

料理レシピの内容と構成が書き手の負担と読み手の効用に 与える影響の研究

The influence of the content and its arrangement of cooking recipes on the burden on
the writer and the utility of the reader

香川璃奈¹ 松原正樹¹ 宮田玲² 馬場雪乃¹ 山肩洋子³

Rina Kagawa¹, Masaki Matsubara¹, Rei Miyata², Yukino Baba¹ and Yoko Yamakata³

¹筑波大学

¹University of Tsukuba

²名古屋大学

²Nagoya University

³東京大学

³University of Tokyo

Abstract: Although the cooking recipe is a procedural document, it is desirable that the contents useful for the reader to make diverse decisions in combination with his / her own taste, such as proposals for alternatives and reasons for the cooking process, should be appropriately described. In this study, for the purpose of clarifying the technology necessary for appropriate support of document description, we clarified the influence of the content and its arrangement of cooking recipes on the burden on the writer and the utility of the reader.

1. 背景と目的

一般的に、読み手が有用と感じる文書の記載は、背景知識の異なる書き手にとって容易ではない。この問題を解決するために、行政、医療、技術文書などの多様な対象で記載支援ツールや構造化記載テンプレート[1]が提案される一方で、不適切な支援は記載情報の欠落や読み手の読解に必要な時間の延長[2]に結びつくことも報告されている。我々は、この問題の一因は、支援技術開発の際に、読み手はどのような文書を有用だと感じるのか、さらに、書き手はどのような文書記載に負担を感じるのか、という基盤的な議論が不十分なことにあると考える。

文書の構造は、どのような内容を、どのように配置するか、の2つの要素から成る[3]。そこで、読み手が有用と感じる文書の記載支援を行うための基盤的な知見として、(1)読み手が有用と感じる内容、(2)その内容がどのように配置されると読み手がより有用と感じるか、という2点を明らかにする必要があると考える。さらに、特に読み手が有用と感じる内容と配置の組み合わせが、従来の記載の慣習と異なる場合は、書き手にとっての負担も同時に考慮する必要がある。記載のための適切な指示を与えること

や書き手が記載しやすいデザインを検討することだけでなく、書き手の負担の程度に応じて、機械的・自動的な支援技術を開発することが重要となる。さらに、読み手がどのような記載内容と配置を有用だと感じるかは、読み手の背景知識や動機によっても変わりうる。

しかし、読み手が感じる有用性に寄与する文書に記載される内容とその配置が、読み手にとってどの程度有用でどの程度書き手に負担を書けるのか、さらに考慮すべき読み手の背景知識や動機も、全て経験的な知見であり、明らかにすることは難しい。

我々は、適切な文書記載支援の設計と技術的課題の抽出のための基盤的な検討として、読み手が有用と感じる文書記載とはどのようなものか、書き手に負担を与える文書記載とはどのようなものか、という2つの問いを立てた。本研究では読み手が感じる有用性と書き手の負担の両側面から、以下を行った。

- (1) 読み手に有用な記載内容の顕在化
- (2) 読み手に有用な記載内容の有無と配置が読み手の感じる有用性に与える影響の可視化
- (3) 読み手に有用な記載内容の有無と配置が書き手の負担に与える影響の可視化
- (4) 読み手が感じる有用性の詳細な検討

本実験では、読み手にとっての有用性の観点が重要かつ多様であり、さらに、書き手の負担軽減が求められるような文書ジャンルとして、料理レシピを対象とする。料理レシピ文書は確かに調理工程を記した手順書である。しかし読み手は、記載された調理工程を厳格に再現するとは限らず、料理レシピ文書を利用しながら、自分の好みや環境に応じて柔軟に自らの調理方法を決定する。またユーザー投稿型レシピサイトの世界的な普及により、記載方法についてフォードバックを受ける機会に乏しいアマチュアの記載者による両例レシピ文書を多様なユーザーが読む機会は世界的に増加している。料理レシピ文書のユーザーの嗜好と利用目的の幅は極めて広く、読み手の背景知識や動機を網羅的に検討することも容易ではない。

2. 実験用レシピデータと実験参加者

NHK みんなのきょうの料理[4] に 2020 年 6 月 1 日時点で公開されている料理レシピ 9,820 個をプロ記載レシピとして利用した。楽天レシピ のうち NII から楽天データセット[5]として提供されているレシピ 113,688 個を投稿型レシピとして利用した。実験参加者はランサーズで募集し、全ての被験者実験は web 画面上で行った。解析に利用したパッケージ等は以下の通りである。Python 3.7.2, NumPy 1.17.2, pandas 0.24.2, NetworkX 2.3, apyori 1.1.1, Community 1.0.0, scikit-posthocs 0.6.6, statsmodels 0.12.2, pingouin 0.3.10。

