

調理手順の類似性に基づく料理レシピの曖昧な表現の補完

福元 颯[†] 松下 光範[†]

[†] 関西大学総合情報学部 〒569-1095 高槻市霊仙寺町2丁目1番1号

E-mail: †k179992@kansai-u.ac.jp, ††mat@res.kutc.kansai-u.ac.jp

あらまし 本研究は、料理レシピにおける曖昧な表現を補完することで、料理に関する知識の少ないユーザに対してレシピの理解を支援することを目指す。現在、ユーザ投稿型レシピサイトには多種多様な料理レシピが投稿されており、ユーザは様々なレシピを手軽に閲覧することができる。しかし、投稿されるレシピには、手順の省略によって曖昧な表現が存在する。このような表現は料理に関する知識の少ないユーザにとって理解が困難である。この問題を解決するためには、曖昧な表現に対して情報を補完し、ユーザの知識レベルに応じた情報を提示する必要がある。本稿では、料理レシピの調理手順に着目し、同一の料理のレシピ群から類似度の高い調理手順の抽出を行うことで、曖昧な表現を補完する手法を提案する。

キーワード 料理レシピ, フローグラフ, 共起分析, 手順理解

Complementing Ambiguous Expressions in Cooking Recipes Based on Similarity of Cooking Procedures

Hayate FUKUMOTO[†] and Mitsunori MATSUSHITA[†]

[†] Faculty of Informatics, Kansai University 2-1-1, Ryozenji-cho, Takatsuki-shi, Osaka 569-1095, Japan

E-mail: †k179992@kansai-u.ac.jp, ††mat@res.kutc.kansai-u.ac.jp

Abstract The purpose of this study is to support understanding of cooking recipes for users with lack of knowledge by complementing ambiguous expressions on the recipes. Currently, various recipes are posted on website by users and easy to browse. However, the submitted recipes are omitted and hard to understand for those kind of users due to ambiguous of procedure. As a first step to support, it is important to present supplemented information according to the user's level of knowledge. This paper proposes a method to complement ambiguous expression of cooking procedure on recipes by extracting high similar procedures from the same recipe group.

Key words cooking recipe, flow graph, co-occurrence analysis, procedure understanding

1. はじめに

インターネットやスマートフォンの普及により、cookpad^(注1)や楽天レシピ^(注2)などのユーザ投稿型レシピサイトが利用可能である。このようなサイトはレシピの閲覧だけでなく、ユーザ自身が作成した料理レシピを投稿することができる。

しかし、ユーザ投稿型レシピサイトに独自の料理レシピを投稿するユーザは、自らの料理知識や技術レベルを基準に料理レシピを投稿するため、調理手順などを詳細に記述しない場合がある。例えば、料理レシピには調理動作の省略や表記ゆれ、オノマトペを使った抽象的な表現などが存在する。具体的には、

表1で示した実線部分はどちらの文章も「切る」という意味の言葉を使用しているが、レシピ1^(注3)では切り方の指定がなく、レシピ2^(注4)では切り方の指定がある。このような曖昧な表現によって、料理に関する知識の少ないユーザは記述されている調理工程を遂行することが困難になる。

この問題を解決するには、料理知識に応じた情報を提示し曖昧な表現を補完する必要がある。本研究では料理レシピの曖昧な表現を補完するため、料理レシピの構成要素を関連付けたデータベースを構築することを目標とする。データベースを構築するためには、料理レシピの構成要素間の関係性を抽出する必要がある。そこで、本稿では、食材と調理動作間の共起関係

(注1) : クックパッド株式会社 : [cookpadhttps://cookpad.com/](https://cookpad.com/)

(注2) : 楽天レシピ : <https://recipe.rakuten.co.jp/>

(注3) : レシピ1 : <https://cookpad.com/recipe/5561470>

(注4) : レシピ2 : <https://cookpad.com/recipe/5538573>

表 1 料理レシピごとの調理手順の細かさや食材の違い (クックパッド株式会社: cookpad より引用)

