ	やわらかさ	やわやか
粘りとぬめり	付着と濃厚感	ねっとり, ねばねば
流動となめらかさ	流れやすさと濃厚感	とろっ, とろとろ, とろみ, とろり

幾何学的特性に分類される用語では、なめらかさに関する用語について、コクに係る食材との共起がみられたものの、サッパリに係る食材との共起はみられなかった（表 4-8）。

表 4-8 幾何学的特性に関する抽出用語

(a) コクに係る食材と共起する用語

中分類	小分類	用語
空気	軽さと膨らみ	ふわっ, ふわふわ
粒子	小さな球	ころころ
	小さな粒や粉	ばらばら
	粉質	ほくほく, ほっくり
なめらかさと均一性	なめらかさと細かさ	クリーミー
	なめらかさとすべり	さらさら

(b) サッパリに係る食材と共起する用語

中分類	小分類	用語
空気	軽さと膨らみ	ふっくら, ふわっ, ふわふわ
粒子	小さな球	ころころ
	小さな球の集まり	ぶちぶち
	粉質	ほくほく

その他の特性では、コクに関係する食材と共起する用語として、油脂の濃厚感に関する用語との共起が顕著であった（表 4-9）。

表 4-9 その他の特性に関する抽出用語

(a) コクに関係する食材と共起する用語

中分類	小分類	用語
脂肪	油脂の濃厚感	クリーミー，こってり，まったり，濃厚
	口どけ	とろける
水分	水分の吸収または喪失	しっとり

(b) サッパリに関係する食材と共起する用語

中分類	小分類	用語
水分	にじみ出る感じの水分	ジューシー

4.3.6 使用率の算出

味覚要素，力学的特性，幾何学的特性，その他の特性に関する味覚用語およびテクスチャー用語別に，コク，サッパリに寄与する特徴語として確認された食材との共起頻度を集計し，その比較を行う。本実験では，コク，サッパリに寄与する食材自体の出現頻度の影響を考慮し，食材毎の各用語の使用率を用いた比較を行うこととした。使用率は，食材毎の味覚要素別の味覚用語・各特性の小分類別テクスチャー用語の共起頻度を分子，食材毎の味覚要素全体の味覚用語・特性毎のテクスチャー用語の共起頻度を分母として算出した。なお，アソシエーション分析結果の各味覚用語およびテクスチャー用語の共起頻度の集計の際，条件部と結論部の違いは考慮せずに集計した。味覚要素別，力学的特性，幾何学的特性，その他の特性の小分類別に使用率を算出した結果を表 4-10，表 4-11 に示す。

表 4-10 コクに寄与する食材と共起する用語の使用率

		コクに関連する食材											平均		
		マヨネーズ	バター	チーズ	ごま油	味噌	牛乳	カレー	米飯	オイスターソース	黒糖	豚肉		粉チーズ	
味覚要素	塩味	0.00%	12.72%	74.07%	0.00%	4.53%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	52.63%	12.00%	
	甘味	0.00%	87.28%	25.93%	0.00%	95.47%	100.00%	0.00%	67.64%	0.00%	100.00%	17.65%	0.00%	41.16%	
	酸味	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	28.95%	0.00%	0.00%	0.00%	0.00%	0.00%	2.41%	
	うま味	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	82.35%	47.37%	10.81%	
	辛味	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	71.05%	32.36%	0.00%	0.00%	0.00%	0.00%	8.62%	
total		0.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	-	
力学的特性	噛みごたえ	引き締まった感じの弾力	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	97.43%	0.00%	0.00%	0.00%	0.00%	8.12%
		ほどよい噛みごたえのある弾力	0.00%	0.00%	11.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.98%
		かたさや強い噛みごたえ	0.00%	12.93%	0.00%	0.00%	0.00%	0.00%	0.00%	1.28%	0.00%	0.00%	0.00%	0.00%	1.18%
	破碎	繰り返しの破碎	0.00%	0.00%	0.00%	0.00%	67.02%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	13.92%
		切れやすさ	0.00%	0.00%	8.48%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	35.82%	3.69%
	凝集の小ささ	破碎や折れやすさ	0.00%	0.00%	11.23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	9.27%
		こわれやすい感じのもろさ	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	17.54%	0.00%	0.00%	0.00%	0.00%	0.00%	1.46%
		もろい感じのやわらかさ	0.00%	72.11%	0.00%	0.00%	0.00%	15.60%	0.00%	0.00%	0.00%	40.00%	0.00%	0.00%	10.64%
	変形しやすさ	膨らんだ感じのやわらかさ	0.00%	14.97%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	60.00%	0.00%	0.00%	6.25%
		やわらかさ	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	57.89%	1.28%	0.00%	0.00%	0.00%	0.00%	4.93%
	流動となめらかさ	流れやすさやすべり	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	24.56%	0.00%	0.00%	0.00%	0.00%	0.00%	2.05%
流れやすさと濃厚感		100.00%	0.00%	68.48%	0.00%	32.98%	84.40%	0.00%	0.00%	0.00%	0.00%	0.00%	64.18%	29.17%	
total		100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	-	
幾何学的特性	空気	軽さと膨らみ	0.00%	17.19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	9.77%	
		小さな球	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.33%	
	粒子	小さな粒や粉	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	16.67%	
		粉質	0.00%	82.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	6.90%	
	なめらかさと均一性	なめらかさと細かさ	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.33%
なめらかさとすべり		0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.33%	
total		100.00%	100.00%	100.00%	0.00%	0.00%	100.00%	100.00%	100.00%	0.00%	100.00%	0.00%	0.00%	-	
その他の特性	脂肪	油脂の濃厚感	100.00%	0.00%	63.51%	0.00%	100.00%	100.00%	0.00%	100.00%	100.00%	0.00%	100.00%	100.00%	63.63%
		口どけ	0.00%	0.00%	36.49%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.04%	
	水分	水分の吸収または喪失	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	8.33%	
total		100.00%	0.00%	100.00%	0.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	-	

表 4-11 サツパリに寄与する食材と共起する用語の使用率

		サツパリに関連する食材												Mean
		梅	シソ	酢	レモン	ポン酢	キュウリ	ワサビ	生姜	豆腐	トマト	大根	柚子胡椒	
味覚要素	塩味	3.27%	100.00%	0.00%	2.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.78%
	甘味	1.29%	0.00%	15.47%	15.19%	26.70%	0.00%	60.44%	0.00%	100.00%	27.63%	0.00%	0.00%	20.56%
	酸味	93.17%	0.00%	84.53%	82.74%	73.30%	0.00%	0.00%	0.00%	0.00%	63.03%	0.00%	0.00%	33.06%
	うま味	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	9.33%	0.00%	0.00%	0.78%
	苦味	2.28%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%
	辛味	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	39.56%	100.00%	0.00%	0.00%	100.00%	100.00%	28.30%
total		100.00%	100.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	-
力学的特性	噛みごたえ	ほどよい噛みごたえのある弾力	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	28.55%	0.00%	2.38%
		かたさや強い噛みごたえ	12.67%	0.00%	0.00%	0.00%	0.00%	16.39%	0.00%	0.00%	100.00%	4.23%	0.00%	11.11%
		かたさと破碎	0.00%	0.00%	0.00%	0.00%	0.00%	21.94%	0.00%	0.00%	0.00%	3.76%	0.00%	2.14%
	破碎	繰り返しの破碎	51.13%	0.00%	0.00%	0.00%	0.00%	37.22%	0.00%	0.00%	0.00%	22.09%	0.00%	9.20%
		切れやすさ	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	1.71%	0.00%	0.00%	8.48%
	凝集の小ささ	破碎や折れやすさ	0.00%	0.00%	0.00%	0.00%	0.00%	24.44%	0.00%	0.00%	0.00%	2.00%	0.00%	2.20%
		こわれやすい感じのもろさ	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	5.52%	0.00%	8.79%
		もろい感じのやわらかさ	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	53.10%	0.00%	0.00%	4.70%	0.00%	4.82%
	変形しやすさ	膨らんだ感じのやわらかさ	0.00%	0.00%	0.00%	0.00%	93.83%	0.00%	0.00%	0.00%	90.91%	0.00%	0.00%	15.40%
		やわらかさ	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	7.38%	0.00%	0.00%	0.61%
	粘りとぬめり	付着と濃厚感	36.20%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	3.02%
流動となめらかさ		0.00%	0.00%	0.00%	0.00%	6.17%	0.00%	0.00%	46.90%	0.00%	0.00%	29.14%	0.00%	6.85%
total		100.00%	0.00%	100.00%	100.00%	100.00%	0.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	-
幾何学的特性	空気	軽さと膨らみ	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	16.67%
		小さな球	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%	0.00%	4.17%
	粒子	小さな球の集まり	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.33%
		粉質	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%	0.00%	4.17%
total		100.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	100.00%	0.00%	100.00%	0.00%	-
その他の特性	水分	にじみ出る感じの水分	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	8.33%
		total	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	-

例えば、コク表現に特徴的な食材であるバターは、味覚要素に関する用語に関して、塩味に関する用語 22 件、甘味に関する用語 151 件の計 173 件の共起頻度が確認されており、各用語の共起頻度 22 件、151 件を分子、味覚要素に関する用語の総合計 173 件を分母として、それぞれの使用率が、12.72%、87.28%、計 100%となることを示している。一方、マヨネーズは、いずれの味覚要素に関する用語とも共起がみられなかったことから、使用率は 0%を示している。

味覚要素に関して、コク表現に特徴的な食材は、サッパリ表現に特徴的な食材と比較して、甘味とうま味に関する用語の使用率が高く、酸味と辛味に関する用語の使用率が低いことが確認できる。

力学的特性に関して、コク表現に特徴的な食材は、サッパリ表現に特徴的な食材と比較して、特に、流れやすさと濃厚感に関する用語の使用率が高いことが確認できる。

幾何学的特性に関しては、コク表現に特徴的な食材と共起される用語は、一部に限られることが確認できる。

その他の特性に関して、コク表現に特徴的な食材は、サッパリ表現に特徴的な食材と比較して、油脂の濃厚感に関する用語の使用率が高いことが確認できる。

次に本使用率を用いて、コクとサッパリに特徴的な味覚用語およびテクスチャー用語の出現傾向について検定を行う。

4.3.7 使用率の検定

食材と共起する特徴的な味覚要素およびテクスチャー特性を発見するため、算出した食材毎の使用率をコクとサッパリにおける味覚要素の味覚別、力学的特性、幾何学的特性、その他の特性の小分類別で平均し、コクとサッパリの 2 群について Welch の t 検定を行った。5%有意水準によって有意に差があることが確認された味覚要素、テクスチャー特性を表 4-12 に示す。味覚要素のうち「酸味」と、その他の特性のうち「油脂の濃厚感」について、有意に差があることが確認された。また、力学的特性のうち「流れやすさと濃厚感」も有意傾向 ($p \text{ value} = 0.0861$) にあることが確認された。

表 4-12 t 検定結果

	コク	サッパリ	
	Mean	Mean	p value
酸味	2.41%	33.06%	0.0275 *
油脂の濃厚感	63.63%	0.00%	0.0008 ***

*p < .05, **p < .01, ***p < .001

4.3.8 コクの特徴語として確認された食品成分の分析

コクとサッパリの各特徴語として確認された食材に含まれる成分に注目し、統計分析を用いて、特徴的な成分を明らかにすることで、コクの特徴の定量的な評価を行う追加実験を行った。実施した実験の手順と結果、そして考察を以下に示す。