3. 読み手に有用な内容の顕在化

3.1 有用性に係る内容的特徴の収集と集約

読み手が役にたつと感じるという観点から、料理レシピに記載されるべき内容を明らかにした。

まず、予備的検討として、プロ記載レシピとユーザー投稿型レシピのどちらかがもう片方と比較して良い点を 40 人の実験参加者が自由記載したところ、プロ記載レシピが投稿型レシピよりも役に立つ点が 592 個、投稿型レシピがプロ記載レシピよりも役に立つ点が 558 個(いずれもユニーク数)取得された。このことから、プロ記載レシピと投稿型レシピのどちらかがより有用であるとは判断できないと考え、プロ記載レシピと素人投稿型レシピを混在させたデータセットを利用した。

ここで、料理レシピに記載されていると読み手が有用だと感じる内容は経験的な知見であり、読み手の動機に触れた研究は認められる[6]ものの網羅的に整理されているとは言い難い。そこで、比較を通じて、対象の特徴をより深く認識できる人間の認知

特性[7]に着目した実験を実施した(図 1)。

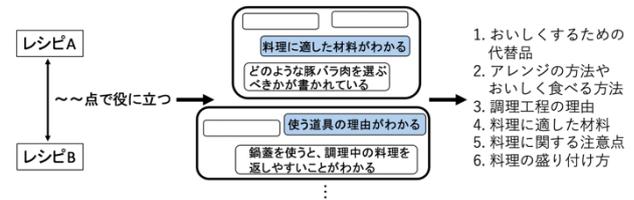


図 1: 読み手に役に立つ内容を顕在化させる手順の概略図

実験参加者はランダムに選択した 2 個のレシピ文書を読み、1 つ目のテキストがもう片方と比較して役に立つ点、および 2 つ目のテキストがもう片方と比較して役に立つ点をそれぞれ「～(の)点で役に立つ」という形式の自由記載で回答した。延べ 134 個のレシピを利用し 20 人が実験に参加した。一度に提示する 2 つのレシピ文書は、レシピに付与されている 3 層のツリー構造から構成されるカテゴリの小分類が一致するものからランダムに選択した。1 人あたり 20 対のレシピを比較した。得られた結果のうち「～(の)点で役に立つ」という記載方法を遵守した記載の「～」に該当する部分のみ抽出し、これを特徴候補と呼ぶ。

次に、特徴候補を類似性に基づいて集約することで読み手の役に立つ記載内容を明らかにしたい。しかし、下記のような類似性を機械的に判断する技術開発は困難である。

- 例:「加熱具合が詳細にわかる」と「鍋の火を止めるタイミングがわかる」が類似している。
- 例:「かぼちゃがグラム表示でも書かれている」と「キャベツの分量が、「〇コ」ではなく、「〇枚くらい」と書かれている」が類似している。

そこで、人手による特徴候補間の類似評価を行い、1 対の特徴候補間が類似しているか否かを多数決で決定することにした。ランダムに選択された 1 つの特徴候補に対して、ランダムに選択された 100 個の特徴候補を示し、類似しているものを全て選択するタスクを、実験参加者延べ 370 人が 25 問ずつおこなった。これにより、各特徴候補をノードとして、類似していると判断された候補のペアにエッジをはって、ネットワークを作った。このネットワークに louvain 法を適用してコミュニティ抽出を行い、各コミュニティが、役に立つ料理レシピに求められる要素と解釈することとした。さらに、各コミュニティを構成する特徴候補の中から、そのコミュニティを最も代表する特徴候補を 1 つ多数決で決定(同数の場合はランダムに選択)し、コミュニティ毎に選定された特徴候補を役に立つ料理レシピに求められる要素の名称とした。50 人の実験参加者が全てのコミュ

ニティに対して作業を行った。要素は、「盛り付けの記載がある」というレシピの記載内容についての要素と、「手順が細かく書かれており詳細にわかる」という表現についての要素に大別できる。要素の名称が「～がわかる」「～の記載がある」「～が書いてある」のいずれかで表記されている要素の「～」の部分のみ抜粋して、役に立つ料理レシピに求められる要素の中でも内容を示すものとみなした。それ以外は、表現についての要素とみなした。特徴候補のネットワークを定義する際には、1 対の特徴候補間を類似していると判断した実験参加者数が 4 人以上のときを、その特徴候補間は類似していると判断した。本研究では、抽出された内容を人間が記載するように支援すべきという結論もありうるため、抽出された内容についてのコミュニティ数が人間が短期記憶で記憶できる要素数とされている 5 以上 9 以下[7]になるように閾値を設定した。

延べ 616 個の回答から、426 個の特徴候補を取得し、13 個のコミュニティが抽出された。そのうち 6 つが内容の要素、7 つが表現の要素であった。内容の要素のうち、「使う道具の理由」のみ、研究の意図を踏まえて「調理工程の理由」に名称を変更した。内容の要素は以下の通りである。おいしくするための代替品、アレンジの方法やおいしく食べる方法、調理工程の理由、料理に適した材料、料理に関する注意点、料理の盛り付け方。以下それぞれ、代替品、アレンジ、理由、材料、注意点、盛り付け、と略す。抽出された 6 つの要素を以後、詳細項目と呼ぶ。

