

コスメ評価サイトにおけるユーザ使用感の時間性分析

合田 美生璃[†] 奥 健太^{††} 角谷 和俊[†]

[†]関西学院大学総合政策学部メディア情報学科 〒669-1337 兵庫県三田市学園2丁目1番地

^{††}立命館大学情報理工学部情報コミュニケーション学科 〒525-8577 滋賀県草津市野路東1丁目1-1

E-mail: [†] {dgx25127, sumiya}@kwansei.ac.jp, ^{††} oku@fc.ritsumei.ac.jp,

あらまし 近年、コスメ商品に特化したレビューサイトが増加している。これらのサイトでは、ユーザがコスメ商品を実際に使用した際の評価値、評価文を書き込むことができ、他の消費者に情報共有することが可能である。コスメは、自分の肌に直接使用するからこそ、商品を購入する際には、レビューが重要な役割を果たしている。しかし、実際のコスメ評価サイトでは、レビュー文が多く存在するため、消費者の望みに合った商品を見極めることは困難である。本研究では、コスメ評価を正確に抽出することを目的とし、時間的要素を含むレビュー文の分析を行う。

キーワード コスメ, レビュー分析, 時間性

1. はじめに

近年、消費者が商品やお店に対しての評価を情報共有することができる口コミサイトが多く存在する。特にコスメ商品においては、コスメに特化したレビューサイトが増えつつある。その理由として、コスメ商品は、自分の肌に直接使用するものであり、商品購入の際にレビューサイトが重要な役割を果たしているからであると考えられる。

コスメ商品のレビューサイトとしては、@cosme や MAQUIA、yahoo! beauty 等があり、幅広い年代の女性を使用している。しかしこれらのコスメ評価サイトには多くのレビューが存在する。その為、自分に合っている商品や、自分と似たようなユーザを見つけることは困難である。また、現在@cosme では、肌質や年齢が類似しているユーザを見つけることは可能だが、本当に同じ悩みを持つユーザを探すことは困難であり、類似ユーザが使用している商品を見つけることも難しい。

@cosme でのコスメレビュー文は、商品を使用した結果、その商品がユーザにとってポジティブなものであった、もしくはネガティブなものであったかが書かれている。このような情報は商品の評価または購入する上で大変重要な情報である。これまでの研究では、レビュー文に出てくるコスメ特有単語や評価のみで研究が行われていた。しかし、このような情報だけではそのコスメ商品のより細かい特性やユーザと商品の相性を判断することが難しい。なぜなら、コスメ商品は、時間とともに効果が現れたり、時間経過によって商品の特性が現れたりする場合があるからである。そこで我々は、コスメレビュー文の中に含まれるコスメ使用感の時間的要素が必要であると考えた。



図1 時間経過によるコスメ評価の変化

コスメレビュー文には図1のように、“くすみ”や“ふっくら”、“透明感”等、多くのコスメ特有単語を含むコスメ評価が記述されている。1つのレビュー文においてこのような記述は、ユーザが商品を使用した直後のコスメ評価が、時間経過と共に変化していく場合がある。

本研究では、「コスメレビュー文は、時系列を踏まえた方がコスメ評価或使用感をより正確に捉えることが可能である。」という仮定の下、分析を行っていく。

そして、今回の研究では、時間的要素を特にコスメ商品“使用直後”と“時間経過後”の二つに着眼点を置いた。

使用直後のコスメ評価に着目した理由は、商品を使用した直後の使用感を見ることで、その商品はユーザにとってどのような仕上がり、または感触になるものであるかを判断することができるからである。そのため、使用直後のコスメ評価を見ることは、商品自体の特徴を捉えやすいということになる。

さらに、時間経過後のコスメ評価に着目することで、より詳しい商品の特性、またユーザと商品の相性を捉えることができる。特に、スキンケア商品だと、何日か使用した後に自分の肌の結果が見られることが多いため、時間経過後のコスメ評価は重要な要素となってくるのだ。