(1) 実験方法

4.2 節で確認されたコク表現に寄与する食材（マヨネーズ、バター、チーズ、ごま油、味噌、牛乳、カレー、米飯、オイスターソース、黒糖、豚肉、粉チーズの計 12 種類）と、サッパリ表現に寄与する食材（梅、シソ、酢、レモン、ポン酢、キュウリ、ワサビ、生姜、豆腐、トマト、大根、柚子胡椒の計 12 種類）を対象とし、表 4-13 で示す、文部科学省が提供する、日本食品標準成分表 2020 年版(八訂)[MEXT 2020]で確認された各食材の成分含有量を実験に用いる。

成分含有量としては、各食材について、[MEXT 2020] より確認される可食部 100g あたりの代表的な成分の含有量を用いる。実験では、水分、タンパク質、脂質、炭水化物、灰分などの一般成分、ビタミン類、アミノ酸合計量、脂肪酸総量、有機酸合計量などを対象とし、各成分の含有量を平均し、コクとサッパリの 2 群について Welch の t 検定を行った。コクとサッパリそれぞれに寄与する食材の成分含有量を表 4-14、表 4-15 に示す。

表 4-13 実験対象となる食材

ラベル	対象となる食材	日本食品標準成分表における食材
コクに関連する食材	マヨネーズ	調味料及び香辛料類/ (ドレッシング類) /マヨネーズ/卵黄型
	バター	油脂類/ (バター類) /食塩不使用バター
	チーズ	乳類/ (チーズ類) /プロセスチーズ
	ごま油	油脂類/ (植物油脂類) /ごま油
	味噌	調味料及び香辛料類/ (みそ類) /米みそ/甘みそ
	牛乳	乳類/ (液状乳類) /普通牛乳
	カレー	調味料及び香辛料類/ (ルー類) /カレールウ
	米飯	こめ [水稻軟めし] 精白米
	オイスターソース	調味料及び香辛料類/ (調味ソース類) /オイスターソース
	黒糖	砂糖及び甘味類/ (砂糖類) /黒砂糖
	豚肉	肉類/ぶた/ [大型種肉] /かた/脂身つき、生
	粉チーズ	乳類/ (チーズ類) /ナチュラルチーズ/パルメザン
	サッパリに関連する食材	梅
シソ		野菜類/しそ/葉、生
酢		調味料及び香辛料類/ (食酢類) /穀物酢
レモン		果実類/ (かんきつ類) /レモン/果汁、生
ポン酢		調味料及び香辛料類/ (調味ソース類) ぽん酢しょうゆ、市販品
キュウリ		野菜類/きゅうり/果実、生
ワサビ		調味料及び香辛料類/わさび/練り
生姜		野菜類/ (しょうが類) /しょうが/根茎、皮むき、生、おろし
豆腐		豆類/だいず/ [豆腐・油揚げ類] /絹ごし豆腐
トマト		野菜類/ (トマト類) /トマト/果実、生
大根		野菜類/ (だいこん類) /だいこん/根、皮むき、ゆで
柚子胡椒		調味料及び香辛料類/ (調味ソース類) /ゆずこしょう

表 4-14 コクに寄与する食材の成分含有量(100gあたり)

成分含有量	コクに関連する食材												平均
	マヨネーズ	バター	チーズ	ごま油	味噌	牛乳	カレー	米飯	オイスターソース	黒糖	豚肉	粉チーズ	
エネルギー(kcal)	686.0	763.0	339.0	921.0	217.0	67.0	511.0	115.0	107.0	356.0	216.0	475.0	397.8
水分(g)	19.7	15.8	45.0	0.0	42.6	87.4	3.0	71.5	61.6	4.4	65.7	15.4	36.0
たんぱく質(g)	2.5	0.5	22.7	0.0	9.7	3.3	6.5	1.8	7.7	1.7	18.5	44.0	9.9
脂質(g)	74.7	83.0	26.0	100.0	3.0	3.8	34.1	0.3	0.3	0.0	14.6	30.8	30.9
炭水化物(g)	0.6	0.2	1.3	0.0	37.9	4.8	44.7	26.4	18.3	90.3	0.2	1.9	18.9
灰分(g)	2.0	0.5	5.0	0.0	6.8	0.7	11.7	0.1	12.1	3.6	1.0	7.9	4.3
ビタミンA(μg)	117.0	1964.0	730.0	0.0	0.0	88.0	154.0	0.0	0.0	27.0	10.0	590.0	306.7
ビタミンD(μg)	0.6	0.7	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.2	0.2	0.2
ビタミンE(mg)	62.4	1.5	1.1	44.8	5.0	0.1	6.5	0.0	0.1	0.0	0.3	0.8	10.2
ビタミンK(μg)	140.0	24.0	2.0	5.0	8.0	2.0	0.0	0.0	1.0	0.0	1.0	15.0	16.5
ビタミンB1(mg)	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.7	0.1	0.1
ビタミンB2(mg)	0.1	0.0	0.4	0.0	0.1	0.2	0.1	0.0	0.1	0.1	0.2	0.7	0.2
ナイアシン(mg)	0.0	0.0	0.1	0.1	1.5	0.1	0.1	0.1	0.8	0.8	4.9	0.1	0.7
ナイアシン当量(mg)	0.5	0.1	5.0	0.1	3.5	0.9	1.0	0.4	0.8	0.9	8.0	10.0	2.6
ビタミンB6(mg)	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.7	0.3	0.1	0.1
ビタミンB12(μg)	0.4	0.1	3.2	0.0	0.1	0.3	0.0	0.0	2.0	0.0	0.4	2.5	0.8
葉酸(μg)	3.0	1.0	27.0	0.0	21.0	5.0	9.0	2.0	9.0	10.0	2.0	10.0	8.3
パントテン酸(mg)	0.4	0.1	0.1	0.0	0.0	0.6	0.4	0.2	0.1	1.4	1.2	0.5	0.4
ビオチン(μg)	7.2	0.3	2.1	0.0	5.4	1.8	4.1	0.3	0.0	33.8	0.0	0.0	4.6
ビタミンC(mg)	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	0.0	0.3
脂肪酸総量(g)	70.4	73.0	23.9	93.8	2.9	3.3	31.4	0.0	0.1	0.0	13.4	26.2	28.2
コレステロール(mg)	140.0	220.0	78.0	0.0	0.0	12.0	20.0	0.0	2.0	0.0	65.0	96.0	52.8
食物繊維総量(g)	0.0	0.0	0.0	0.0	5.6	0.0	6.4	1.1	0.2	0.0	0.0	0.0	1.1
食塩相当量(g)	2.0	0.0	2.8	0.0	6.1	0.1	10.6	0.0	11.4	0.1	0.1	3.8	3.1
アルコール(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
グルタミン酸(mg)	550.0	91.0	5000.0	0.0	1800.0	690.0	3000.0	0.0	3900.0	110.0	0.0	9900.0	2086.8
アミノ酸合計(mg)	2500.0	470.0	24000.0	0.0	9900.0	3300.0	6400.0	0.0	7200.0	800.0	0.0	48000.0	8547.5
でんぷん(g)	0.1	0.0	0.0	0.0	0.0	0.0	25.0	0.0	0.0	0.0	0.0	0.0	2.1
ぶどう糖(g)	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.6	0.0	0.0	0.1
果糖(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	1.0	0.0	0.0	0.1
ガラクトース(g)	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
しょ糖(g)	0.0	0.0	0.0	0.0	0.0	0.0	9.3	0.0	0.0	87.3	0.0	0.0	8.1
麦芽糖(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
乳糖(g)	0.0	0.6	0.0	0.0	0.0	4.4	0.2	0.0	0.0	0.0	0.0	0.0	0.4
トレハロース(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
有機酸計(g)	0.5	0.0	1.3	0.0	0.0	0.2	0.4	0.0	0.0	0.0	0.0	0.0	0.2

表 4-15 サッパリに寄与する食材の成分含有量(100gあたり)

成分含有量	サッパリに関連する食材												平均
	梅	シソ	酢	レモン	ポン酢	キュウリ	ワサビ	生姜	豆腐	トマト	大根	柚子胡椒	
エネルギー(kcal)	30.0	37.0	25.0	26.0	61.0	14.0	265.0	70.0	62.0	19.0	18.0	49.0	56.3
水分(g)	72.2	86.7	93.3	90.5	77.0	95.4	39.8	81.6	88.5	94.0	94.8	64.5	81.5
たんぱく質(g)	0.9	3.9	0.1	0.4	3.7	1.0	3.3	0.7	5.3	0.7	0.5	1.3	1.8
脂質(g)	0.7	0.1	0.0	0.2	0.0	0.1	10.3	0.8	3.5	0.1	0.1	0.8	1.4
炭水化物(g)	8.6	7.5	2.4	8.6	10.8	3.0	39.8	16.0	2.0	4.7	4.0	9.3	9.7
灰分(g)	17.6	1.7	0.0	0.3	7.6	0.5	6.8	0.9	0.7	0.5	0.5	24.1	5.1
ビタミンA	14.0	22880.0	0.0	20.0	2.0	689.0	16.0	30.0	0.0	1129.0	0.0	605.0	2115.4
ビタミンD(μg)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ビタミンE	2.1	3.9	0.0	0.1	0.0	0.3	0.0	2.1	3.4	1.1	0.0	2.4	1.3
ビタミンK(μg)	9.0	690.0	0.0	0.0	0.0	34.0	0.0	0.0	9.0	4.0	0.0	0.0	62.2
ビタミンB1(mg)	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.0	0.1
ビタミンB2(mg)	0.0	0.3	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.1
ナイアシン(mg)	0.4	1.0	0.1	0.1	0.7	0.2	0.7	0.5	0.2	0.7	0.2	0.9	0.5
ナイアシン当量(mg)	0.5	2.4	0.1	0.1	0.7	0.4	1.2	0.6	1.6	0.8	0.3	1.1	0.8
ビタミンB6(mg)	0.0	0.2	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.2	0.1
ビタミンB12(μg)	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
葉酸(μg)	0.0	110.0	0.0	19.0	17.0	25.0	0.0	5.0	12.0	22.0	33.0	13.0	21.3
パントテン酸(mg)	0.0	1.0	0.0	0.2	0.2	0.3	0.0	0.1	0.1	0.2	0.1	0.2	0.2
ピオチン(μg)	0.8	5.1	0.1	0.3	3.1	1.4	0.0	0.5	3.5	2.3	0.3	3.6	1.8
ビタミンC(mg)	0.0	26.0	0.0	50.0	0.0	14.0	0.0	1.0	0.0	15.0	9.0	2.0	9.8
脂肪酸総量(g)	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	3.1	0.1	0.0	0.0	0.3
コレステロール(mg)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
食物繊維総量(g)	3.3	7.3	0.0	0.0	0.3	1.1	0.0	7.4	0.9	1.0	1.7	6.2	2.4
食塩相当量(g)	18.2	0.0	0.0	0.0	7.8	0.0	6.1	0.0	0.0	0.0	0.0	25.2	4.8
アルコール(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
グルタミン酸(mg)	44.0	420.0	0.0	47.0	1500.0	240.0	280.0	0.0	1100.0	240.0	150.0	0.0	335.1
アミノ酸合計(mg)	630.0	3600.0	0.0	330.0	3600.0	760.0	2200.0	0.0	6000.0	570.0	430.0	0.0	1510.0
でんぷん(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0
ぶどう糖(g)	0.5	0.0	0.0	0.6	2.5	0.9	0.0	0.0	0.0	1.4	1.3	0.0	0.6
果糖(g)	0.4	0.0	0.0	0.6	2.3	1.0	0.0	0.0	0.0	1.6	1.0	0.0	0.6
ガラクトース(g)	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
しょ糖(g)	0.0	0.0	0.0	0.3	2.0	0.1	0.0	0.0	0.7	0.0	0.1	0.0	0.3
麦芽糖(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
乳糖(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
トレハロース(g)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
有機酸計(g)	4.3	0.0	0.0	6.7	1.8	0.3	0.0	0.0	0.2	0.4	0.0	0.0	1.1