3.2 実際の料理レシピ文書における詳細項目の記載

各詳細項目が現状でどの程度記載されているか確認した。プロ記載レシピからランダムに選択した 100 個のうち、他の料理レシピを参照しているレシピを除外した 86 個のレシピを利用した。訓練されたアノテーター 2 名がマニュアルに基づき各詳細項目の記載にタグを付与した(各項目の記載例は表 1 を参照)。意見が一致しなかった箇所は話し合いにより結果を一意に決定した。各詳細項目について、記載が認められたプロ記載レシピの数は、代替品、アレンジ、理由、材料、注意点、盛り付けの順に、22 個、44 個、13 個、59 個、84 個、31 個であった。レシピ文書あたりの記載箇所の数の平均値と標準偏差は順に、0.36(0.75)、1.08(1.43)、0.16(0.40)、1.16(1.16)、3.53(2.61)、0.36(0.48)であった。この結果から、料理レシピは十分に構造化された手順書と解釈されることもあるが、読み手が有用と感じる記載が十分に記載されているとは言い難い現状が示唆された。

表 1：各詳細項目の記載例

代替品	ハーブの一種。赤い茎、葉脈が特徴。 なければルッコラ、マージョなどを使ってもよい。
アレンジ	水を加え、1～2 時間おく(ごく弱火にかけておいてもよい)。
理由	一度火を止めてアルミ箔と器を入れる。 アルミ箔を敷くと、熱くなった器を安全に取り出せる。
材料	魚 (切り身)1 切れ(80g) *さわら、たい、まながつおなど。
注意点	竹ぐしがスツと刺さる柔らかさになったら火を止め、そのまま冷ます。
盛り付け	6 を 5 の器に盛り、ベーコンを添え、イタリアンパセリを散らす。

3.3 詳細項目の妥当性検討

得られた詳細項目に対して、実際にレシピの利用者が各詳細項目を役に立つ項目として認識しているか確認する、質的な評価をおこなった。

1 つ目の質的評価として、料理レシピ(紙媒体でも可、記載・閲覧は問わない)の利用経験がある 200 人に対して、「役に立つ料理レシピのテキストにはどのような内容が記載されていますか？」という自由記載でのアンケートをおこなった。自由記載の中に、Step1-3 で抽出した要素が含まれているかどうか 1 名のアノテーターが確認した。その結果、各詳細項目について、役に立つレシピに記載されている内容として言及した人数の割合は順に、0.255、0.35、0.165、0.43、0.595、0.045 であった。6 つの詳細項目以外に、記載されているべき内容について言及された回答は見つからなかった。

2 つ目の質的評価として、抽出した 6 つの要素について、それぞれが記載されているレシピは役に立つと思うか、料理レシピ(紙媒体でも可、記載・閲覧は問わない)の利用経験がある 400 人に対して、10 件法(10：役に立つ～1:役に立たない)で調査を行った。各詳細項目が記載されている料理レシピは役に立つと感じるか 10 件法の平均値(標準偏差)は順に、8.39(1.65)、8.33(1.68)、8.16(1.80)、8.48(1.51)、9.02(1.22)、6.98(2.06)であった。

以上より、3.1 で得た詳細項目は妥当と判断した。

4. 詳細項目の記載の有無が読み手と書き手に与える影響

4.1 読み手が感じる有用性

3.2 節でタグを付与した各レシピ文書に対して、各詳細項目の記載を削除した実験用文書(以後、欠落文書と呼ぶ)を作成した。削除により助詞の接続が不自然になる、主語が欠落するなどの場合のみ最低限の表現の調整を行った。各詳細項目の記載の存在が、同一の読み手が感じる有用性を向上させることを確認するために、オリジナルの料理レシピ文書(以下、オリジナル文書)と欠落文書を同時に閲覧し、オリジナル文書と欠落文書のそれぞれについて、役に立つと思うか、という質問を 10 件法 (10 : 役に立つ ~ 1 : 役に立たない) で回答する実験を行った。このスコアを以後、有用性スコアと呼ぶ。実験参加者は、6 種類の詳細項目からランダムに選択された 1 種類について、文書の比較を行った。ただし比較する文書の違いなどの詳細については実験参加者に一切明らかにしていない。実験参加者 1 人あたりが比較する問題の数は詳細項目ごとに異なる。実験参加者の負担を鑑み、50 個以上のオリジナル文書に記載が存在した材料と注意点は、ランダムに抽出した 40 個のオリジナル文書と対応する欠落文書を対象とした。それ以外の詳細項目は該当する記載が存在した全てのオリジナル文書と対応する欠落文書を実験材料として用いた。各詳細項目について、オリジナル文書と欠落文書の有用性スコアの平均値が等しいという帰無仮説に対する有意差検定(t 検定)を行った。オリジナル文書と欠落文書のいずれかの有用性スコアの回答に欠損がある場合は、解析対象から除外した。延べ 400 人の実験参加者が作業した。

