

アラビア語レシピ中のアレルゲン認識に関する一考察 A Study on Allergen Recognition in Arabic Recipes

大坪 愛佳*
Manaka Ootsubo

アルハビブ ファティマ*
Fatima Abdulhakim a Al Habib

大和矢 悠仁*
Yuji Yamatoya

延澤 志保*
Shiho Hoshi Nobesawa

1. 研究背景

海外旅行の楽しみの1つとして食があげられるが、アレルギーを持つ人は、蕁麻疹や喘息、アナフィラキシーショック症状へ至る場合もあり、外国料理を気軽に楽しむことができない。本研究では、アレルギーの原因となる食材(アレルゲン)の摂取を避けながら外国料理を安心して食べられるように、外国料理に含まれるアレルゲンの自動推定を行う。本稿ではアラビア料理のアレルゲン自動認識について考察する。

2. 含有推定対象とするアレルゲン

2.1. アラビア料理のための含有推定対象アレルゲン

日本では加工食品で表示義務があるまたは表示が推奨されているアレルギー食材が合計 28 品目定義されている¹。本研究では、イスラム教でハラムとして禁止されている食べ物は入っていないものと考え、日本のアレルギー表示対象 28 品目から豚肉、カニ、あわび、イカを除外し 24 品目を含有推定の対象とする(表 1)。

表 1: アラビア料理含有推定対象アレルゲンリスト

卵	乳	牛肉	鶏肉	ゼラチン
さけ	そば	えび	いくら	
小麦	大豆	そば	ごま	やまいも
桃	りんご	くるみ	落花生	アーモンド
バナナ	キウイ	オレンジ	まつたけ	カシューナッツ

2.2. アレルゲン含有レベル

料理の食材はその料理に必要な食材とアレンジで入れる食材に分けられる。花井らは、複数のレシピの比較においてあまり使用されない食材はそのレシピの特徴を示すものとし、材料の希少度に着目した [1]。本研究では、花井らのように食材の出現頻度を考慮することでアレルギーを持つ人が各自でより判断しやすくなるよう、アレルゲンの有無を食材の出現頻度などから 4 段階の含有レベルに分類することを提案する(表 2)。

表 2: アレルゲン含有レベル

含有確実	当該料理が必ず含有する食材
含有濃厚	当該料理が含有する可能性が高い食材
含有可能	当該料理が含有する可能性を持つ食材
含有僅少	当該料理が含有する可能性が十分に低い食材

本研究では、複数のレシピを用いて料理の食材を推定することで表 2 の判定を可能とする。

3. 外国料理レシピの原材料の特徴

3.1. 外国料理のレシピ検索

日本語でアラビア料理を検索しても、cookpad²や世界中のさまざまな伝統料理が書かれているレシピサイ

*東京都市大学, Tokyo City University

¹食物アレルギーの原因食物とアレルゲン表示, ニッポンハム食の未来財団, <https://www.miraizaidan.or.jp/allergy/cause.html>.

²cookpad, <https://cookpad.com/>.

ト³でさえ、有名な料理のみの、少ない料理数しか出てこないため、アレルゲンの含有を判断するには料理数が少ない。また、日本語のレシピでは、食材や調味料など日本人の好みに合わせた日本風アレンジがされているため本場の海外料理とは異なるものであり、実際に旅行に行った際のアレルゲン含有の参考にはならない。

そこで本研究では、海外の料理について現地のレシピを調べるものとする。本稿ではアラビア料理を対象とするため、レシピサイトとしてアラビア語の justfood⁴を用いる。本研究では日本語で外国の料理名を入力し、これを現地語に翻訳して現地語のレシピを検索し、その結果を日本語に翻訳してアレルゲンの含有の有無を示す。

3.2. アラビア語レシピの食材の記述

justfood で料理「コシャリ」のレシピ 10 件を調べ、材料欄に出現した食材を出現レシピ数順に並べたものを表 3 に示す。コシャリのアレルゲンは鶏肉、卵であ