(2) 成分含有量の検定

コクとサッパリの2群について Welch の t 検定を行い、5%有意水準によって有意に差があることが確認された成分を表 4-16 に示す。

表 4-16 t 検定結果

成分表	コク	サッパリ	p value
	Mean	Mean	
エネルギー(kcal)	397.750	56.333	0.0013 **
水分(g)	36.008	81.525	0.0003 ***
脂質(g)	30.883	1.392	0.0156 *
脂肪酸総量(g)	28.196	0.285	0.0137 *
コレステロール(mg)	52.750	0.000	0.0251 *

*p < .05, **p < .01, ***p < .001

食品成分のうち、「エネルギー」、「水分」、「脂質」、「脂肪酸総量」、「コレステロール」について、有意に差があることが確認された。

以上で示した通り、コクが持つ成分特徴としては、エネルギー、脂質、脂肪酸総量、コレステロールが関係していることが明らかになった。また、「ビタミン C(mg)」も有意傾向 ($p \text{ value} = 0.0534$) にあることが確認された。

4.3.9 実験結果の考察

本実験では、コクの特徴語として確認された食材に対して、共起するコク以外の味覚やテクスチャー用語をアソシエーション分析により抽出、集計した結果をもとに、食材と各用語の使用率を算出した。コク表現に特徴的な食材とサッパリ表現に特徴的な食材と共起する用語の使用率を比較した結果、コクが持つ特性として、油脂の濃厚感について、有意に差があることが確認された。油脂については、その味を甘いなどの直接的な味覚表現で表すことは難しく、油脂から感じられた味覚をコクと表現する可能性が示唆されたといえる。また、コクが持つ特性として確認された油脂の濃厚感に関連する用語としては、「クリーミー」、「こってり」、「まったり」、「濃厚」が本実験では確認されている。一般生活者にとっては、これらの表現を用いて、明確に定義のできない味覚表現であるコクと同等の感覚として表現される可能性が示唆されたといえる。

さらには、コクとサッパリの各特徴語として確認された食材の成分データを用いた比較から、コクが持つ特性として脂質などに関して統計的に有意に差があることが確認された。脂質は油脂に多く含まれるものであり、味覚やテクスチャー用語を用いたテキストマイニング手法による本実験結果は、成分データを用いた実験結果と同等の結果を示すことが確認されたといえる。

また、一連の実験手順により、コク以外にも明確な定義がない味覚表現といえる、サッパリについても一定の特徴語やサッパリ以外の味覚要素である酸味が確認されたことから、本分析手法は、明確に定義のできない味覚表現の特徴を定量的に評価する手法として有用であることが示されたといえる。

4.4 おわりに

本研究では、曖昧な味覚表現の例としてコクを取り上げ、明確に定義ができない味覚表現に対する、共通的な意図説明を目的とした定量的評価手法の提案を行った。本実験では、コクと対の表現になることが確認された、サッパリ表現との比較を行う形で一連の分析を進めた。まず、独自の変更で BERT に適用させた LIME を用いて、テキスト分類に寄与する特徴語を可視化することで、コクに関係するとされる特定の食材の存在を確認した。次に、それら特定の食材から感じられるとした、コク表現以外の定義された味覚やテクスチャー用語を抽出し、それらの共起頻度から算出した使用率を用いて、t 検定による比較を実施した。そして、最終的にコクの特性として、油脂の濃厚感が大きく寄与することを確認した。また、一般生活者は、クリーミー、こったり、まったり、濃厚といった表現を用いて、明確に定義のできない味覚表現であるコクと同等の感覚を表現している可能性が示唆された。本章で示した一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現の共通的な意図説明に向けた定量的評価を可能にしたといえる。

第5章 他者情報が味覚表現に与える影響の 定量的評価手法

5.1 はじめに

前章では、一般生活者が明確な定義なく使用していた味覚表現の共通的な意図理解のための定量的評価を行い、明確に定義ができない味覚表現例のコクの特性として、油脂の濃厚感が大きく寄与することを確認した。しかし、油脂の濃厚感が大きく寄与するとされるコクなどの明確に定義ができない味覚表現が、どのように一般生活者に広く共有され、使用されるのかといった合意形成に至る点に関して評価出来ていない。前章で取り上げたコクという表現は、明確な定義がないにもかかわらず、我々の多くは、コクという表現を多く目や耳にし、コクのある食品として、煮込み料理やラーメン、カレー、ビールなど広範囲にわたって連想することが可能なものである。そのため、明確に定義ができない味覚表現は、他者が用いた表現の影響を受けている可能性が考えられる。そこで、本章では、一般生活者にとって明確に定義ができない味覚表現は、他者情報からの影響を受けて、一般生活者の間で合意形成されていくことの傾向有無を定量的に評価する手法を提案する。

本研究では、明確に定義ができない味覚表現の例として、前章と同じくコクを取り上げる。そして、料理レシピ共有サイトに投稿されたレシピ情報上におけるコク表現の有無と、投稿されたレシピに対する口コミ情報上におけるコク表現およびその他の味覚表現との関係性を分析する。本分析手法によって、レシピ投稿者によって用いられたコク表現が、一般生活者が用いる味覚表現に与える影響評価を可能にし、明確に定義ができない味覚表現が、どのように一般生活者に広く共有され、使用されるのかといった合意形成に至る点に関する定量的評価を可能にすることが期待できる。

5.2 他者情報が味覚表現に与える影響の検証

本節では、分析手法および本実験に用いるデータセット作成に関して説明する。

5.2.1 使用データ

本研究では、4章と同じく、国立情報学研究所より、研究利用目的に限定公開されているクックパッドデータ（表 5-1）を使用する。

表 5-1 クックパッドのデータ概要

レシピ	材料	料理手順	ロコミ
1,715,595	12,724,287	8,769,210	9,449,478

5.2.2 分析方法

本実験では、クックパッドに投稿されたレシピ情報上におけるコク表現の有無を目的変数、投稿されたレシピに対するロコミ情報上におけるコク表現およびその他の味覚表現の出現頻度を用いた指標を説明変数とし、ロジスティック回帰分析を行う。ロジスティック回帰分析では、個々の説明変数をオッズ比で示すことができるため、特徴的な説明変数を捉えることが可能である。本実験では、説明レシピ情報上におけるコク表現の有無が、投稿されたレシピに対する味覚表現に与える影響の検証を行うことから、本実験において、本分析手法を用いることは適当といえる。本実験によって、明確に定義ができない味覚表現に対する、他者情報の影響有無の定量的評価を目指す。

5.2.3 実験環境

本実験では、ロコミ情報に関する前処理には Python を用い、形態素解析には MeCab と、新語・固有表現に強いとされる単語分かち書き辞書である mecab-ipadic-Neologd を用いた。ロジスティック回帰分析には R を用い、変数選択には、パッケージ MASS の関数 stepAIC を用いた。

5.2.4 データセットの作成

本実験で用いるデータセットの作成手順およびデータセットの概要を示す。本実験では、ロコミ情報上にコク表現がみられたレシピを対象に分析を行う。

(1) コク表現を含むロコミ情報の抽出

今回対象とする「コク」をキーワードとして、つくれぽの中にコクという表現が直接的に含まれるレシピ情報を抽出した。その結果、直接的にコクという表現が含まれるロコミ情報 49,342 件を得た。次に、コク表現が直接みられたレシピ id をキーにして、つくれぽデータを抽出した。これにより、コクと表現されたつくれぽ以外に、コク以外の味覚やテクスチャーなどで表現されたつくれぽが含まれる。また、前章 4.2.4 項における前処理手順と同様に、コクの対比表現と考えられる、サッパリ、あっさりを含むつくれぽデータは除く形とした。結果、583,272 件を得た。

(2) 味覚表現の抽出

本実験では、特徴的に表れやすい味覚表現に限定することとした。味覚に関する特徴語として、「直接的な味覚用語」「テクスチャー用語」「シズルワード」を選定する。味覚用語としては、味ことばとして収集された「塩味」「甘味」「酸味」「うま味」「苦味」「辛味」「渋味」の各味覚要素に関する用語[Seto 2003] を用いる。テクスチャー用語としては、ISO11036 Texture Profile の 3 要素と対応する形で分類された「力学的特性」「幾何学的特性」「その他の特性」に関する用語[Hayakawa 2013a], [Hayakawa 2013b] を用いる。シズルワードは、「味覚系」「食感系」「情報系」の 3 つの領域

に分類されたもの[Ohashi 2016], [B. M. FT Kotoba Lab 2016] を用いる。以上で対象とした味覚表現のうち、先に抽出された 583,272 件の口コミ情報の中から、出現頻度ベースで上位 85%を占める味覚表現を抽出対象とした。結果、119,719 件を得た。なお、テクスチャー用語である、「とろとろ」、「とろり」、「とろっ」などは、流れやすさと濃厚感として一つにカウントするなど、各用語は小分類の単位で一つとした。抽出した味覚表現を表 5-2 に示す。

表 5-2 抽出された味覚表現

抽出された味覚表現	口コミ上で用いられた味覚表現
甘味	甘い, 甘味, 甘辛い, 甘酸っぱい, 甘口
流れやすさと濃厚感	濃厚, とろとろ, クリーミー, とろり, とろっ, とろみ, まったり
コク	コク, コク旨, コク深い
やわらかさ	やわらかい, 軽い
膨らんだ感じのやわらかさ	ふわふわ, ふんわり, ふわっ, ふっくら, ふかふか, ほわほわ
辛味	辛い, ピリ辛, 辛味, スパイシー, 辛口, 激辛, 旨辛
もろい感じのやわらかさ	ほくほく, ほっこり, ほっくり, ほこほこ
うま味	うまみ
水分の吸収または喪失	しっとり, しなびた, 湿った
折れやすさ	カリカリ, カリッ, パリパリ, バリバリ
まろやか	まろやか
繰り返しの破碎	シャキシヤキ, シャキッ, ザクザク, シャクシヤク, しゃっきり
油脂の濃厚感	濃厚, クリーミー, こってり, 油っこい, ぎとぎと
酸味	酸っぱい, 酸味

(3) 単語頻度を用いた値の算出

抽出された味覚表現の出現頻度をレシピ毎に集計する。本実験では、該当レシピへの口コミ数によって味覚表現の出現傾向が影響されてしまうことを考慮し、各味覚表現の出現頻度を分子、全体の味覚表現数を分母にして算出した、各味覚表現の出現率を用いることとした。なお、前述の通り抽出した味覚表現以外の味覚表現はその他としてカウントし、出現率の算出の際、分母に含める形とした。また、算出された出現率が 0.1 以下の値になることが多く確認された。その場合、1 増加した場合の変化があまりにも大きくなるため、元の値のままロジスティック回帰分析を行うと、オッズ比が極端に大きくなり、その意味が解釈しにくくなることが知られている。そのた