表 2: オリジナル文書(オリジナル)と欠落文書(欠落)の有用性スコアの平均値と標準偏差

	オリジナル	欠落	p 値	効果量
代替品	7.54(1.87) (n = 1,257)	7.34(1.82) (n = 1,257)	9.08e-10	0.11
アレンジ	7.03(1.89) (n = 2,879)	6.63(1.88) (n = 2,879)	3.79e-73	0.21
理由	7.61(1.94) (n = 1,105)	7.11(1.88) (n = 1,105)	5.35e-52	0.26
材料	7.05(1.93) (n = 3,046)	6.93(1.93) (n = 3,046)	9.41e-20	0.06
注意点	7.80(1.83) (n = 1,890)	7.23(1.73) (n = 1,890)	1.48e-76	0.32
盛り付け	7.22(2.21) (n = 2,161)	6.89(2.25) (n = 2,161)	4.63e-65	0.15

全ての内容において、記載があるオリジナル文書の方が欠落文書よりも有用性スコアが有意に高かった(表 2)ことから、3.1 節で取得した詳細項目は確か

に読み手が感じる有用性に寄与すると考える。特にアレンジ、理由、注意点の記載は有用性に寄与する程度が強いと考えられる(図 2)。

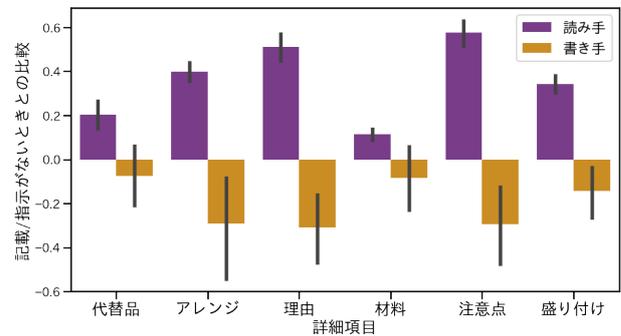


図 2: 詳細項目の記載が読み手の効用と書き手の負担に与える影響。読み手: 欠落文書と基準とした、欠落文書とオリジナル文書の有用性スコアの差。書き手: 記載の指示がない場合を基準とした、記載の指示がない場合とある場合の書きやすさスコアの差。

4.2 書き手の負担

各詳細項目を記載するよう書き手に指示することが書き手に与える影響を調査するため、自由に記載した場合と比較して、特定の詳細項目の記載を指示した場合に書き手が感じる負担の変化を測定した。

延べ 120 人の実験参加者が、プロのシェフが家庭料理のレシピ共有のために web 上に公開している調理動画[9]からランダムに選択した調理動画を見た上で、動画の内容のレシピを記載した。まず、何も指示を与えず、自由にレシピを記載してもらった。その後その実験参加者は、6 種の詳細項目のうちランダムに選択した 1 種類について、「このレシピをおいしくするための代替品があれば、上で書いたレシピに追記をしてください。」のような記載の指示にしたがって、レシピに追記する。各詳細項目は必ずしも全てのレシピに対して記載できるわけではないため、追記は義務とはしておらず、実験参加者が追記する内容がないと判断した場合には、指示がない状態で記載したレシピを新しい記載欄にコピーアンドペーストしてもらった。与える指示は動画ごとにランダムに指定した。それぞれの場合に、10 件法で書きやすさの主観的評価(10: 書きやすい ~ 1: 書きにくい)を計測した。以下、書きやすさスコアと呼ぶ。1 人あたり 5 つの動画について作業した。レシピが単純すぎて詳細項目に言及するまでもない動画ばかりでは実験にならないため、手順がある程度複雑で模倣も容易すぎない調理動画集を実験に利用した。

全ての詳細項目において、読み手が感じる有用性は有意に増加する一方で、書きやすさの主観的評価の平均値は全ての詳細項目で低下した(図 2)。指示し

ない時と比べた p 値(効果量)は順に 0.077 (0.06)、0.037 (0.08)、0.0012 (0.17)、0.320 (0.05)、7.13e-05 (0.13)、0.0075 (0.07)であり、代替品、アレンジ、材料、盛り付けの記載の指示は書きやすさを有意 ($p < 0.005$) には低下させなかった。しかし詳細項目を問わない傾向として、読み手にとって有用な詳細項目ほど書き手の負担が大きい(図 2)ことが推測される。

文書記載支援の1つの方向性は、記載すること自体を支援することである。4 節の結果から、代替品、アレンジ、材料、盛り付けを記載するよう書き手を促すことにより、読み手の有用性が向上する文書を書き手が負担なく作成できる可能性が考えられた。

5. 詳細項目の配置が読み手と書き手に与える影響

5 節ではまず、必ずしも調理工程そのものとは言えない詳細項目を記載する際に、どのような配置で実際に記載されたかを整理した。その配置が読み手の感じる有用性にもたらす影響と書き手に与える負担を測定した。