レシピ	調理手順
1	玉ねぎ, 人参, ピーマンをみじん切りにします。
2	玉ねぎを切って色が変わるまで軽く炒める。

を分析した。

2. 関連研究

2.1 料理レシピの構成要素を扱う研究

料理レシピの構成要素を扱う研究では, 料理レシピのオントロジーや構成要素のデータフレームを構築し, 構成要素間の関係から料理レシピの省略された部分の補完などを試みている。

Nanba らは, 料理レシピテキストと特許データベースを情報源とし, 統計的言語処理技術を用いることによって, 料理オントロジーの構築を試みた [2]。オントロジー構築では, 辞書の構築やカテゴリ設定など5つの段階で行っている。収集した関連語を手で選出することで同義語辞書を構築し, 同義語 5,023 語を獲得した。料理オントロジーの同義語辞書を用いることで, 言語処理精度の改善の可能性についても示している。

赤澤らは, 冷蔵庫にある食材を利用した料理レシピをランキング形式で推薦する手法を提案し, システムの実装を行った [4]。ユーザが効率よく冷蔵庫にある食材だけで作ることができる料理レシピを検索できるように, 冷蔵庫の食材と料理レシピの食材に注目している。ランキングの決定には, 食材の充足率, レシピの特異度が高い, 低いといった尺度を用いている。

志土地らは, 料理レシピにおける代替可能な食材を推薦する方法を提案した [6]。この研究では, 料理レシピ中で適用される調理手順が類似している食材を, 代替して使用できる素材と考えている。そのために, 各調理動作とその素材を対応づけ, 素材ごとに各料理レシピ内で適用される調理動作を列挙することで, 素材に対する調理手順を抽出している。そして, 同一カテゴリ内の料理レシピ群における特徴的な調理手順の類似度を算出することにより代替可能な食材を推薦している。

2.2 料理レシピのフローを扱う研究

料理レシピのフローを扱う研究では, 調理手順をフローグラフやマークアップ言語などを用いて表現し, システムで料理レシピにおける調理手順の省略や調理手順の関係性を扱えるようにすること (以下, 構造化と記す) を試みている。

Mori らは, 料理レシピを用いて手順を示したフローグラフから手順を示す文を自動生成する手法を提案した [1]。この研究の目的は, 学習コーパスからテンプレートを作成し, フローグラフを入力とすることで統計的な文生成を行うことである。料理レシピ中の固有表現である「食材」や「道具」に対してタグ付けを行い, 抽出した単語を接点として扱うことで, 図1のようなフローグラフを作成した。

浜田らは, 料理レシピの構造解析とそのフローグラフの自動生成を行った [9]。構造解析では, 料理レシピに記述されている名詞と動詞から辞書の構築を行った。辞書の構築後に名詞と動

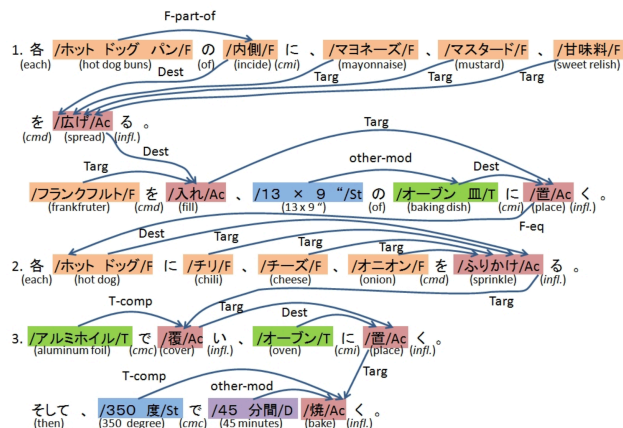


図 1 料理レシピにおけるフローグラフの例 (文献 [1] より引用)