め、算出した出現率を 10 倍した数値を説明変数として用いることとした。

(4) レシピ情報上におけるコク表現の有無

クックパッドデータで確認できるレシピ情報のうち、「レシピ名」、「レシピ説明」、「作り方」、「コツ・ポイント」、「このレシピの生い立ち」のいずれかにコク表現が直接みられたレシピを、他者情報としてコク表現が含まれるレシピとして分類した。WEB 上のクックパッドの投稿レシピ画面の概要を図 5-1 に示す。



図 5-1 クックパッドの投稿レシピ画面の概要

(5) データセットの概要

本実験では、1 レシピあたりに用いられる味覚表現の分布状況に関する評価が必要となるため、60 件以上の口コミ情報が確認されたレシピを対象とした。結果、コク情報のあるレシピ 70 種類、コク情報のないレシピ 314 種類も計 384 件、総計 75,459 件の口コミ情報によるデータセットを作成した。

5.3 ロジスティック回帰分析による検証結果

本節では、5.2 節で作成したデータセットに含まれる料理カテゴリーの出現状況や味覚表現に関する基礎統計量を確認し、本データセットを用いてロジスティック回帰分析を行う。

5.3.1 料理カテゴリーの属性比較

本データセットの料理カテゴリー別の出現分布状況を表 5-3 に示す。各料理カテゴリーの分布状況が、コク情報の有無でラベリングされた各群間で差があるのか確認するために、 χ^2 検定を行った。本検定によって分布状況に複数の差が確認された場合、今回目的とする、当該レシピの口コミに与える影響は、レシピ情報上におけるコク表現の有無ではなく、料理カテゴリーの差によるものと考えられる可能性があるためである。検定結果、麺についてのみ有意差が確認されたが、その他の料理カテゴリーでは有意差は確認されなかった。また、各群間の料理カテゴリーの出現頻度の χ^2 検定については、有意差は確認されなかった (p value = 0.360)。以上から、本データセットによる実験によって、料理カテゴリーの分布の差ではなく、今回目的とする、レシピ情報上におけるコク表現の有無による、当該レシピの口コミに与える影響の分析が可能になるといえる。

表 5-3 コク情報有無別の料理カテゴリーの出現状況

#	Category	Label - KOKU	Label - Not KOKU	p value
		(n = 70)	(n = 314)	
		with	with	
1	お肉のおかず	20	82	0.786
2	野菜のおかず	14	68	0.885
3	スープ・シチュー	5	23	1.000
4	麺	5	5	0.026 *
5	みそ汁	4	5	0.104
6	サラダ	3	25	0.415
7	お菓子	3	24	0.462
8	たまご	3	10	0.924
9	パスタ	3	9	0.812
10	ご飯もの	3	5	0.335
11	カレー	2	10	1.000
12	魚介のおかず	2	10	1.000
13	豆腐	1	14	0.400
14	パン	1	10	0.689
15	大豆のおかず	1	6	1.000
16	飲み物	0	3	0.944
17	鍋もの	0	2	1.000
18	粉もの	0	2	1.000
19	おもち	0	1	1.000

*p < .05, **p < .01, ***p < .001

5.3.2 味覚表現の属性比較

本データセットの味覚表現に関する基礎統計量を表 5-4 に示す。コク情報有無の各群から確認された味覚表現毎の関係性について独立な t 検定を用いて検証した結果、コクについては、コク情報有の場合に出現の割合が有意に多く、辛味、折れやすさ、繰り返しの破碎については、コク情報無の場合に出現の割合が有意に多いことが確認された。

表 5-4 味覚表現に関する基礎統計量

The extracted taste expression	Label - KOKU (n = 70)	Label - Not KOKU (n = 314)	p value
	Mean(SD)	Mean(SD)	
甘味	1.518(1.968)	1.935(2.265)	0.125
流れやすさと濃厚感	1.496(1.787)	1.267(1.744)	0.335
コク	1.782(1.918)	0.547(0.693)	0.000 ***
やわらかさ	0.687(1.509)	0.822(1.585)	0.508
膨らんだ感じのやわらかさ	0.441(1.449)	0.486(1.335)	0.820
辛味	0.247(0.690)	0.536(1.442)	0.014 **
もろい感じのやわらかさ	0.713(1.296)	0.426(1.071)	0.090
うま味	0.415(0.827)	0.221(0.430)	0.062
水分の吸収または喪失	0.128(0.393)	0.210(0.678)	0.181
折れやすさ	0.099(0.366)	0.285(0.870)	0.005 ***
まろやか	0.315(0.502)	0.232(0.604)	0.233
繰り返しの破碎	0.045(0.106)	0.332(1.061)	0.000 ***
油脂の濃厚感	0.834(1.184)	0.740(1.128)	0.548
酸味	0.142(0.387)	0.217(0.630)	0.207
その他	1.138(1.412)	1.746(1.915)	0.003 *

*p < .05, **p < .01, ***p < .001

5.3.3 ロジスティック回帰分析の実行

本実験では、ステップワイズ法を用いて、AIC が最も小さい説明変数の組み合わせを最も当てはまりがよいロジスティック回帰モデルとして選定し、外部情報としてのコク情報の有無が、選択された説明変数に及ぼす影響を確認する。ロジスティック回帰分析の結果を表 5-5 に示す。

ステップワイズ法による変数選択を行った結果、流れやすさと濃厚感、コク、膨らんだ感じのやわらかさ、もろい感じのやわらかさ、うま味、繰り返しの破碎が選択された。

なお、10 以上で多重共線性の影響があるといわれる VIF 値については、いずれも 10 未満であり、多重共線性の影響は無いと考えられる。

表 5-5 ロジスティック回帰分析の結果

	Estimate	SE	p value	OR	ORlow	ORhigh
(Intercept)	-2.759	0.323				
流れやすさと濃厚感	0.173	0.085	0.042 *	1.19	1.01	1.40
コク	0.883	0.146	0.000 ***	2.42	1.82	3.22
膨らんだ感じのやわらかさ	0.162	0.105	0.124	1.18	0.96	1.44
もろい感じのやわらかさ	0.204	0.122	0.095	1.23	0.96	1.56
うま味	0.488	0.229	0.033 *	1.63	1.04	2.55
繰り返しの破碎	-2.160	1.167	0.064	0.12	0.01	1.14

*p<.05, **p<.01, ***p<.001

選択された変数のうち、コクについて、正の有意な影響があることが確認された。オッズ比は 2.42 (95%CI =1.82 - 3.22) であり、選択された変数の中で最も影響のある変数であることが確認された。また、流れやすさと濃厚感とうま味についても、正の有意な影響があることが確認された。オッズ比はそれぞれ、1.19 (95%CI =1.01 - 1.40) と 1.63 (95%CI =1.04 - 2.55) であった。

5.4 料理レシピ材料の類似度を用いた比較分析

5.3 節におけるロジスティック回帰分析は、レシピ情報上におけるコク表現の有無が該当レシピへの口コミに与える影響を分析したものであった。そのため、本来口コミ上ではコクと表現されるべきレシピであったにもかかわらず、レシピ情報上にコク表現が無かった場合の、該当レシピへの口コミ情報に与える影響の分析が十分ではないことが考えられる。そこで、この問題に対処するために、料理レシピ材料の類似度を用いた比較分析を行う。

5.4.1 データセットの作成

上述の問題に対処するために、本データセットにおける、コク情報の有無によって分類された 2 群のそれぞれのレシピに用いられた材料間の類似度を算出し、最も高い

類似度が確認されたレシピ同士をペアとして抽出，そして，対応のある t 検定を用いて，味覚表現毎の比較検証を行うこととした。今回類似度の指標には，各レシピ間の材料の種類数の考慮が必要であることから，2つの集合の差集合の要素数が影響する Jaccard 係数による分析を行った。Jaccard 係数は，目安として 0.3 以上で強い共起があることが知られているが，本実験では，材料の種類数の違いが該当レシピの味に与える影響を考慮し，0.4 以上を分析対象にすることとした。結果，28 組のレシピを抽出した。抽出した 28 組のレシピを表 5-6 に示す。