5.1 配置の分類

4.2 節で、詳細項目のいずれかの記載の指示のあとに作成された文書において、指示がない場合の文書と比較して記載が追加されていた場合に、記載はどのような配置であったか、1 名のアノテーター(筆頭著者)が整理した。4 節の成果物のうち、実験参加者の最後の一回のタスクの成果物、かつ 10 件法の回答が正しく記録されていたレシピは 117 個であった。この中で指示の前後で記載の追加があったのは 53 個であった。このうち、4 節の実験条件に起因すると想定される偏りとして、指示のあとに記載されたテキストに「追記」と書かれた上で指示に従う内容が記載されていたレシピが 4 個存在した。これらは検討対象から除外した。残りの 49 個の文書と、それぞれに対応する指示前の文書を比較したところ、詳細項目の記載の配置が表 3 の 5 種類に分類された。なお、(3)と(4)を共に含む文書が 1 個存在した。

5.2 詳細項目の配置と有用性

3.2 節でタグを付与した 86 個のオリジナルレシピ文書に対して、各詳細項目を 5.1 節で明らかにした 5 種類の配置に再配置した実験用文書(図 3)を作成した。1 つのレシピから詳細項目ごとに最大 5 種類の実験用文書が作成される。実験意図を鑑みて、詳細項目を再配置した実験用文書がオリジナル文書と全く同じ場合には、実験用文書を作成しなかった。

4.1 節と同様に、検討する配置ごとに 400 人の実験

参加者が、オリジナル文書と各実験用文書との比較を行い、オリジナル文書と実験用文書のそれぞれについて、役に立つと思うか、という質問に 10 件法(10:役に立つ ~ 1:役に立たない)で回答した。さらに、同じ内容の配置間の違いをより詳細に検討するために、実験参加者の全作業へのメンタルワークロードを日本語版 NASA-TLX [10] で計測した。実験参加者の負担を鑑み、NASA-TLX と高い相関が知られている Raw-TLX[11]を利用した。また、認知的負荷の定義[12]を参照し、行動指標として比較ごとの作業時間も計測した。3,600 秒以上の作業時間は外れ値として除外した。

表 3: 配置の分類。見出しと視覚的符号を合わせて強調とした。

	位置	改行	強調	レシピ数
(1)	手順ごと	あり	あり	10
(2)	手順ごと	あり	なし	14
(3)	手順ごと	なし	なし	3
(4)	最後にまとめて	あり	あり	16
(5)	最後にまとめて	あり	なし	7

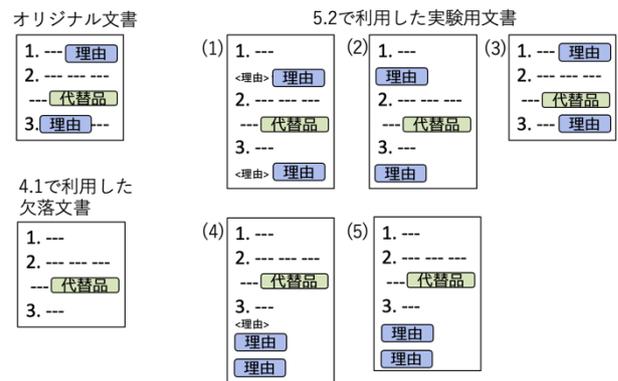


図 3: 各配置に基づいた実験用文書の概略図。理由を対象にした場合の例。

各詳細項目について、各配置の実験用文書のスコアの平均値がオリジナル文書の有用性スコアの平均値と等しいという帰無仮説に対する有意差検定(t 検定)を行った。また、各詳細項目において配置間の有用性スコアの多重検定(Games-Howell 検定)も行った。日本語版 NASA-TLX と作業時間は、オリジナル文書と実験用文書それぞれで取得していない。そこで、各詳細項目について、配置によって読み手の負担が有意に($p < 0.001$)変化するか確認するため、各詳細項目において配置間のスコアの多重検定を行った。日本語版 NASA-TLX には Games-Howell 検定、作業時間には Steel-Dwass 検定を用いた。

オリジナル文書の有用性スコアと比較して、代替品の配置(4)、理由の配置(1)、材料の配置(1)と(4)、盛り付けの配置(1)と(4)において、有用性スコアが有意($p < 0.001$)に上昇した。以上のうち、同じ詳細項目における配置間の多重検定により、同じ詳細項目の中で他の配置と比較して有意($p < 0.001$)に有用性スコアまたは NASA-TLX または作業時間が変化した配置は認められなかった。

注意点の読みやすさスコアが配置を問わず低下したことは、3.2 節において「豚肉を鍋に入れて約 10 分煮込む」の「約」などを全て注意点とタグ付けして再配置したことが原因と考えられた。