詞の関係をグループ化し, それらを調理手順の順番につなげ, フローグラフ化することにより料理レシピの構造化を行っている。

吉川らは, 料理レシピをデータフロープログラミング言語を用いることで料理レシピの構造化を行った [7]。料理の材料や料理工程の流れを視覚的に判断しながら構造化されたレシピをプログラミングに対応させることを目標に, 料理の全てのデータや関数をオブジェクトとして表現し, オブジェクト同士を線でつなぐことで有向グラフとしてプログラムを作成できるようにした。

2.3 本研究の位置づけ

2.1 節の料理レシピの構成要素を扱う研究では, 構成要素の調査やその関係性に注目することで料理レシピ中の省略された部分の補完などを行っている。これらの先行研究を参考にデータベースの構築を行い, 料理動作と食材の共起関係を調査する。

また, 2.2 節の料理レシピの構造を扱っている先行研究においては, 料理レシピを構造化することによって調理手順の把握や支援がされている。しかし, 調理手順が複雑な構造である場合, フローグラフでは対応できない点や, 提案されている構造化の手法に置いて料理の構成要素同士の関係の全てを網羅することはできない点から不完全である。そのため, 料理調理動作における曖昧性を解消することはできていない。

以上の研究に対し, 本研究では曖昧性を補完するため, 料理レシピの構成要素間の関係を網羅したデータベースを構築することを目標とする。

3. 料理レシピを構成する要素

料理レシピは基本的にテキストと写真で構成されている。料理レシピの中にはタイトル, 料理の説明, 材料, 料理の写真, 手順を主に記載している (図 2 参照)。料理レシピの中には動画を用いたものも存在するが, 本研究では料理レシピの調理手順における内容を対象とするため, 本研究で扱う料理レシピはテキストと写真, もしくはテキストのみで構成されているものとする。さらに, 調理工程におけるユーザの困難を取り除くことを目的とすることから調理手順を扱う。調理手順とは料理を

ふわふわオムレツ。からのオムライス。 タイトル 



バターの良い香りとチーズのコク、そしてふわわり卵と最強オムライス！

料理の説明 

材料 (1人分)

★卵	2個
★粉チーズ	大きじ1
★牛乳	大きじ1
バター	4g
好きなライス	1人分

料理の写真 材料

手順

-  チキンライスであれば、下記レシピを参照してみてください！
レシピID:5631250
-  ★をボールに入れ、よく混ぜる。
-  フライパンを中弱火にかけバターを溶かしフライパン全体に広げ、2を入れ、箸で端から真ん中へと卵を集めていく。
-  と、こうなるので、これをゆすってそのままライスにかけるのもOK!!中をとりとさせなければ包み込んで写真のように。

図 2 料理レシピの構成 (クックパッド株式会社:cookpad より図引用)






-  レトルトご飯を2分レンジする。ミートソースは、袋の半分を器に移し、1分レンジする。
-  フライパンを温め、油をひいて温めたご飯を入れる。軽く炒めたら、ミートソースを大きじ4杯くらい入れ、さらに炒める。
-  ご飯とソースをよく馴染ませながら炒めたら、塩を2つまみ入れ、味を整える。
-  炒めたライスは皿に移し、卵1個軽く溶き、フライパンに油をひいて薄焼きの半熟にする。
-  炒めたライスの上に、薄焼きした卵焼きを乗せて、残りのミートソースをかけて、出来上がり♪

図 3 料理レシピの手順 (テキストのみ) (クックパッド株式会社:cookpad より図引用)

作するために何をしなければならないかを示している。基本的に調理手順は複数の工程に分かれており、料理レシピの投稿者が工程のまとめごと順番に番号をつけて調理手順として記載する。料理の調理手順の表し方は二通りある。(1) テキストのみで記載されている場合 (図 3 参照)、(2) テキストと写真で記載されている場合、である (図 4 参照)。調理手順の表し方は料理レシピを投稿するユーザが決定するため、書き方が様々になる。