表 5-6 抽出した 28 組のレシピ名・材料および Jaccard 係数

#	Label - KOKU		Label - Not KOKU		Jaccard 係数
	レシピ名	材料	レシピ名	材料	
1	サツマイモのお味噌汁	さつまいも, 味噌, 玉葱, ネギ, 油揚げ, 出汁	とろあま玉ねぎと油揚げの味噌汁	味噌, 玉葱, 酒, ネギ, みりん, 油揚げ, 出汁	0.625
2	柔らかい! 豚ヒレ肉のさっぱり照り焼き	砂糖, 片栗粉, 塩, サラダ油, 酒, 酢, みりん, 豚肉, 醤油	すっぱ甘辛の豚小間焼き	砂糖, 片栗粉, 白ごま, サラダ油, 酢, 豚肉, 醤油	0.600
3	バターなしで簡単ヘルシーMビシソワーズ	水, ジャガイモ, 塩, コショウ, 玉葱, コンソメ, 牛乳	私のヴィシソワーズ	水, バター, 生クリーム, ジャガイモ, ホワイトペッパー, 塩, 玉葱, コンソメ, 牛乳	0.600
4	大根と薄切り豚バラのさっぱりコク旨煮	砂糖, 豚バラ肉, 豆板醤, 大根, 酒, ネギ, みりん, 調味料, 出汁, 醤油	煮るだけ簡単とろっと幸せ大根	マヨネーズ, 砂糖, 豆板醤, 大根, もやし, 酒, ネギ, みりん, 醤油	0.583
5	かぼちゃのチーズクリームホットサラダ★	マヨネーズ, ブラックペッパー, カボチャ, 塩コショウ, チーズ, 牛乳	かぼちゃのマヨチーズ焼き	マヨネーズ, カボチャ, 塩コショウ, 玉葱, チーズ	0.571
6	コクウマ! 白ねぎの豚肉巻き	砂糖, 豚バラ肉, 酒, ネギ, みりん, 醤油	家族喜ぶ♪豚バラこんにゃく	砂糖, 白ごま, 豚バラ肉, こんにゃく, 出汁醤油, みりん, ネギ, 醤油	0.556
7	おばあちゃんのかぼちゃの煮物 ~黒糖使用	砂糖, カボチャ, 昆布茶, 酒, みりん, 黒糖, 醤油	ほっこり♪甘&塩が絶妙『かぼちゃの煮物』	水, 砂糖, カボチャ, 塩, 酒, みりん, 醤油	0.556
8	簡単美味じゃがいもと玉ねぎの甘辛バター煮	バター, 砂糖, ジャガイモ, マーガリン, 出汁の素, 本だし, 玉葱, みりん, 醤油	バターが決め手。玉葱人参じゃが芋の煮物。	バター, 砂糖, ジャガイモ, 玉葱, みりん, ニンジン, 出汁, 醤油	0.545
9	コクのある野菜コンソメスープ♪	水, パセリ, バター, ジャガイモ, ローリエ, ベーコン, 玉葱, キャベツ, コンソメ, オリーブオイル, ニンジン	玉ねぎとニンジンの食べるコンソメスープ	水, パセリ, バター, 玉葱, コンソメ, ニンジン	0.545
10	■旨い! 豚肉と白菜のこく旨クリーム煮■	水, オイスターソース, 白菜, 砂糖, 片栗粉, 酒, 豚肉, 牛乳, 醤油	ヘルシー☆白菜と春雨たっぷり豚肉炒め	オイスターソース, 水, 白菜, 砂糖, 豚バラ肉, 春雨, サラダ油, 酒, 鶏がらスープの素, 豚肉, 醤油	0.538
11	鱈のネギ照りマヨ焼き	マヨネーズ, 鱈, 砂糖, 片栗粉, 塩, 酒, ネギ, みりん, 醤油	甘辛こってり豆腐ステーキ	マヨネーズ, 砂糖, 片栗粉, 豆板醤, サラダ油, 酒, ネギ, みりん, 豆腐, 調味料, 醤油	0.538
12	■鶏むね&長ねぎのオイスター黒胡椒炒め■	オイスターソース, バター, 砂糖, 片栗粉, ケチャップ, ごま油, 鶏ムネ肉, 塩コショウ, 酒, ネギ, 赤唐辛子, 醤油	鶏むね肉としめじのオイスター炒め	オイスターソース, 砂糖, 片栗粉, ごま油, 鶏ムネ肉, 酒, しめじ, 醤油	0.538
13	【農家のレシピ】我が家の無水肉じゃが	バター, 砂糖, ジャガイモ, 牛肉, マーガリン, サラダ油, 玉葱, こんにゃく, ニンジン, 醤油	バターが決め手。玉葱人参じゃが芋の煮物。	バター, 砂糖, ジャガイモ, 玉葱, みりん, ニンジン, 出汁, 醤油	0.500
14	♪驚くほど甘いさつま芋黒コショウ炒め♪	塩, さつまいも, ブラックペッパー, オリーブオイル	豚肉巻きトマトステーキ	トマト, ブラックペッパー, 豚バラ肉, 塩, オリーブオイル	0.500
15	簡単☆薩摩芋と茹で卵のほっこりサラダ	マヨネーズ, さつまいも, 卵, 酢, ごま, 醤油	揚げない 簡単大学芋	さつまいも, ごま油, 酢, ごま, 黒糖, 醤油	0.500
16	おからの炊いたん (卵の花♪)	干し椎茸, 砂糖, おから, サラダ油, 玉葱, こんにゃく, ネギ, みりん, ニンジン, 出汁, 醤油	バターが決め手。玉葱人参じゃが芋の煮物。	バター, 砂糖, ジャガイモ, 玉葱, みりん, ニンジン, 出汁, 醤油	0.462
17	コクが濃厚! 『うまうま中華風焼きうどん』	マヨネーズ, 砂糖, 卵, ごま油, 青海苔, 豆板醤, 酒, うどん, ネギ, 醤油	煮るだけ簡単とろっと幸せ大根	マヨネーズ, 砂糖, 豆板醤, 大根, もやし, 酒, ネギ, みりん, 醤油	0.462
18	我が家の絶品カレー	オイスターソース, 水, ジャガイモ, サラダ油, 玉葱, 牛乳, 鶏もも肉, にんにく, ニンジン	シチューの素はもういない!!これでOK	パセリ, 水, カボチャ, ジャガイモ, 小麦粉, 玉葱, コンソメ, 鶏もも肉, 牛乳, ニンジン	0.462
19	ご飯が進む♪豚じゃがオイスター炒め♪	ピーマン, オイスターソース, 砂糖, 片栗粉, ジャガイモ, コショウ, サラダ油, 玉葱, 酒, 豚肉, 醤油	ご飯が進む! 鶏むね肉のマヨ照り焼き♪	マヨネーズ, 砂糖, 片栗粉, コショウ, 鶏ムネ肉, サラダ油, 酒, 醤油	0.462
20	食べ過ぎ注意☆鶏肉のねぎマヨボン炒め	マヨネーズ, 片栗粉, 塩コショウ, 酒, ポン酢, ネギ, 鶏肉, 醤油	■とろうま! 豚肉と茄子のマヨ生姜焼き■	ナス, マヨネーズ, 砂糖, 片栗粉, 塩コショウ, 酒, 豚肉, 醤油	0.455
21	タモさんのピーマン	ピーマン, 水, 出汁の素, サラダ油, 酒, みりん, かつお節, 醤油	さつまいもと油揚げの煮物	水, さつまいも, 砂糖, 出汁の素, 酒, みりん, 油揚げ, 醤油	0.455
22	かぼちゃの塩麴煮	水, 砂糖, カボチャ, 薄口醤油, 酒, 麴	ほっこり♪甘&塩が絶妙『かぼちゃの煮物』	水, 砂糖, カボチャ, 塩, 酒, みりん, 醤油	0.444
23	むね肉のくせに...ウマウマ!!	マヨネーズ, 小麦粉, 鶏ムネ肉, 酢, 醤油	ピリリと辛い鶏むね肉のマヨソテ♪	マヨネーズ, おろしにんにく, 小麦粉, 豆板醤, 鶏ムネ肉	0.429
24	熱々ふわふわ♪豆腐&卵の納豆グラタン	マヨネーズ, 卵, 納豆, チーズ, 豆腐, 醤油	簡単★豆腐グラタン	マヨネーズ, 豆腐, チーズ, 味噌	0.429
25	ブロッコリーとゆで卵のオイマヨサラダ	マヨネーズ, オイスターソース, ブロッコリー, 塩コショウ, ゆで卵	ブロッコリーのサクサク! マヨパン粉焼き	マヨネーズ, ブロッコリー, 塩コショウ, パン粉, ソーセージ	0.429
26	ホクホクじゃがいもといんげんの炒め煮	砂糖, ジャガイモ, 醤油, ごま油, 塩, 出汁, イングテン	甘辛♪ 粉ふきいも	砂糖, ジャガイモ, 醤油	0.429
27	キャベツのゴマみそ汁	キャベツ, 白ごま, ごま油	母の味♪キャベツの味噌汁	ごま油, キャベツ, 出汁, 味噌	0.400
28	キャベツと鶏胸肉炒め♪レモンバターソース	バター, 片栗粉, 鶏ムネ肉, キャベツ, 酒, 油, レモン汁, 醤油	ごぼうの甘辛バター醤油焼き	バター, 砂糖, 片栗粉, 酒, ゴボウ, 醤油	0.400

5.4.2 比較分析の実行

本データセットにおける検証結果として、5%有意水準によって有意に差があることが確認された味覚表現を表 5-7 に示す。分析結果、コクについてのみ、コク情報有の場合に出現の割合が有意に多いことが確認された。

表 5-7 t 検定結果

	Label - KOKU	Label - Not KOKU	
	平均	平均	p 値
コク	1.986	0.914	0.034 *

*p < .05

5.4.3 実験結果の考察

ロジスティック回帰分析および材料類似度を用いた比較分析の結果、他者情報としてコクに関する情報がレシピ上に用いられる場合、該当レシピへの口コミ上において、コクという味覚表現が有意に用いられる傾向にあることが確認された。一連の本分析手法によって、明確に定義ができない味覚表現は、他者の影響を受ける可能性があることが定量的に確認されたといえる。

また、うま味については、ロジスティック回帰分析において、一定の有意な正の影響があることが確認された。他者情報の影響を受けない一般生活者にとっては、一部がコクと表現するものに対して、うま味を用いて表現する傾向にある可能性や、コクと聞いてうま味を想起させる傾向にあることの可能性が考えられる。うま味は、基本味として用いられる場合のほか、おいしさ自体を表す場合にも用いられることもあるため、多くの一般生活者にとっては、「うまみ」という味の正確な認識や表現がし難い状況にあるものといえる。今後、本分析手法などを用いた、一般的な生活者にとってのうまみの解明が必要であるといえる。

5.5 おわりに

本章では、明確に定義ができない味覚表現が、どのように共有され使用されるのかといった合意形成に至る点に関して評価を行うため、料理レシピ共有サイトの口コミ情報を対象に、明確に定義ができない味覚表現は、他者情報からの影響を受けて合意形成されていくことの傾向有無を示す、定量的評価手法を提案した。曖昧な味覚表現の例としてコクを取り上げ、ロジスティック回帰分析および材料の類似度を用いた比較分析を行った。分析結果、レシピ投稿者による他者情報としてコクに関する情報がレシピ上に用いられる場合、該当レシピの口コミ上において、コクという味覚表現が有意に用いられる傾向にあることが確認された。本章で示した一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現は、他者情報の影響を受けて、一般生活者の間で合意形成されていくことを示す定量的評価を可能にしたといえる。

第6章 結論

本論文は、一般生活者の味覚によって判断され、用いられる、明確に定義ができない味覚表現の共通的な意図理解を目的とした、定量的評価手法の提案を行ったものである。

近年、消費者をとりまく経済・社会的環境の変化に伴い、食の外部化、簡便化、高級化など食生活の変化が進み、消費者の食へのニーズはますます多様化、高度化しているといわれている。そうした環境下、飲食業界にとっては、消費者へのPR戦略の一つとして、商品の魅力をオノマトペやシズルワードを用いて表現することで、商品の差別化や販売促進が図られている。また、そのようにPRされた商品である飲食物のおいしさが、消費者にどのように感じられたのか、といったことを、消費者間のSNSや口コミ投稿サイトなどによる口コミ情報の分析・評価により、商品開発に繋げる動きがみられ、おいしさを表す表現に注目が集まっている。

一般的に、味覚は、味、香り、食感、色など味わう対象から五感を通じて感じられるものや、食の経験や味わう際の環境、各種媒体などから収集される食に関する情報など様々な要因が複雑に作用して、形成されると考えられている。そのため、人間の持つ感性の一つとなる味覚には個人差があるとされる。その一方で、多くの一般生活者にとっては、飲食物は、個人または家族や友人、職場の同僚などの身近な存在と味わう機会が主であり、その飲食物から感じられた具体的、詳細なおいしさを他者に積極的に伝える必要性がなく、おいしいやうまい、まずいといった一般的な表現を用いて表現することが多い。そのため、飲食物から感じたおいしさを具体的、詳細に表現する機会は少なく、実際に、飲食物の味を表現する場合、その味をおいしいやまずいといった一般的な表現以外を用いて表現することが難しいといった事態が生じる。例えば、日本酒は、糖分が多く含まれたものが甘口、糖分が少ないものが辛口といわれる。しかし、甘口は甘みを感じるものとして容易に他人と共有できる一方で、辛口は辛いと感じる成分が含まれているわけではなく、その味を何の味であるか表現しづらいものといえる。さらには、説明された味の情報として多く用いられる、キレやフルーティ、コク、まろやかなどといった表現は、明確な定義がないものである。また、たとえおいしいやまずいといった一般的な表現以外を用いて、おいしさを表現するこ