5.3 詳細項目の配置と書き手の負担

4.2 節と同様の実験を行った。レシピ記載の指示を与える際に、5 種類の配置のいずれかで記載するように指示を行った。なお、見出しや視覚的符号[13]を利用する配置では、見出しの利用を指示した。

配置ごとに実験参加者の料理レシピ記載経験の有無を問わない 100 人の実験参加者が参加した。実験参加者が指示の内容を適切に把握して記載するために不要な混乱を招かないように、参加者毎に指示する配置は統一したが、指示の内容は 4.2 節と同様にランダムに変化させた。実験参加者は動画ごとに日本語版 NASA-TLX を計測した。動画ごとの作業時間も計測し、7,200 秒以上の作業時間は外れ値として除外した。各詳細項目について、指示があるときとない時とで各スコアの平均値が等しいという帰無仮説に対する有意差検定を行った。書きやすさスコアには t 検定、日本語版 NASA-TLX には Steel-Dwass 検定、作業時間には Games-Howell 検定を用いた。

全ての配置と詳細項目の組み合わせで、書きやすさスコアは指示なく記載した場合と比較して、平均して減少していた(図 4)。5.3 節で、オリジナル文書と比較して有意に点数が高かった詳細項目と配置の組み合わせは全て、有意($p < 0.005$)に書きやすさスコアが低下しなかった。各詳細項目における配置間の多重検定で、他の全ての配置よりも有意に書きやすさスコアまたは NASA-TLX または作業時間が変化した配置は認められなかった。

文書記載支援の方向性として、記載された文節を適切な配置に再配置する、あるいは適切な配置への記載を促す、などの方向性も考えられる。5 節の結果から、代替品、理由、材料、盛り付けは、人間が経験的に記載している配置で記載されたオリジナル文書よりも、読み手が有用だと感じる配置が存在することがわかった。それぞれ適切な配置に再配置を促す支援が有効であると考えられ、さらにその配置での記載を書き手に負担なく促すことができる可能性

も考えられた。詳細項目ごとに適した配置は異なる一方で、見出しの利用が、読み手の効用を増加させ書き手の負担を増加させない、共通した特徴であることが示唆された。

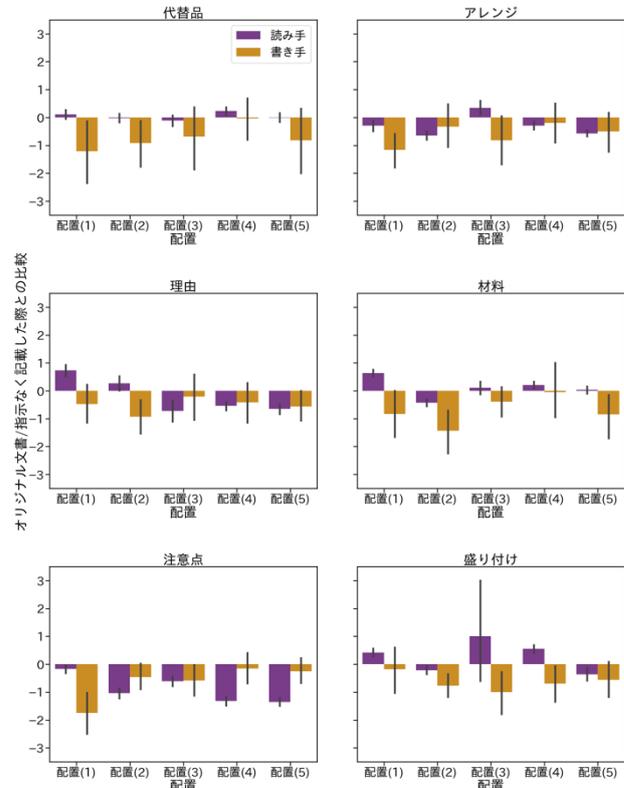


図 4：詳細項目と配置の組み合わせが読み手の効用と書き手の負担に与える影響。読み手：オリジナル文書と、各配置の実験用文書との有用性スコアの差。書き手：指示なく記載したときと指示して記載した際の書きやすさスコアの差。

6. 読み手が感じる有用性の詳細分析

読み手が感じる有用性を構成しうる要素は多岐にわたる。そこで、システム開発の要求仕様書が満たすべき品質として国際的に標準化されている IEEE Std830-1998 [14]を参照して決定した質問項目(表 4、以下、詳細評価項目と呼ぶ)を利用して、読み手が感じる有用性を詳細に検討した。各詳細項目の各詳細評価項目(10 件法)のスコアについて、実験用の文書のスコアの平均値がオリジナル文書のスコアの平均値と等しいという帰無仮説に対する有意差検定(t 検定)を行った。実験用の文書として、4.1 節で用いた欠落文書と、5.2 節で用いた実験用文書をそれぞれ利用し、4.2 節および 5.2 節と同様の実験に、検討する配置ごとに 400 人ずつの実験参加者が参加した。