4. 調理動作に関するアンケート

4.1 アンケートの概要

料理レシピに対するユーザの理解度に影響を及ぼす原因を明らかにするため、アンケート調査を行った。アンケートの対象者は大学生 10 名 (男性 7 名, 女性 3 名) とした。事前アンケートの質問項目を以下に記す。今回の調査では、料理レシピへの理解度が高い状態を「作業工程を容易に想像できること」

-  ☆を合わせておく。
-  フライパンに油を敷き、卵を伸ばし炒め1を入れ包む。
-  お皿に盛ったらケチャップをかけて召し上がれ。

図 4 料理レシピ手順 (テキストと写真) (クックパッド株式会社:cookpad より図引用)

及び「調理手順の内容が整理して記述されていると感じること」と定義した。調査では、まず料理経験に関するアンケート (以降、事前アンケート) に回答してもらった。次に、同じ料理のレシピを 5 パターン読んでもらい、それぞれの料理レシピの理解度に関するアンケート (以降、本アンケート) に回答してもらった。さらに、理解度が低い箇所には、自由記述式でその理由を記載してもらった。事前アンケートの内容を以下に記す。

- 1). 料理頻度
- 2). 何を参考にするか (レシピ, 本など)
- 3). 具体的な料理経験
- 4). 料理レシピを見て理解困難と感じた体験

本アンケートでは 5 パターンの料理レシピを読んでもらい 2 項目の質問に答えてもらった。料理レシピは一般的に知られていて、調理動作が比較的単純な「オムライス」を採用した。質問は、調理動作「切る」、「炒める」に該当する、全ての記述に対して回答してもらった。質問の内容を以下に記す。

- 1). 調理工程を容易に想像できるか
- 2). 料理レシピとして内容や記述法が整理されているか

これらの項目に対し、1. 理解度が低い、5. 理解度が高いとし 5 段階評価で回答してもらった。さらに、5 段階中 1, 2 を選んだ実験参加者には、選んだ理由と料理レシピ中の低評価の原因となった箇所を記述してもらった。

4.2 アンケートの結果

事前アンケートで得られた参加者の料理頻度を表 2 に、ユーザが料理レシピを読んだ際の理解度を計測するためのアンケートで得られた調理手順の評価を表 3 に示す。

- 調理動作「切る」について

表 3 より、質問「作業工程を容易に想像できたか」「内容、記述用法が整理されていたか」の評価がいずれも高い。よって、オムライスの料理レシピにおいて調理動作「切る」の理解度が高いことがわかる。原因は提示したレシピのほとんどが「みじん切り」をするというように、切り方が具体的であったことである。「作業工程を容易に想像できたか」について評価が低い回答例は「玉ねぎを切って」という文章に対して「切り方が指定されておらず、切る大きさがわからない」であった。「内容、

表 2 事前アンケート対象者の料理頻度

参加者	料理頻度
A	普段はしない
B	普段はしない
C	週に一回程度
D	週に一回程度
E	2日に一回程度
F	月に一回程度
G	週に一回程度
H	毎日
I	普段はしない
J	週に一回程度

表 3 調理動作の理解度に関する5段階評価の平均値

調理動作	質問	評価
切る	作業工程を容易に想像できたか	4.22
	内容、記述方法が整理されていたか	4.30
炒める	作業工程を容易に想像できたか	3.70
	内容、記述方法が整理されていたか	3.38

記述方法が整理されていたか」について評価が低い回答例は料理レシピ上の調理動作「切る」と「きざむ」に対して、「切り方の表記が混合して用いられているため、整理された文章とは言えない」という回答が得られた。