とがあつたとしても、同一の飲食物の味を表現する際には、異なる味覚に関する表現が用いられるといった事態が生じることがある。例えば、味覚に関して「甘み」や「辛み」「コク」などといった表現を用いる。この中で、「甘み」が砂糖の味であり「辛み」が唐辛子の味であるように容易に他人と共有できる言葉である一方、その逆に「コク」のように何の味であるかが定義しづらい言葉もある。これはコクに対して明確な定義がないためである。しかし、コクに類似する味覚表現として、「深みのある味」や「濃厚な味」、「複雑な味」などを想起することができる。また、コクのある食品として連想されるものとしては、煮込み料理、ラーメン、カレーやビールなど広範囲にわたり、これらは、「深みのある味」や「濃厚な味」、「複雑な味」としても連想が可能といえる。コクのニュアンスや感覚などは、曖昧ではあるにもかかわらず、おおよそ我々の多くに共有されているものといえる。

こうした背景から、飲食業界などでは自社商品の味覚的な特徴を、多くの消費者へ共有させる必要があるが、一般生活者各人が各々の味覚によって判断して用いられる味覚表現は、曖昧性を含んでいるために、客観的な評価を困難にするという問題が生じる。そこで、本研究では、一般生活者各人の味覚によって判断された味覚に関する表現が記載された、感性の情報源といえる口コミ情報を対象に、テキストマイニングによって、各人の味覚によって判断され用いられる、明確に定義ができない味覚表現の共通的な意図理解を目的とした、一連の定量的評価手法の提案を行った。これまで生化学分野を中心に、味覚を明らかにすることは試みられてきたが、これらは化学的な分析による手法であった。本論文で示した一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現の共通的な意図理解に向けた定量的評価を可能にした点は本論文の重要な成果である。

第3章では、日本酒を味わった際の明確に定義ができない味覚表現の共通的な意図理解を目的とした定量的評価手法の提案を行った。重回帰分析を用いた日本酒の味覚表現の検証の結果、辛口とされる日本酒の場合、辛口の代わりに、キレやサッパリといった用語で表現される可能性があることが確認された。また、甘口とされる日本酒の場合、なめらかやフルーティ、やわらかといった表現が用いられることが確認された。その他、日本酒の味に濃醇さが感じられる場合、甘味やコク、まろやかと表現される傾向にあることが示唆された。一連のテキストマイニングによる分析手法によっ

て、「日本酒度」「酸度」「甘辛度」「濃淡度」と、日本酒に関する各種味覚表現における関係性について、定量的な評価を可能にした。

第4章では、曖昧な味覚表現の例としてコクを取り上げ、明確に定義ができない味覚表現に対する、共通的な意図解明を目的とした定量的評価手法の提案を行った。本実験では、コクと対の表現になることが確認された、サッパリ表現との比較を行う形で一連の分析を進めた。まず、独自の変更でBERTに適用させたLIMEを用いて、テキスト分類に寄与する特徴語を可視化することで、コクに関係するとされる特定の食材の存在を確認した。次に、それら特定の食材から感じられるとした、コク表現以外の定義された味覚やテクスチャー用語を抽出し、それらの共起頻度から算出した使用率を用いて、t検定による比較を実施した。そして、最終的にコクの特性として、油脂の濃厚感が大きく寄与することを確認した。また、一般生活者は、クリーミー、こったり、まったり、濃厚といった表現を用いて、明確に定義ができない味覚表現であるコクと同等の感覚を表現している可能性が示唆された。一連のテキストマイニングによる分析手法によって、一般生活者が明確な定義なく使用していた味覚表現の意図理解のための定量的評価を可能にした。

第5章では、明確に定義ができない味覚表現が、どのように共有され使用されるのかといった合意形成に至る点に関して評価を行うため、料理レシピ共有サイトの口コミ情報を対象に、明確に定義ができない味覚表現は、他者情報からの影響を受けて、一般生活者の間で合意形成されていくことの傾向有無を示す、定量的評価手法を提案した。前章と同様に曖昧な味覚表現の例としてコクを取り上げ、ロジスティック回帰分析および材料の類似度を用いた比較分析を行った。分析結果、他者情報としてコクに関する情報がレシピ上に用いられる場合、該当レシピの口コミ上において、コクという味覚表現が有意に用いられる傾向にあることが確認された。一連のテキストマイニングによる分析手法によって、明確に定義ができない味覚表現は、他者の影響を受けて一般生活者の間で合意形成されていくことを示す可能性があることを示す定量的評価を可能にした。

今後の取り組みとしては、様々な飲食物や明確に定義ができない味覚表現を対象に、各章で提案した一連のテキストマイニングによる本分析手法の適用を図り、更なる明確に定義ができない味覚表現の共通的な意図理解について取り組んでいく必要がある。3章では、日本酒の日本酒度、酸度、甘辛度、濃淡度の各指標と、日本酒に関

する各種味覚表現における関係性について定量的評価を行い、日本酒に関する表現の特徴を明らかにしたが、日本酒以外の飲食物にも同手法の展開が望まれる。例えば、ワインでは、ワイン中に内在する糖分である残糖分によって甘口や辛口に分類されるため、今回3章において提案した手法を用いて、ワインにおける甘辛度とワインに関する味覚表現の関係性の解明などが期待できる。また、ビールでは苦味に関する指標であるIBU (International Bitterness Units : 国際苦味単位) など、その他飲食物に対しては、食品分量やECサイトによる商品への5段階評価、または飲食物の価格などの数値が、今回提案手法に適用させることが可能と考えられる。これら指標と、各種飲食物に関する味覚表現との関係性について分析することで、評価対象を拡大していくことが重要な課題といえる。4章では、明確に定義ができない味覚表現としてコクを取り上げ、コクに関する特徴を明らかにしたが、その他、まろやか、マイルド、フルーティ、軽さ、重さ、キレ、スッキリなど多くの明確に定義のできない味覚表現に対して、4章での提案手法を用いた同様の評価が望まれる。また、今回使用した国立情報学研究所より、研究利用目的に限定公開されているクックパッドデータは、2014年9月30日までにクックパッドに掲載されたものである。味覚は、時代によってトレンドがあり、変化していくものであるといわれており、時間軸の観点を取り入れた分析が必要といえる。最新の時点におけるデータを用いて同手法を適用した分析結果と、今回4章における分析結果との比較など、一般生活者における明確に定義ができない味覚表現に関する時系列的な変化による特徴の評価は、今後の重要な課題である。5章では、4章と同様に、明確に定義ができない味覚表現としてコクを取り上げ、コクは他者情報の影響を受けて一般生活者の間で合意形成される傾向にあることを定量的に明らかにしたが、上で述べたことと同様に、その他、まろやか、マイルド、フルーティ、軽さ、重さ、キレ、スッキリなど多くの明確に定義のできない味覚表現に対して、5章での提案手法を用いた同様の評価が望まれる。また、時間軸の観点を取り入れ、明確に定義ができない味覚表現が、他者情報の影響を受けながら、一般生活者の間で合意形成されていく傾向などの時系列的な変化などの分析についても、今後の重要な課題といえる。そして、これらの分析を通じて、個人差のある味覚の定量的評価による、各人が持つ味覚の嗜好性を反映させた、最適な料理レシピの提案や商品開発への応用などへの実務的な展開を図っていくことが求められる。

謝辞

本論文は著者が、筑波大学大学院 理工情報生命学術院 システム情報工学研究群 リスク・レジリエンス学位プログラムに在籍中の研究成果をまとめたものである。

本研究を遂行するにあたり、国立情報学研究所の IDR データセット提供サービスにより楽天株式会社から提供を受けた「楽天データセット」およびクックパッド株式会社が国立情報学研究所に提供した「クックパッドデータ」を利用しました。この場を借りて御礼申し上げます。

研究を進めるにあたっては、同学位プログラム教授 津田和彦先生には研究方法、進め方など全ての段階において多大なるご指導をいただいたおかげで本博士論文としてまとめることができました。修士時代から計 5 年間の長きにわたりご指導頂き、心より深謝申し上げます。また副査として同学位プログラム教授 倉橋節也先生、同学位プログラム教授 木野泰伸先生、同学位プログラム教授 古川宏先生、帝京大学教授 藤田昌克先生には予備審査や最終審査を通じて、様々な有益な指摘や助言を戴きました。深く御礼申し上げます。

また、津田研究室 OB の原田氏をはじめ、研究室 OB・OG およびゼミ生各位からは、研究手法や論文の作成方法などについてアドバイスをいただくとともに、ゼミ時の闊達な助言によって考察を深めることができました。誠にありがとうございました。

最後に、仕事と研究の両立に理解を示してくれた家族（妻&娘）に心より感謝申し上げます。

参考文献

[Agrawal 1994]

R. Agrawal and R.Srikant: “Fast algorithm for mining association rules”, Proc. 20th Int. Conf. Very Large Data Bases, VLDB, Vol. 1215, pp.487–499 (1994)

[Akiguchi 2018]

秋口いくみ, 王 元元, 河合由起子, 角谷和俊: 「料理レシピ動画の時間特性抽出による難易度判定」, 第10回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2018), C4-4 (2018)

[ASBC 1992]

American Society of the Brewing Chemists, Sensory analysis: “In Methods of Analysis of the American Society of Brewing Chemists”, American Society of the Brewing Chemists, St Paul (1992)

[Bahdanau 2014]

D. Bahdanau, K. Cho and Y. Bengio: “Neural Machine Translation by Jointly Learning to Align and Translate”, arXiv preprint arXiv:1409.0473. (2014)

[B. M. FT Kotoba Lab 2016]

B.M.FT ことばラボ編著: 「ふわとろ SIZZLE WORD 「おいしい」 ことばの使い方」, B.M.FT 出版部 (2016)

[Breiman 2001]

Leo Breiman: “Random Forests”, Machine Learning October 2001, Vol. 45, Issue1, pp.5-32 (2001)

[Chan 2013]

K. Chan, E. Tong, D. Tan and A. Koh: “What do love and jealousy taste like?”, *Emotion*, Vol.13, No.6, pp.1142-1149 (2013)

[Devlin 2019]

J. Devlin, M. Chang, K. Lee, and K. Toutanova: “BERT: Pre-training of deep bidirectional transformers for language understanding”, In *NAACL-2019*, pp. 4171–4186 (2019)

[Elman 1990]

J. L. Elman: “Finding structure in time”, *Cognitive science*, Vol.14, No.2, pp.179-211 (1990)

[Eskin 2011]

K. Eskine, N. Kacvnik and J. Prinz: “A bad taste in the mouth: Gustatory disgust influences moral judgement”, *Psychological science*, Vol. 22, No.3, pp.295-299 (2011)

[Fukumoto 2016]

J. Fukumoto: “Analysis of Taste Expressions of Japanese Sake”, *The 30th Annual Conference of the Japanese Society for Artificial Intelligence* (2016) (in Japanese)

[Fukumoto 2017]

J. Fukumoto: “Analysis of Metaphorical Taste Expressions of Japanese Sake”, *The 31st Annual Conference of the Japanese Society for Artificial Intelligence* (2017) (in Japanese)

[Fukumoto 2018]

J. Fukumoto: “Analysis of Taste Expressions Similarity of Japanese Sake”, *The 32nd Annual Conference of the Japanese Society for Artificial Intelligence* (2018) (in Japanese)

[Fukushima 2015]

H. Fukushima and C. Nagano: “Analyzing Semantic Functions of Sound Symbolic Words of Taste”, The 29th Annual Conference of the Japanese Society for Artificial Intelligence (2015) (in Japanese)

[Fukushima 2016]

H. Fukushima and S. Tanaka: “The Roles of Sound Symbolisms in the Tasting Descriptions”, Journal of Japanese Society for Artificial Intelligence, Vol.31, No.6, pp.1-8 (2016) (in Japanese)