表 3: 「コシャリ」の食材と出現レシピ数

食材	数	食材	数	食材	数
塩	10	唐辛子	5	コリアンダー	1
トマト	9	水	5	卵	1
黒コショウ	8	酢	5	パセリ	1
玉ねぎ	8	麵	4	じゃがいも	1
油	7	鶏肉	3	パブリカ	1
ご飯	7	レモン	2	石油	1
クミン	7	マーガリン	2		
レンズ豆	6				
にんにく	6				

る。コシャリはゆで卵を添えて提供されることがあり、レシピの1つでは、手順に「ゆで卵と一緒にお召し上がりください」との記述が出てくるが、このレシピの材料欄にゆで卵は記載されていない。このように、材料欄にはないが手順の中に含まれていた食材もある。

レンズ豆はアレルゲンリストの大豆や落花生と同じ豆類だが、豆はソラマメ属やヒラマメ属などの属で分かれており、大豆、青大豆、黒大豆、黒豆はダイズ属、落花生、南京豆はラッカセイ属、レンズ豆はヒラマメ属に分類されるため、レンズ豆は大豆や落花生とは異なり、アレルゲンには当たらない。このように、アレルゲンの推定では細かい分類まで調べる必要がある。

表 3 の「石油」は「油」の誤訳である。日本語レシピのタイトル、材料、手順を英語に翻訳した際のレシピ翻訳の誤り分類の研究 [2] によると、ここでの石油の誤訳は、意味が近い単語が出力される置換誤りに分類される。このように、アラビア語レシピからの翻訳でも誤りが起こり得る。アラビア語のレシピを翻訳する上で、翻訳の間違いが問題になる可能性がある。

同様に justfood で料理「マンディ」のレシピ 9 件を調べ、材料欄に出現した食材を出現レシピ数順に並べたものを表 4 に示す。マンディの食材の中でアレルゲ

³世界の料理レシピ, <http://jp.ndish.com/recipe/>.

⁴アラビアレシピサイト justfood, <https://www.justfood.tv/>.

表4: 「マンディ」の食材と出現レシピ数

食材	数	食材	数	食材	数
黒コショウ	9	ウコン	3	ベイリーフ	1
カルダモン	6	オイル	3	ナツメグ	1
クローブ	6	ニンニクペースト	3	サフラン	1
塩	5	唐辛子	3	タイム	1
鶏肉	5	レモン	2	バルガー	1
生姜	4	ターメリック	2	パプリカ	1
クミン	4	月桂樹	2	ニンジン	1
コリアンダー	4	水	2	トマト	1
シナモン	4			バセリ	1
玉ねぎ	4			レーズン	1
ご飯	4			石炭	1
バター	4				

ンに当たるものは、鶏肉、バルガー、バターの原材料である乳である。バルガーとはアラブ諸国で料理に用いられる小麦の加工品で、日本では馴染みのない食材であり、表1での表記も「小麦」のみのため、これがアレルギーであることを推定するためには小麦であることを自動的に検出する必要がある。またバターも表1には記述がなく、含有アレルギーを正しく推定するためにはバターの原材料まで自動的に検出する必要がある。また表4の「石炭」は、食材として表記されているのではなく、炭火焼などにする際の材料として書かれていた。このように、食材でなく材料が書かれてある問題の解決も必要であることがわかった。

4. 材料欄の複合原材料の推定

複合原材料とは、2種類以上の原材料からなるものである⁵。例えば、マヨネーズは複合原材料であり、その原材料は油や卵である。表4の材料欄にあるバターをWikipediaで調べると、乳からできていると記載されているため、バターにはアレルギーが含まれていることがわかる。このように、複合原材料は、アレルギーを発見するために原材料を調べる必要がある。