- 調理動作「炒める」について

表3より、調理動作「切る」に比べていずれの質問も評価が低いことがわかる。調理動作「炒める」は「切る」に比べ、食材・調味料を追加、火加減や終了するタイミングなどの構成要素が関係する。よって、調理工程を完全に読み取ることは困難である。「作業工程を容易に想像できたか」について評価が低い回答例は「加えて炒める」という文章に対して「動作を終了するときの判断基準がわからない」であった。「内容、記述方法が整理されていたか」について評価が低い回答例は「具材を加えて炒めます。ご飯を炒め」という箇所に対して「ご飯を炒めるタイミングがわからない」や「ご飯を入れる前にどの程度炒めればよいかわからない」であった。

5. デザイン指針

本アンケート調査の結果、一般的に知られている料理のレシピにもユーザが理解しにくいと感じる箇所があることがわかった。原因は調理手順に食材や火加減、また調理動作の終了時間などの条件の省略である。本研究では、このような記述の省略をユーザの料理レシピの理解を妨げる、曖昧な表現と定義する。

料理に関する知識の少ないユーザは食材が曖昧に表現されている事によって、料理をする上で様々な問題に直面する。材料が曖昧に表現されることで、食材や調味料の量が適切に理解できず、味の濃薄による料理の失敗が起こる可能性などがある。また、調理手順が曖昧に表現されている場合、調理動作や使用する調理器具などが正確に把握できないため、調理工程を遂行することが困難である。本研究では、これらの料理レシピの曖昧な表現の内、調理工程を遂行する上でより影響を及ぼす調理

タグ	意味	備考
F	食材	代名詞・中間・最終生成物を含む
T	道具	調理動作や器など（代名詞を含む）
D	継続時間	概数表現を含む
Q	分量	概数表現を含む
Ac	調理者の動作	語幹のみ
Af	食材の動作	語幹のみ
Sf	食材の状態	
St	道具の状態	温度設定など

表 4 レシピ用語とタグ一覧

ラベル	意味	概説
Agent	主語 (ガ格)	主に「が」や「は」で表される動作と主語の関係
Targ	対象 (ヲ格)	主に「を」で表される動作と対象の関係
Dest	方向 (ニ格)	主に「に」で表される動作と方向や場所の関係
T-comp	道具デ	主に「で」で表される動作とその手段の関係
F-comp	食材デ	主に「で」で表される動作とその手段の関係
F-eq	同一の食材	既出の食材とそれに対する参照表現
F-part-of	食材の一部	既出の食材とその一部に対する参照表現
F-set	食材の集合	既出の複数種の食材とその全体に対する参照表現
T-eq	同一の道具	既出の道具とその一部に対する参照表現
T-part-of	道具の一部	既出の道具とその
A-eq	同一の動作	既出の動作とそれに対する参照表現
V-tm	動作のタイミング	別の動作を行う条件やタイミングを示す句の動詞
other-mod	その他の修飾語句	

表 5 辺のラベルの一覧

手順を扱う。

調理手順の曖昧な表現には様々な種類が存在する。「玉ねぎ」と「オニオン」のような食材の表記ゆれや調理器具、調理動作の省略などがある。そこで、調理する際に必要となる知識である食材、調理器具、調理動作などの構成要素がユーザの料理に関する知識のレベルに合わせて提示できれば、料理レシピを読んで調理が困難になることはなくなる。そのため、料理レシピ構成要素の調理動作と食材との共起関係を抽出し分析する。

6. フローグラフコーパス

本章ではの共起関係分析で扱うフローグラフコーパスの詳細について述べる。

6.1 定義

フローグラフコーパス (FG corpus) では 3. 章で述べたレシピを構成する要素の内, 手順の意味内容の表現形式で,

- 頂点
食材や調理動作といったレシピ特有の表現
- 辺
レシピ用語間の関係

この 2 つの要素からなる根つき無閉路有向グラフで表される (図 1 参照). 無閉路有向グラフの根 (図 3 では「焼/Ac」) が, 最終的に生成される料理に対応する.

6.2 レシピ用語とタグ

フローグラフの頂点は, 食材や動作を表す単語列 とタグの組である (表 4 参照). 食材や道具といった名詞句だけでなく, 調理者の動作のような用語もレシピ用語として扱っている. さらに, 同一性の簡単な判定のために, 活用語は語幹のみをレシピ用語としている.