全ての詳細項目において記載の有無で比較を行うと、詳細評価項目のうち「曖昧な記載がない」と「調

理に必要な情報が全て記載されている」のスコアが、欠落文書と比較してオリジナル文書の方が有意 ($p < 0.001$)に高くなった(図 5)。アレンジと理由においては「調理に関係ない記載はない」スコアが有意 ($p < 0.001$)に低くなり、アレンジと理由の記載は、レシピ文書の有用性には寄与する(4.1 節参照)ものの、調理する上で関係がない記載だと読み手が感じる事が示唆された。代替品のみ「内容を自由に変更できる」スコアが有意 ($p < 0.001$)に高く、代替品の記載は、読み手が記載内容を自由に変更して調理できることに寄与すると示唆された。

表 4：読み手が感じる有用性の詳細評価項目。全て「～と思いますか」の形式で質問をおこなった。

(問 1) 調理する上で関係ない記載が書かれている
(問 2) 曖昧な記載がある
(問 3) 調理に必要なことが全て書かれている
(問 4) 矛盾がある
(問 5) 記載内容を自由に変更して調理できる

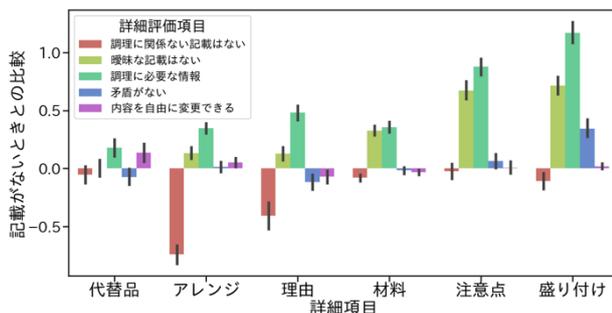


図 5：各評価項目の記載の有無による各詳細評価項目のスコアの変化

配置を変化させた場合、4.1 節と 4.2 節で明らかになった、読み手の有用性を他の配置と比較しても有意に向上させ、かつ書き手の書きやすさで有意な低下を認めない詳細項目と配置の組み合わせの全てで、オリジナル文書よりも有意($p < 0.001$)に高値を認めた詳細評価項目は存在しなかった。「調理に必要なことが全て書かれている」スコアは、有意($p < 0.001$)に高値を認めたその組み合わせの数をもっとも多く、理由の配置(1)、材料の配置(1)と(4)、盛り付けの配置(1)であった。

4 節と 5 節の結果とあわせると、読み手が感じる有用性と特に類似する傾向を示した詳細評価項目は、詳細項目や配置に依らず「調理に必要な情報が全て記載されている」と考えられる。読み手が何を調理に必要な情報と考えるかは、読み手の背景知識や動機に依存するため、今後は、読み手の背景知識や動

機を取得した上での検討が必要である。

なお、盛り付けの配置(3)のエラーバーが大きい(図 6)原因は、実験用文書とオリジナル文書が全く同一の場合に実験対象外としたところ、実験対象となる文書が 1 通りになってしまったことと推測される。

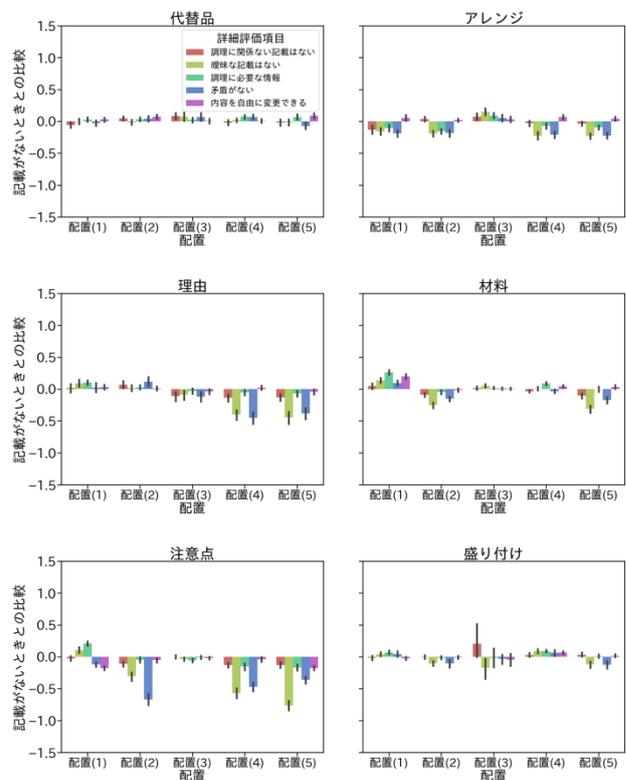


図 6：オリジナル文書を基準とした際の、各詳細項目の配置における各詳細評価項目のスコアの変化。

7. 今後の課題

本研究では 3.2 節のタグ付けの際に、人間の読み手がどの部分をどのように考えるか、個別の文生成をはじめとする自然言語処理技術に依存しない知見を得るという目的を重視した。その結果として、タグが付与された範囲が、箇条書きを含む単語、フレーズ、文と多様である。代替品や材料は「鶏ささ身 * 鶏むね肉でもよい」のように単語単位の記載が多いため、単語間の類似度などを定義して代替として利用されやすい材料を知識として獲得しておくことなどが有効とも考えられる。しかし、そもそもプロ記載レシピですら既存のレシピ文書あたりの記載箇所数が平均して 0.36 または 1.16 であり、利用可能なデータの少なさが問題になる可能性がある。