5. レシピ材料欄の表記の揺れ

日本語からアラビア語にする際、翻訳の揺れが問題となる。表5にコシャリのレシピ10件の材料欄に出現した食材のうち、類似した食材を含むものを示す。ト

表5: 「コシャリ」食材の表記の揺れ

米	洗米	ご飯	エジプト米	バスマティライス
パスタ	スパゲッティ	沸騰麺	カップ麺	
トマト	トマトペースト	トマトジュース	トマトソース	
豆	ひよこ豆	レンズ豆	黒レンズ豆	
鶏肉	チキン	チキンストック		
にんにく	ニンニク	ガーリック		
玉ねぎ	中玉ねぎ	オニオンパウダー		
油	オリーブオイル			
黒胡椒	黒コショウ			
水	お湯			
酢	白酢			
唐辛子	トウガラシ			
レモン	レモンジュース			

マトなどの食材の後にパウダー、ペーストなどの言葉が付いたもの(表5)のような加工食品も、含有アレルギー推定という目的の下では、表記の揺れの一つと考えることができる。同様にマンディのレシピ9件について類似を調べたものを表6に示す。表5、表6を見ると、似た食材がさまざまな語で表現されていることがわかる。

⁵加工食品における複合原材料の表示方法について、消費者庁、https://www.cao.go.jp/consumer/history/03/kabusoshiki/syokuhinhyouji/doc/k140605_shiryou3.pdf。

表6: 「マンディ」食材の表記の揺れ

コショウ	黒胡椒	黒コショウ	ペッパーシード
鶏肉	チキン	ひき肉	ラム
米	ご飯	バスマティ米	
唐辛子	トウガラシ	カブシカム	
玉ねぎ	中玉ねぎ		
生姜	ジンジャーペースト		
カルダモン	カルダモンシード		
クミン	クミンシード		
シナモン	シナモンスティック		
ニンニクペースト	ガーリックパウダー		
バター	マーガリン		
油	オイル		
水	お湯		

本研究では食材の記述に現れる表記の揺れを7種類に分類する(表7)。日本語の字種異なりでは、Wikipedia

表7: 食材の表記の揺れの種類

字種異なり	唐辛子/トウガラシ	黒胡椒/黒コショウ
同音異表記	バルガー/ブルガー/ブルグル	
外来語換言	鶏肉/チキン	油/オイル
細分類	米/エジプト米	コショウ/黒コショウ
状態等付記	米/ご飯/洗米	玉ねぎ/中玉ねぎ
加工食材	トマト/トマトジュース/ トマトペースト/トマトソース	
類似食材	バター/マーガリン	

で検索すると例えば「唐辛子」は香辛料、「トウガラシ」は植物種として別ページで説明されているなど、表記によって検索結果に違いがある。表4の材料欄にあるバルガーは、ブルガー、ブルグルなどさまざまなカタカナで表記されるため、外来語の表記の揺れとして問題となる。表7にある状態等付記とは、洗米や沸騰麺など、動作や状態を示す文字列が入っているものである。食材のアレルギー推定を行うには、まずこれらの表記の揺れを解消する必要がある。

表7の表記の揺れで特に問題となるのは類似食材である。他は原材料が同一であるため、同じ食材としてアレルギー推定をすればよい。それに対して、類似食材の例として挙げられているバターとマーガリンは、バターの主原材料は乳、マーガリンの主原材料は大豆油や牛脂、乳である。このように、材料欄に現れる類似食材は含有アレルギーが異なる場合があり、同じ食材として扱うかどうかの判定が課題となる。

6. まとめ

アレルギーを持つ人がアラビアの伝統料理を安心して食べられるよう、本研究では、アラビア語レシピからのアレルギー推定を行う。アラビア語レシピについて調査した結果、アラビア語から日本語への翻訳に問題があり、油を石油と翻訳した例など、翻訳誤りの修正の必要性があることがわかった。また、レシピには材料の原材料にアレルギーが含まれることがあり、アレルギーの判定には原材料の検索が必要だが、バルガーなどのアラビア料理独自の食材の検索が困難であるなど問題点があることがわかった。

参考文献

- [1] 花井俊介, 難波英嗣, 灘本明代, “主食材と主調味料を考慮した類似レシピクラスターリング,” DEIM2016, no.E2-7, 2016.
- [2] 佐藤貴之, 原島純, 小町守, “レシピに対する日英機械翻訳の誤り分析,” 情処研報, vol.2016-NL-228, no.3, pp.1-9, 2016.