6.3 レシピ用語間の関係

フローグラフの辺は, 用語間に関係があることを表す. 有向グラフで示され, 関係は大きく 4 つある (表 5 参照).

- 述語項関係
調理手順の内容は主に「食材」と「動作」の関係であるため, これらの述語項関係は最も重要である. Agent, Targ, Dest, T-comp, F-comp が該当する.
- 2 つの食材や道具の参照関係
既に出現している食材や, 食材を加工した物体を指すレシピ用語を関連付ける. F-eq, F-part-of, F-set, T-eq, T-part-of が該当する.
- 2 つの動作の関係
実際には一度しか行われない動作を繰り返して記述している部分や, 動作のタイミングを関連付ける. A-eq, V-tm が該当する.
- その他
その他の修飾関係を Other-mod としている.

辺は同一文内の頂点の間だけではなく, 異なる文の頂点の間に対しても付与される. したがって, 調理手順全体の構造を表現することが可能になる. 例えば, 各食材からの辺を辿ることで, その食材がどういった手順で処理されていくのかがわかる.

7. 曖昧性解消へ向けた取り組み

7.1 調理手順ごとの類似度算出

調理手順における曖昧な表現には, 様々な種類があり, 調理手順ごとにその理解度や記述内容の粒度は異なる. 理解度が低い調理手順に対して, 構成要素が詳細に書かれている調理手順を推薦することで, ユーザは理解困難な箇所を調べることなく調理工程を遂行することが可能になる. よって類似度の高い調理手順を算出し, 調理手順ごとの記述方法の違いを調査する.

7.1.1 分析対象

料理レシピサイト cookpad より投稿された料理レシピの内, 「オムライス」のレシピを選択した. 分析には Mori らが作成したフローグラフコーパスのうち, 読み込み可能であった 3,470

件の json ファイルを用いる.

7.1.2 分析方法

オムライスの料理レシピ 3,470 件から調理手順を抽出し, その手順を一文ごとに分けた. さらに, MeCab を用いて形態素に分解した. その際, 「あれ」や「いつ」などの料理レシピの調理手順の理解度に影響のないものが含まれることを防ぐため, Slothlib [8] に設定されている名詞群をストップワードに設定した.

次に形態素に分解した単語の中から名詞と動詞のみを取得し, TF-IDF 法を用いて単語ごとの重要度を算出した. TF-IDF 法は, 文章処理におけるキーワードの重み付け手法として広く利用されている [5]. ある文章中の単語の出現頻度 TF 値と, 単語の文章中における頻度の偏りを表す IDF 値の積で計算される [3]. 単語 i に対する TF 値は下式 (1) により計算される. $n_{w,s}$ は調理手順 s における単語 w の出現頻度を表し, X は調理手順 s に出現する単語の総出現数を表す.

$$TF_{w,s} = \frac{n_{w,s}}{X} \quad (1)$$

単語 w に対する IDF 値は下式 (2) により計算される. 対象となる調理手順 s の総手順数を N とし, df_w は単語 s を含む文書数を表す.

$$IDF_w = \log_e \frac{N}{df_w} + 1 \quad (2)$$

式 (1), (2) から単語 w に対する TF-IDF 値は下式 (3) により計算できる.

$$TFIDF = TF_{w,s} \times IDF_i \quad (3)$$

さらに, TF-IDF 法によって計算された各語の重みをもとに類似度を算出するため, コサイン類似度を扱う. コサイン類似度とは, ベクトルの類似性を示す指標の一つである. コサイン類似度の計算式は, 文章 A,B の TF-IDF 値ベクトル v_A, v_B を用いると式 (4) で表現できる.

$$\cos(v_A, v_B) = v_A \cdot v_B = \frac{\sum_i w_i A w_i B}{|v_A| |v_B|} \quad (4)$$

7.1.3 分析結果

類似度を算出した結果, 様々な調理手順のペアが現れた. その内, 多く現れた 3 つの例を紹介する.