以下、第2章ではコスメ商品推薦やレビュー分析についての関連研究の紹介を行う。第3章では予備実験について、第

4章ではパターン分類手法、第5章ではデータ分析について説明する。そして第6章ではまとめと今後の課題について述べていく。

2. 関連研究

本章では、コスメ商品推薦やレビュー文解析に関する関連研究について紹介していく。

近年、お店や映画、商品など、様々なものに対して口コミサイトが増加している。口コミは消費者が実際に体験した事実に基づき書かれるため、他の消費者がお店に行く際や商品を購入する際の参考にする場合に役立つ。特に最近では、コスメ商品の口コミに特化した@cosmeが有名である。@cosmeは、日本最大のコスメレビューのデータベースを保持している。

松波ら[2]は、コスメアイテムに関する口コミおよびそのレイティングを分析し、コスメアイテムに対するポジティブ・ネガティブ表現に関する辞書を構築した。この辞書を用いて、評価関連ごとの自動レイティングを行い、口コミの可視化を行った。そして、既存の口コミに対するレイティングのみならず、未知の口コミデータに対するレイティングも行った。本研究では、コスメアイテムに対するポジティブ・ネガティブ表現に加えて、時間的要素を含むコスメ評価表現について分析する。これにより、より正確で有用性のあるコスメ評価表現を抽出することを目的とする。

安部ら[3]は、ユーザに適したレビューを提示するため、自分とユーザの属性の近さを分析した。化粧品ブランドにおけるレビューの関係性の類似度を算出した後、これを元にユーザに提示するシステムを提案した。ここでは、ブランド・商品・レビューの関係を明らかにし、レビュー文書に含まれるユーザのプロファイルなども考慮に入れ、分析している。

中野ら[4]は、レビューの要約として、レビューの文章中の商品のある部分を表す表現を属性、属性の評価値や主観評価を表す表現を意見とし、二つが対応する属性-意見ペアとして抽出する手法を提案している。そして、これを抽出するために機械学習を用いている。本研究では、時間表現とコスメ評価の二つの表現を抽出しているため、機械学習を利用することも可能である。

松尾ら[5]は、レビューが評価値を決定した根拠となる商品属性を提示する手法を提案している。商品の機能や特徴といった商品属性に関する評価文を抽出する。そして、抽出した評価文から属性ごとのスコアを算出し、重回帰分析を行い、その結果を利用してレビュー評価の根拠となっている商品属性を提示している。

さらに松尾ら[6]は、レビューの評価値と評価表現を含む文に着目し、レビュー文書から重要文を自動で選択する手法を提案している。レビュー文書から商品の属性に関する評価表現が記述されている箇所を抽出し、それが含まれる文に対して、商品の評価を表す単語のスコアや文節数などの重要度を表す指標を算出している。また、ここでは実験結果から名詞

や動詞を評価語として追加することを検討している。しかし本研究では、コスメレビュー文を形態素解析し、品詞分解をした結果、品詞によるコスメ評価の抽出は難しいと判断した。

このように、@cosmeに限らず、口コミレビュー解析について、ユーザの様々な観点からレビューを分析するという研究が多くされていた。松波らは、@cosmeのサイトのレイティング方式を生かした研究であった。本研究では、コスメ商品の時間経過による変化に着目し、コスメ評価サイトのみ抽出可能な観点を用いてレビュー分析を行う。

3. 予備実験

本章では、レビュー文をコスメ商品の使用直後と時間経過後に着目して分類した、パターン分けについて述べる。さらに、データから見えてきたさらなる時間経過後のパターンについて述べていく。

3.1 使用直後と時間経過後に着目したパターン

本節では、使用直後と時間経過後に着目した際に分類されるパターンについて述べる。

コスメレビュー文から、コスメ商品の使用直後と時間経過後に着目した際、レビュー文は16パターンに分けることができることが分かった(表1)。

表1 コスメレビュー文における時間表現および
コスメ評価表現の出現パターン

		時間経過後				
		時間表現	t	t	f	f
	時間表現	コスメ評価	t	f	t	f
	使用直後	T	T	TT-tt(22)	TT-tf(1)	TT-ft(8)
T		F	TF-tt(5)	TF-tf	TF-ft	TF-ff
F		T	FT-tt(18)	FT-tf	FT-ft(7)	FT-ff(0)
F		F	FF-tt(72)	FF-tf	FF-ft(27)	FF-ff