[Funakubo 2016]

K. Funakubo, F. Kobayashi, R. Uno, K. Shinohara and S. Kotake: “Consistency between "crispy" and "edible" textures in the outer packaging”, Japan Society of Bioscience, Biotechnology and Agrochemistry Kanto (2016) (in Japanese)

[Furumoto 2002]

古本健太, 難波英嗣, 角谷和俊: 「類似レシピの手順分析 による料理アドバイスの抽出方式」, 第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2016), P3-1 (2016)

[Furusho 2002]

古庄 重樹: 「ビール官能評価法について」, 日本醸造協会誌, Vol.97, No.3, pp.178-187 (2002)

[Gawel 2014]

R. Gawel, A.Schulkin, P. A. Smith and E. J. Waters: “Taste and textural characters of mixtures of caftaric acid and Grape Reaction Product in model wine”, Australian journal of grape and wine research, Vol. 20, No.1, pp. 25-30 (2014)

[Goto 2020]

後藤奈美：「ワインにおけるコク　ーボディ感ー」，コク研究会シンポジウム 2020 講演要旨集，WEB シンポジウム，2020-08-25，コク研究会（2020）

[Harashima 2018]

J. Harashima and Y. Yamada: “Two-step validation in character-based ingredient normalization”, Proc. 10th Workshop on Multimedia for Cooking and Eating Activities (CEA 2018), pp.29-32 (2018)

[Harashima 2019]

原島純：「レシピサービスと情報処理」，人知能学会論文誌，Vol.34, No.1, pp.3-8 (2019)

[Harris 1954]

Z. S. Harris: “Distributional Structure”, WORD, 10:2-3: pp.146-162 (1954)

[Hayakawa 1999]

早川文代，岩政由布子，畑江敬子，島田淳子：「食感覚の擬音語・擬態語の収集と選定」，日本家政学会誌，Vol.50, No.5, pp.481-490 (1999)

[Hayakawa 2005]

F. Hayakawa, K. Ioku, S. Akuzawa, M. Saito, K. Nishinari, Y. Yamano and K. Kohyama: “Collection of Japanese Texture Terms”, Nippon Shokuhin Kagaku Kaishi, Vol. 52, No.8, pp.337-346 (2005)

早川文代，井奥加奈，阿久澤さゆり，齊藤昌義，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，Vol.52, No.8, pp.337-346 (2005)

[Hayakawa 2013a]

F. Hayakawa, Y. Kazami, K. Nishinari, K. Ioku, S. Akuzawa, Y. Yamano, and Y.

Baba: “Classification of Japanese Texture Terms”, *Journal of Texture Studies*, Vol. 44, pp.140-159 (2013)

[Hayakawa 2013b]

F. Hayakawa: “Development of Japanese Texture Lexicon for Food Texture Studies”, *Nippon Shokuhin Kagaku Kaishi*, Vol.60, No.7, pp.311-322 (2013)

早川文代：「日本語テクスチャー用語の体系化と官能評価への利用」, *日本食品科学工学会誌*, Vol.60, No.7, pp.311-322 (2013)

[Hirakawa 2017]

平川芽依, 牛尼剛聡, 角谷和俊：「料理画像の色情報を用いたレシピ選別支援」, 第9回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2017), P8-3 (2017)

[Hochreiter 1995]

S. Hochreiter and J. Schmidhuber: “Long Short Term Memory”, Technical Report FKI-207-95, Fakultat fur Informatik, Technische Universitat Munchen (1995)

[Hochreiter 1997]

S. Hochreiter and J. Schmidhuber: “Long Short-Term Memory”, *Neural Computation*, Vol.9, No.8, pp.1735–1780 (1997)

[Hong 2010]

L. Hong and B.D. Davison: “Empirical study of topic modeling in Twitter”, in Proc, SOMA’10, pp.80-88, ACM (2010)

[Iori 2000]

庵功雄, 高梨信乃, 中西久実子, 山田敏弘：「初級を教える人のための日本語文法ハンドブック」, スリーエーネットワーク (2000)

[Itoen 2019]

株式会社伊藤園：「レポート 02 お茶と出汁，うま味の相乗効果～和食の味わいを引き立てる “お茶のうま味成分” ～」，

<<https://www.itoen.co.jp/company/research/report02/>>， Accessed December 17, 2019.

[Jones 1972]

K. S. Jones: “A statistical interpretation of term specificity and its application in retrieval”, *Journal of Documentation*, Vol.28, No.1, pp.11-21 (1972)

[Kagawa 2022]

香川璃奈，原悠輔，姜志勲，山肩 洋子：「付加的な説明文の生成と提示による料理レシピ執筆支援手法の提案と評価」，*電子情報通信学会論文誌 D*, No.12, pp.736-746

(2022)

[Kanauchi 2016]

金内 萌，難波英嗣，角谷和俊：「投稿型レシピサイトにおけるレビュー情報に基づく料理タイトル自動生成」，第 8 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2016)，P3-2 (2016)

[Kaneda 2014]

H. Kaneda: “Searching a Taste of the Beer”, *Journal of the Brewing Society of Japan*, Vol.109, No.6, pp.417-425 (2014) (in Japanese)

金田弘挙：「ビールの美味しさを求めて」，*日本醸造学会誌*，Vol.109, No.6, pp.417-425 (2014)

[Kanno 2018]

Y. Kanno, T. Minetoki, T. Bogaki T. Onodera and K. Toko: “Visualization of Flavor of Sake by Sensory Evaluation and Statistical Method”, *Sensors and Materials*, Vol.30, No.5, pp.1197-1206 (2018)

[Kanno 2019]

Y. Kanno, T. Onodera, T. Kamiya and K. Toko: “Development of Brewed Beverage Search System and Program”, *Sensors and Materials*, Vol.31, No.7, pp.2405-2413 (2019)

[Kanno 2020]

Y. Kanno, M. Hirata, K. Mitani, S. Shirota, T. Onodera and K. Toko: “Visualization of Flavor of Sake Using Taste Sensor and Gas Chromatography-Mass Spectrometry”, *IEEJ Transactions on Sensors and Micromachines*, Vol.140, No.11, pp.294-304 (2020)

[Kato 2015a]

A. Kato, Y. Fukuzawa and T. Mori: “Serendipitous recommendation of restaurant using sense related onomatopoeia”, *Journal of Japanese Society for Artificial Intelligence*, Vol.30, No.1, pp.216-228 (2015) (in Japanese)

加藤亜由美, 深澤佑介, 森武俊: 「五感と関連するオノマトペを用いた意外性の高い飲食店推薦」, *人口知能学会論文誌*, Vol.30, No.1, pp.216-228 (2015)

[Kato 2015b]

加藤大介, 宮部真衣, 荒牧英治, 灘本明代: 「インターネット上のメディア毎のおいしさ表現比較」, *DEIM Forum* (2015)

[Kin 2017]

金明哲: 「Rによるデータサイエンス 第2版」, 森北出版 (2018)

[Kin 2018]

金明哲: 「テキストアナリティクス」, 共立出版 (2018)

[Kiyomaru 2018]

清丸寛一，黒橋禎夫，遠藤充，山上勝義：「料理レシピとクラウドソーシングに基づく基本料理知識ベースの構築」，言語処理学会第24回年次大会発表論文集，pp.662-665 (2018)

[Kuroda 2018]

黒田 素央：「コク増強因子の一つである kokumi 物質の受容機構と官能特性」，月刊フードケミカル，Vol.403, No.11, pp.38-42 (2018)

[Kuroda 2021]

M. Kuroda. (2021) “Perceptive mechanism and sensory characteristics of kokumi Substances: Effect on the sensory characteristics of low-fat foods”, Journal of Lipid Nutrition, Vol.30, No.11, pp.38-42. (in Japanese)

[Kurohashi 2016]

黒橋禎夫，柴田知秀：「自然言語処理概論」，サイエンス社 (2016)

[Kusumi 2004]

楠見孝：「味覚のメタファー表現への認知的アプローチ」，日本言語学会第127回大会予稿集，pp.9-14 (2004)

[Lan 2020]

Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma and R. Soricut: “ALBERT: A Lite BERT for Self-supervised Learning of Language Representations”, Proc. ICLR 2020 (2020)

[Liu 2019]

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer and V. Stoyanov: “RoBERTa: A Robustly Optimized BERT Pretraining Approach”, arXiv preprint arXiv:1907.11692. (2019)

[Luong 2015]

M. Luong, H. Pham and C. D. Manning: “Effective Approaches to Attention-based Neural Machine Translation”, arXiv preprint arXiv: 1508.04025. (2015)

[Maeta 2015]

H. Maeta, T. Sasada and S. Mori: “A Framework for Procedural Text Understanding”, in Proc. of IWPT15, pp.50-60 (2015)

[MEXT 2020]

文部科学省 : 「日本食品標準成分表 2020 年版 (八訂)」 (2020),
<https://www.mext.go.jp/a_menu/syokuhinseibun/mext_01110.html>, Accessed
February 23, 2021.

[Mikolov 2013a]

T. Mikolov, K. Chen, G. Corrado and J. Dean: “Efficient estimation of word representations in vector space”, arXiv preprint arXiv:1301.3781. (2013)

[Mikolov 2013b]

T. Mikolov, I. Sutskever, K. Chen, G. Corrado and J. Dean: “Distributed representations of words and phrases and their compositionality”, In Advances in Neural Information Processing Systems. (2013)

[Mori 2013]

森信介, 山肩洋子, 笹田鉄郎, 前田浩邦 : 「レシピテキストのためのフローグラフの定義」, 自然言語処理研究発表会, Vol. 2013-NL-214, No.13, pp.1-7 (2013)

[Muto 2003]

武藤彩加 : 「ことばは味を超える 美味しい表現の探求」, 海鳴社, pp.241-300 (2003)

[Nanba 2013]

H. Nanba, Y. Doi, M. Tsujita, T. Takezawa and K. Sumiya “Summarization of Multiple Cooking Recipes”, IEICE Technical Report, Vol. 113, No. 338, NLC2013-41, pp.39-44 (2013) (in Japanese)

難波英嗣，土居洋子，辻田美穂，竹澤寿幸，角谷和俊：「複数料理レシピの自動要約」，電子情報通信学会，言語理解とコミュニケーション研究会，Vol.113, No.338, NLC2013-41, pp.39-44 (2013)

[Nanba 2014]

H. Nanba, Y. Doi, T. Takezawa, K. Sumiya and M. Tsujita: “Construction of a cooking ontology from cooking recipes and patents”, in Proc, Workshop on Smart Technology, for Cooking and Eating Activities, CEA2014, pp.507-516 (2014)

[National Tax Agency 2018]

National Tax Agency. (2018) “National Survey of Alcoholic Beverages” (2018) < <https://www.nta.go.jp/taxes/sake/shiori-gaikyo/seibun/2019/pdf/001.pdf> > Accessed April 1, 2021

[NICT 2020]

国立研究開発法人 情報通信研究機構 (National Institute of Information and Communications Technology (NICT)) : “NICT BERT 日本語 Pre-trained モデル” (2020), <<https://alaginrc.nict.go.jp/nict-bert/index.html>> Accessed December 31, 2022.