- 調理手順の内容と記述方法が一致する場合

「玉ねぎはみじん切りに」, 「玉ねぎはみじん切りにします」のペアを例に挙げると, 食材と調理動作の内容が一致している.

- 調理手順の内容が一致しない場合

「ここで火を切ります」, 「材料を切ります」のペアを例に挙げると, 「切ります」という調理動作の部分が一致しコサイン類似度が高くなっているが, 調理動作の意味と内容が異なっている.

- 調理手順の内容が一致し, 記述方法が異なるもの

「ベーコンは 1 センチ幅くらいに切る」, 「ベーコンは 5 ミリ~1 センチ幅の短冊切り」のペアを例に挙げると, ベーコン

を切るという調理手順の内容は同じだが、前者では切り方は指定されておらず、一方で後者は切り方が指定されている。このようなパターンは料理初心者が調理手順によって理解度に差が生じると考えられる。

7.2 使用するデータの検討

調理動作「切る」を補完するために、切り方として適切に表現されている別のレシピの情報を用いる。そこで、食材と調理動作の共起関係を分析する。しかし、共起したすべてのデータを使用すると出現頻度の少ないデータが含まれる可能性がある。そのため、食材の出現頻度を分析することで、共起関係を分析する際にどのようなデータが有用であるかの検証を行った。

7.2.1 分析対象

料理レシピサイト cookpad より投稿された料理レシピの内、「オムライス」のレシピの中からランダムに 1000 件を選択した。分析には Mori らが作成したフローグラフコーパスの内、読み込み可能であった 713 件の json ファイルを用いる。

7.2.2 分析方法

読み込み可能であった 713 件の json ファイルの「食材/F」をすべて読み込み食材リストを作成した。さらに、その食材リストをトークンとし NLTK (Natural Language Tool Kit) ver.3.4 を用いて出現頻度分布グラフ (図 9 参照) を作成し、出現頻度が 1 回のみデータの除いた食材リストを作成した。

7.2.3 分析結果

オムライスの料理レシピ 713 件から、食材のリストを作成した結果、単語数 12,735 個のリストを作成することができた。その食材リストには 1637 個の異なり語が存在することがわかった。食材の頻度分布のグラフからは、オムライスを作る上で一般的な食材は頻度が高くなっていることがわかった。さらに、出現頻度が 1 回の食材のリストには、「手のひら」や「ラグビーボール」など、食材ではないものも含まれていた。このように、作成した食材リストの中には、食材の表記ゆれや料理レシピ独自の表現 (図形など) などが含まれている。そのため、食材と調理動作の共起関係の分析には頻度の高い食材を使用すべきである。

7.3 食材と調理動作の共起関係分析

類似度の高い調理手順の「調理内容が一致し、記述が異なるもの」は記述方法によって、調理手順の具体性に違いがあった。ユーザにとって必要な情報である構成要素がより詳細に書かれている場合、ユーザは調理手順の理解が容易となる。よって、調理動作に関連づいている構成要素が多く、記述が具体的であれば、調理工程をより理解しやすいと考えられる。よって、調理手順において食材と調理動作の共起関係を分析する。

7.3.1 分析対象

料理レシピサイト cookpad より投稿された料理レシピの内、「オムライス」のレシピの中からランダムに 1000 件を選択した。分析には Mori らが作成したフローグラフコーパスの内、読み込み可能であった 713 件の json ファイルを用いる。