※()内の数字はデータ数を表す

T,Fは使用直後に関して該当表現が存在する(T)、存在しない(F)を表す。t,fは時間経過後に関して該当表現が存在する(t)、存在しない(f)を表す。これら2種類ずつを組み合わせ、コスメ商品使用直後・時間経過後のパターンとしてそれぞれ4種類に分類することができる。さらに、これら(使用直後4種類×時間経過後4種類)を組み合わせることで、図のような16パターンに分類することができた。

グレー部分はそのパターンに該当するレビュー文が存在せず、かつ存在しても今回の研究では不必要な表現であるため、省く。そのため、今回は全10パターンでコスメレビュー文を分類した。

表2 使用直後および時間経過後に関する表現の例

パターン	使用直後	時間経過後
TT-tt	塗った直後:カサカサ	少し時間が経つ:テカリ
TT-tf	つけた後:サッパリ, ベタつかず	使って2週間:
TT-ft	仕上がり:肌密着, 良く	—; 崩れにくく
TT-ff	使用感:サラッと, 肌染み込んでいく	—; —
TF-tt	朝と夜, 使用後:—	昼過ぎ:乾燥せず, 肌すべすべ, ハリ
FT-tt	—;カバー力ある, 重たく感じません	時間が経つと:テカリ
FT-ft	—;サラサラなる	—;乾燥しなかった, 崩れ少なく
FT-ff	具体例なし	具体例なし
FF-tt	—;—	夕方:くすまなく, 透明感
FF-ft	—;—	—;柔らかく, ふっくらしてきました

※各セルの表記は、「時間表現 ; コスメ評価表現」を表す。

「—」は表現が含まれていないことを表す。

(a) パターン TT-tt

パターン TT-tt は、[使用直後の時間的要素あり, コスメ評価あり], かつ[時間経過後の時間的要素あり, コスメ評価あり]のパターンである。

例えば、「塗った直後はちょっとカサカサに見えるくらいなのに、少し時間が経つとテカります」という文章がある。この文からキーワードを抜き出すと、「塗った直後, カサカサ, 少し時間経つ, テカリ」と表すことができる。そして、「塗った直後, カサカサ」は使用直後に分類され、「少し時間経つ, テカリ」は時間経過後に分類される。さらに、「塗った直後」「少し時間が経つ」は時間的要素、「カサカサ」「テカリ」はコスメ評価に分けられる。

(b) パターン TT-tf

パターン TT-tf は、[使用直後の時間的要素あり, コスメ評価あり], かつ[時間経過後の時間的要素あり, コスメ評価なし]のパターンである。

例として、「使ってみて2週間程です。つけた後もサッパリしていてベタつかずいい感じです！」という文がある。キーワードは、「使って2週間, つけた後, サッパリ, ベタつかず」である。さらに、「使って2週間」は時間経過後に分類され、「つけた後, サッパリ, ベタつかず」は使用直後に分類される。そして、「使って2週間」「つけた後」は時間的要素、「サッパリ, ベタつかず」はコスメ評価となる。

(c) パターン TT-ft

パターン TT-ft は、[使用直後の時間的要素あり, コスメ評価あり], かつ[時間経過後の時間的要素なし, コスメ評価あり]のパターンである。

「肌にピタッと密着して仕上がりも良く崩れにくくなります」という文では、キーワードとして「肌に密着, 仕上がり, 良く, 崩れにくく」が抜き出される。そして、「肌に密着, 仕上がり, 良く」は使用直後に分類され、「崩れにくく」は時間経

過後に分類される。「崩れる」という言葉は、コスメ商品を使用する際、ある程度の時間が経過してから起こる現象であるため、時間経過後に分類する。さらに、「仕上がり」はコスメ商品を使用直後に使われる言葉であるため、時間的要素に分類され、「肌に密着, 良く」「崩れにくく」はコスメ評価に分類される。

(d) パターン TT-ff

パターン TT-ff は、[使用直後の時間的要素あり, コスメ評価あり], かつ[時間経過後の時間的要素なし, コスメ評価なし]のパターンである。

「サラッとした使用感で、肌に染み込んでいくのが分かります」という文では、キーワードは「サラッと, 使用感, 肌染み込んでいく」となる。そして、これらのキーワードは全て使用直後に分類される。また、「使用感」は使用直後を表している言葉なので時間的要素、「サラッと, 肌染み込んでいく」はコスメ評価となる。