同じ内容でも冗長に感じたり簡潔に感じたりする、表記の問題を不問としたことは本研究の限界である。特に 6 節で示された、読み手が感じる有用性と類似した傾向を示す「曖昧な記載はない」という詳細評

価値指標は、同じ内容であっても表記の問題にも依存することが予想される。既存の料理レシピ文書をもとに自動で文生成を行うことが主流である今日の技術的背景から考えて、どのような表記のレシピが読み手に好まれるのか議論することは記載支援において重要だと考える。本研究で取得した、同じ料理動画上に基づく複数の料理レシピを利用して、表記の面でもどのような文書にどのような反応があるのか検討を実施する予定である。

特定の料理レシピサイトの文書を用いていること、また、日本語の文書のみを利用していることによる未知の偏りが生じている可能性は否定できない。他言語のレシピ文書においても、内容と配置が読み手の効用と書き手の負担に影響を及ぼすのか検討を行いたい。

8. まとめ

本研究では、読み手にとって真に有用な文書記載支援技術を検討するために、料理レシピ文書において、読み手が有用だと感じる内容を6種類抽出した。その中で、代替品、アレンジ、材料、盛り付けが記載されていると読み手が有用と感じる上に、それを記載するよう指示された書き手の負担も、記載を指示されない場合と比較して増加しないことを可視化した。さらに、文書の記載内容の配置が読み手の効用と書き手の負担に与える影響を測定し、可視化した。これにより、詳細項目ごとに適した配置は異なる一方で、見出しの利用が、読み手の効用を増加させ書き手の負担を増加させない、共通した特徴であると示唆された。また、読み手が感じる有用性を構成する要素は多岐にわたるが、「調理に必要な情報が全て記載されている」という要素が、有用性ともっとも類似する傾向を示した。

今後は文生成や記載の再配置のための技術開発を行うと同時に、記載支援が有効だと考えられる多様な文書ジャンルに同様の検討を拡張したい。

倫理的配慮と謝辞

筑波大学医学医療系医の倫理委員会より承認を得ています。JST 未来社会創造事業(JP19211284)、JSPS 科研費(19K19347)、筑波大学 F-MIRAI の支援を受けています。アノテーションでは住本美香さんと大久保憲子さんにご尽力いただきました。

参考文献

- [1] Weed, L. L. *Medical Records, Medical Education and Patient Care: The Problem Oriented Record as a Basic Tool*. Cleveland: Press of Case Western Reserve University. (1970).

- [2] Varpio, L., Rashotte, J., Day, K., King, J., Kuziemsy, C., & Parush, A. The EHR and building the patient's story: A qualitative investigation of how EHR use obstructs a vital clinical activity. *Int J Med Inform*, 84(12), pp. 1019-1028. (2015)
- [3] Reiter, E., & Dale, R. *Building Natural Language Generation Systems*. Cambridge: Cambridge University Press. (2000)
- [4] みんなのきょうの料理 <https://www.kyounoryouri.jp/> (2021年2月20日閲覧)
- [5] 楽天株式会社: 楽天レシピデータ. 国立情報学研究所情報学研究データリポジトリ. (データセット). <https://doi.org/10.32130/idr.2.4> (2019)
- [6] Shidochi, Y., Takahashi, T., Ide, I., & Murase, H., Finding replaceable materials in cooking recipe texts considering characteristic cooking actions. *Proceedings of the ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities*. pp. 9-14. (2009)
- [7] Gentner, D., & Markman, A. B. Structure mapping in analogy and similarity. *Am Psychol*, 52(1), pp. 45-56. (1997)
- [8] Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), pp. 81-97. (1956)
- [9] NHK らいふ <https://www.nhk.or.jp/lifestyle/cooking/index.html> (2020年9月1日閲覧)
- [10] Miyake, S., & Kumashiro, M. Subjective mental workload assessment technique: An introduction to NASA-TLX and SWAT and proposal of simple scoring methods, *Japanese Journal of Ergonomics*, 29, pp. 399-408. (1993)
- [11] Hart, S. G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 50(9), pp. 904-908. (2006)
- [12] Sweller, J., Van Merriënboer, J. J., & Paas, F. G. Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), pp. 251-296. (1998)
- [13] Lorch, R. F., Jr., & Lorch, E. P. Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology*, 87(4), pp. 537-544. (1995)
- [14] Software Engineering Standards Committee of the IEEE Computer Society. *IEEE recommended practice for software requirements specifications*. 830. IEEE. (1998)