[Nishio 2014]

西尾泰和：「word2vec による自然言語処理」，オライリージャパン (2014)

[Nishimura 2018a]

西村 敏英：「コクとコク味は違うー正しい使い方と科学的根拠ー」，月刊フードケミカル， Vol.403, No.11, pp.21-27 (2018)

[Nishimura 2018b]

西村敏英：「食肉のコクとその形成・増強メカニズム」，食肉の科学, Vol.56, No.2, pp.9-15 (2018)

[Nishimura 2019]

西村敏英：「食品のコクとおいしさへの活用」，醤油の研究と技術, Vol.45, No.3, pp.161-168 (2019)

[Nishimura 2021a]

西村敏英：「コク」と「とろみ」を科学する」，全国発酵乳乳酸菌飲料協会機関紙乳酸菌ニュース, No.511 (2021)

[Nishimura 2021b]

西村敏英，黒田素央編著：「食品のコクとは何か おいしさを引き出すコクの科学」，恒星社厚生閣 (2021)

[Nishimura 2022]

西村敏英：「食品のコクの定義と見える化への挑戦」，コク研究会シンポジウム 2022 講演要旨集，女子栄養大学駒込キャンパス，2022-08-29，コク研究会 (2022)

[Ohashi 2016]

M. Ohashi: “The current situation of sizzle words”, The 30th Annual Conference of the Japanese Society for Artificial Intelligence (2016) (in Japanese)

大橋正房：「シズルワードの現状と変化」，The 30th Annual Conference of the Japanese Society for Artificial Intelligence (2016)

[Okumura 2010]

奥村学：「自然言語処理の基礎」，コロナ社 (2010)

[Otsuka 2015]

H. Otsuka, M. Suwa and K. Yamaguchi: “Studies of Expressions to Taste Japanese Sake by Creating Onomatopoeia”, The 29th Annual Conference of the Japanese Society for Artificial Intelligence (2015) (in Japanese)

[Pennacchiotti 2010]

M. Pennacchiotti and A. M. Popescu: “A Machine Learning Approach to Twitter User Classification”, in Proc, ICWSM’11, pp.281-288 (2010)

[Ribeiro 2016]

M. T. Ribeiro, S. Singh, and C. Guestrin: “Why Should I Trust You? Explaining the Predictions of Any Classifier”, Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp.1135-1144 (2016)

[Saito 2018]

斎藤康毅: 「ゼロから作る Deep Learning②—自然言語処理編」, オライリージャパン (2018)

[Sakaurai 2000]

櫻井広幸: 「香り表現における感覚用語—共感的表現について—」, 立正大学文学部論叢, Vol.111 (2000)

[Sakiyama 2021]

崎山一哉: 「チョコレートの呈味関与物質とコク」, コク研究会シンポジウム 2021 講演要旨集, 女子栄養大学駒込キャンパス, 2021-08-31, コク研究会 (2021)

[Sakauchi 2017]

G. Sakauchi, S. Tsuchida, M. Asada, N. Sato, K. Suzuki, K. Shibuya and H. Inaba: “Do others' comments affect the taste and electric taste thresholds of co-eaters?”, Niigata Journal of Health and Welfare, Vol.17, No.1, pp.39-39 (2017) (in Japanese)

[Salton 1983]

G. Salton and M. J. McGill: “Introduction to modern information retrieval”, McGraw-Hill Book Company (1983)

[Sanh 2019]

V. Sanh, L. Debut, J. Chaumond and T. Wolf: “DistilBERT: “a Distilled Version of BERT:Smaller, Faster, Cheaper and Lighter”, arXiv preprint arXiv:1910.01108. (2019)

[Seto 2003]

瀬戸賢一編著：「ことばは味を超える 美味しい表現の探求」，海鳴社（2003）

[Seto 2005]

瀬戸賢一，山本隆，楠見孝，澤井繁男，辻本智子，山口治彦，小山俊輔：「味ことばの世界」，海鳴社（2005）

[Shibata 2019]

柴田知秀，河原大輔，黒橋禎夫：「BERTによる日本語構文解析の精度向上」，NLP 2019, F2-4 (2019)

[Shimizu 2019]

清水慎太郎：「カレーのコクについて，各種食品のコクの見える化ーその概念と評価法の確率に向けて」，コク研究会シンポジウム 2019 講演要旨集，女子栄養大学駒込キャンパス，2019-08-22，コク研究会（2019）

[Shigeta 2018]

重田識博，難波英嗣，竹澤寿幸：「分散表現を用いた複数料理レシピからの典型手順の抽出」，第 10 回データ工学と情報 マネジメントに関するフォーラム（DEIM Forum 2018），P8-4 (2018)

[Sumiya 2019]

角谷和俊, 難波英嗣, 牛尼剛聡, 若宮翔子, 王元元, 河合由起子: 「料理レシピデータのメディア特性分析と利活用」, 人知能学会論文誌, Vol.34, No.1, pp.32-40 (2019)

[Sutskever 2014]

I. Sutskever, O. Vinyals and Q. VLe: “Sequence to Sequence Learning with Neural Networks”, NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems, Vol.9, pp.3104–3112 (2014).

[Tachibana 2013]

橘明穂, 若宮翔子, 角谷和俊: 「料理レシピサイトにおける料理名の修飾語に着目したネーミングコンセプト抽出」, 第 5 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2018), C4-4 (2013)

[Tachibana 2020]

S. Tachibana and K. Tsuda: “A Study of a Method to Understand the Intention of Taste Expressions through Text Mining”, Procedia Computer Science, Vol.176, pp.1793-1802 (2020)

[Tachibana 2021a]

S. Tachibana, T. Harada and K. Tsuda: “A method to understand the intention of taste expressions from word-of-mouth data of cooking recipe sites using text mining”, IEEJ Transactions on Electronics, Information and Systems, Vol.141, No.10, pp.1115-1124 (2021) (in Japanese)

[Tachibana 2021b]

S. Tachibana and K. Tsuda: “A Study of a Method to Understand the Intention of Taste Expressions of Sake through Text Mining”, Procedia Computer Science, Vol.192, pp.1687-1695 (2021)

[Tachibana 2022]

S. Tachibana and K. Tsuda: “A Study on Analysis Model for Influence of Expressions from Others on Taste Expressions in Word-of-Mouth Data of Cooking Recipe Website”, *Procedia Computer Science*, Vol.207, pp.1500-1508 (2022)

[Tanimura 2002]

谷村修也 : 「ビールのこくについて」, *味と匂学会誌*, Vol.9, No.2, pp.143-146 (2002)

[The National Research Institute of Brewing 2013]

The National Research Institute of Brewing, an Independent Administrative Institution: “Glossary of Terms on Sake Bottle Labels” (2013)

< https://www.nrib.go.jp/sake/pdf/nlziten_e.pdf > Accessed April 1, 2021.

[Tohoku University 2020]

Tohoku University: “Pretrained Japanese BERT models” (2020)

<<https://github.com/cl-tohoku/bert-japanese>> Accessed December 31, 2022.

[Uno 2017]

R. Uno, F. Kobayashi and S. Kotake: “Analysis of ideophones for physically perceived and imagined rice crackers”, *The 31st Annual Conference of the Japanese Society for Artificial Intelligence* (2017) (in Japanese)

[Utsunomiya 2006]

H. Utsunomiya: “Flavor Terminology and Reference Standards for Sensory Analysis of Sake”, *Journal of the Society of Brewing*, Vol.101, No.10, pp.730-739 (2006) (in Japanese)

[Utsunomiya 2007]

H. Utsunomiya: “Smell and Perfume Related to the Sensory Evaluation of Sake”,

Journal of Japan Association on Odor Environment, Vol.38, No.5, pp.352-360
(2007) (in Japanese))

[Vanik 1998]

V. N. Vanik: “Statistical Learning Theory”, John Wiley & Sons (1998)

[Vaswani 2017]

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L.Kaiser
and I. Polosukhin: “Attention Is All You Need”, Proc. NIPS 2017, pp.5998-6008
(2017)

[Vidal 2003]

S. Vidal, L. Francis, S. Guyot, N. Marnet, M. Kwiatkowski, R. Gawel, V. Cheynier
and E.J. Waters: “The mouth-feel properties of grape and apple proanthocyanidins
in a wine-like medium”, Journal of the Science of Food and Agriculture, Vol.83,
No.6, pp.564-573 (2003)

[Wakamiya 2014]

S.Wakamiya, Y. Kawai, H.Nanba and K. Sumiya: “Extracting naming concepts by
analyzing recipes and the modifiers in their titles”, Trans. on Engineering
Technologies, pp.321-335 (2014)

[Watanabe 2015]

C. Watanabe and S. Satoshi: “Onomatoperori: ranking cooking recipes by using
onomatopoeias which express their tastes and textures”, Journal of Japanese
Society for Artificial Intelligence, Vol. 30, No.1, pp.340-352 (2015) (in Japanese)
渡辺智恵美, 中村聡史: 「オノマトペロリ: 味覚や食感を表すオノマトペによる料理レ
シピのランキング」, 人知能学会論文誌, Vol.30, No.1, pp.340-352 (2015)

[Weng 2010]

J. Weng, E. P. Lim, J. Jiang and Q. He: “Twitter rank: Finding topic-sensitive influential Twitterers”, in Proc, WSDM’10, pp.261-270, ACM (2010)

[Yamakata 2013]

山肩洋子, 今堀慎治, 杉山祐一, 田中克己: 「レシピフローグラフを介したレシピ集合の要約と特徴抽出」, 電子情報通信学会技術研究報告, Vol.113, No.214, DE2013-40, pp.43-49 (2013)

[Yamakata 2016]

山肩洋子, 今堀慎治, 森信介, 田中克己: 「ワークフロー表現を用いたレシピの典型性評価と典型的なレシピの生成」, 電子通信学会論文誌, J99-D(4), pp.378-391 (2016)

[Yamakata 2019]

Y. Yamakata, H. Nanba and S. Mori: “Cooking Recipe Text Processing and its Applications”, Journal of Japanese Society for Artificial Intelligence, Vol.34, No.1, pp.15-23 (2019) (in Japanese)

[Yamamoto 2005]

山本隆: 「味ことばの世界」, 海鳴社, pp.54-87 (2005)

[Yamazaki 2014]

山崎健史, 森信介, 河原達也: 「フローグラフからのレシピ文自動生成」, 情報処理学会自然言語処理研究会, NL-219 (2014)

[Yoshida 1976]

吉田重厚: 「第1章 ビールの一般成分」, 日本醸造協会誌, Vol.71, No.7, pp.505-510 (1976)

関連業績

第 3 章

【査読付国際学会】

Shinichi Tachibana and Kazuhiko Tsuda: “A Study of a Method to Understand the Intention of Taste Expressions of Sake through Text Mining”, *Procedia Computer Science*, Vol.192, pp. 1687-1695 (2021)

第 4 章

【査読付学術論文】

立花 伸一, 原田 智彦, 津田 和彦: 「テキストマイニングを用いたレシピ口コミ分析による味覚表現の意図解明」, *電気学会論文誌C (電子・情報・システム部門誌)*, Vol.141, No.10, pp.1115-1124 (2021)

第 5 章

【査読付国際学会】

Shinichi Tachibana and Kazuhiko Tsuda: “A Study on Analysis Model for Influence of Expressions from Others on Taste Expressions in Word-of-Mouth Data of Cooking Recipe Website”, *Procedia Computer Science*, Vol.207, pp. 1500-1508 (2022)