7.3.2 分析方法

オムライスの料理レシピ 713 件から手順を抽出し、前処理として頻度の多い食材の中でも、表記ゆれがある食材を word2vec

食材	回数
たまねぎ	177
にんじん	57
鶏肉	33
ピーマン	29
1	27
野菜	19
にんにく	19
バター	16
ベーコン	15
材料	12

表 6 オムライスの食材と調理動作「みじん切り」の共起頻度

を使って算出した類似度を参考に手作業で表記を統一した。例えば、「玉葱」や「玉ねぎ」は「たまねぎ」に統一している。次にオムライスにより使用されている切り方の頻度を分析した。その結果、「みじん切り」の頻度が高かったため、共起関係の分析には 7.2 節で分析した頻度が高い食材と調理動作を使用し、組み合わせを算出した。さらに、切り方が省略されている食材を明らかにするため、調理動作「切る」と食材との共起関係も分析した。

7.3.3 分析結果

調理動作「みじん切り」と共起している食材の頻度の上位 10 個を表 6 に記す。調理動作「みじん切り」と食材の共起した回数は 867 回であった。表において「たまねぎ」、「にんじん」、「鶏肉」などの食材が調理動作と多く共起していることがわかる。「たまねぎ」、「にんじん」はオムライスに含まれるチキンライスを作る際、一般的にみじん切りされるため、分析結果が妥当であることがわかる。しかし、みじん切りすべきではない食材である「鶏肉」においても、すべてのオムライスのレシピにおける出現頻度が高いため、みじん切りとの共起頻度が多くなっている。このような、ただ頻度が高い食材の処理についても今後は考慮しなければならない。

8. 終わりに

本研究では料理レシピの曖昧な表現を解消するため、料理レシピ構成要素のデータベースを構築することを目標とし、その端緒として、アンケート結果から料理レシピに対する料理者の理解度を評価し、本研究で扱う曖昧な表現の定義を行った。また、料理レシピ構成要素の調理動作と食材との共起関係を分析した。その結果、曖昧性を補完するための課題が明らかになった。今後の方針として、抽出した構成要素の関係性を元にレシピ上の曖昧な表現に対して情報提示することで調理動作を補完することを目標とする。

謝辞

本研究では、クックパッド株式会社と国立情報学研究所が提供する「クックパッドデータ」を利用した。ここに感謝の意を記す。

文 献

[1] Mori, S., Maeta, H., Yamakata, Y. and Sasada, T.: Flow

- Graph Corpus from Recipe Texts., *The International Conference on Language Resources and Evaluation*, pp. 2370–2377 (2014).
- [2] Nanba, H., Doi, Y., Tsujita, M., Takezawa, T. and Sumiya, K.: Construction of a cooking ontology from cooking recipes and patents, *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, ACM, pp. 507–516 (2014).
- [3] 鈴木啓, 大内紀知: テキストマイニングによる学会の特徴比較分析, 経営情報学会全国研究発表大会要旨集 2016 年秋季全国研究発表大会, 一般社団法人経営情報学会, pp. 79–82 (2016).
- [4] 赤澤康之, 宮森恒: 冷蔵庫食材を考慮した料理レシピ検索システムの提案, データ工学と情報マネジメントに関するフォーラム (2011).
- [5] 日本図書館情報学会: 図書館情報学用語辞典, Vol. 第 4 版, 日本図書館情報学会用語辞典編集委員会 (2013).
- [6] 志土地由香, 高橋友和, 井手一郎, 村瀬洋: 調理レシピテキストからの代替素材の発見, 人工知能学会全国大会論文集, 一般社団法人 人工知能学会, pp. 6–6 (2008).
- [7] 吉川祐輔, 宮下芳明: グラフィカルデータフローによる調理レシピプログラミング言語の提案, 研究報告ヒューマンコンピュータインタラクション, No. 4, pp. 1–7 (2010).
- [8] 大島裕明, 中村聡史, 田中克己: SlothLib: Web サーチ研究のためのプログラミングライブラリ, 日本データベース学会 Letters, Vol. 6, No. 1, pp. 113–116 (2007).
- [9] 浜田玲子, 井手一郎, 坂井修一, 田中英彦ほか: 料理教材における手順の構造化, 第 60 回全国大会講演論文集, Vol. 2000, No. 1, pp. 9–10 (2000).