(e) パターン TF-tt

パターン TF-tt は、[使用直後の時間的要素あり, コスメ評価なし], かつ[時間経過後の時間的要素あり, コスメ評価あり]のパターンである。

例として、このような文章がある。「夜と朝に使用した後の昼過ぎ、いつもは肌が乾燥してくる時間帯にも乾燥せず、それどころか肌がすべすべしているようでビックリです。代謝が良くなったかのような肌ものハリも感じられました。」2文で時間の流れとコスメ評価を表している。ここでのキーワードは、「夜と朝, 使用後, 昼過ぎ, 乾燥せず, 肌すべすべ, ハリ」である。「夜と朝, 使用後」は使用直後に分類され、「昼過ぎ, 乾燥せず, 肌すべすべ, ハリ」は時間経過後に分類される。さらに、「夜と朝, 使用後」「昼過ぎ」は時間的要素となり、「肌すべすべ, ハリ」はコスメ評価となる。

(f) パターン FT-tt

パターン FT-tt は、[使用直後の時間的要素なし, コスメ評価あり], かつ[時間経過後の時間的要素あり, コスメ評価あり]のパターンである。

例として、「カバー力もばっちりあるのに、私は重たく感じませんでした。時間がたてば必ずテカる私の肌、もちろんこれを使ってもテカります。笑」という文がある。キーワードは、「カバー力ある, 重たく感じません, 時間たてば, テカリ」である。ここでは「カバー力ある, 重たく感じません」は使用直後に分類され、「時間たてば, テカリ」は時間経過後に分類される。さらに、「時間たてば」は時間的要素、「カバー力ある, 重たく感じません」「テカリ」はコスメ評価となる。

(g) パターン FT-ft

パターン FT-ft は、[使用直後の時間的要素なし, コスメ評価あり], かつ[時間経過後の時間的要素なし, コスメ評価あり]のパターンである。

例として、「サラサラになるけど、乾燥はしなかったですし、崩れも少なく嬉しいです。」という文がある。ここでのキーワ

ードは、「サラサラなる、乾燥しなかった、崩れ少なく」である。さらに、「サラサラなる」は使用直後に分類され、「乾燥しなかった、崩れ少なく」は時間経過後に分類される。「乾燥する」という言葉はパターン TT-ft で述べた「崩れる」と同様に、ある程度の時間が経過してから起こる現象であるため、時間経過後に分類する。そして、これらのキーワードは全てコスメ評価に分類される。

(h)パターン FT-ft

パターン FT-ft は、[使用直後の時間的要素なし、コスメ評価あり]、かつ[時間経過後の時間的要素なし、コスメ評価なし]のパターンである。

このパターンは、今回抽出したデータの中に例はなかったが、コスメ商品の仕上がりや使用感を表すコスメ評価のみが書かれているレビューは必要であると判断したため、このパターンも考え、分類を行った。

(i)パターン FF-tt

パターン FF-tt は、[使用直後の時間的要素なし、コスメ評価なし]、かつ[時間経過後の時間的要素あり、コスメ評価あり]のパターンである。

例として、「夕方にくすむのが気になるがコレは全然くすまなく、ずっと綺麗な透明感を保ってくれています。」という文がある。キーワードは、「夕方、くすまなく、透明感」となる。さらに、これらは全て時間経過後に分類される。そして、「夕方」は時間的要素、「くすまなく、透明感」はコスメ評価となる。

(j)パターン FF-ft

パターン FF-ft は、[使用直後の時間的要素なし、コスメ評価なし]、かつ[時間経過後の時間的要素なし、コスメ評価あり]のパターンである。

例として、「毛穴ケアに良いとのことを使い始めたら肌が柔らかくふっくらしてきました。」という文がある。ここでのキーワードは、「柔らかく、ふっくらしてきました」である。これらは両方時間経過後、かつコスメ評価に分類される。

3.2 さらなる時間経過後のパターン

本節では、データを分類した際に見えてきたさらなる時間経過後のパターンについて述べていく。

3.1 節では、使用直後と時間経過後の二つの観点からキーワードを抜き出し、分類した。しかし、コスメレビュー文にはそれだけではなく、さらなる時間経過後について書かれている文章も多く見られた。

表3 さらなる時間経過後を含む表現の例

パターン	使用直後	時間経過後	さらなる時間経過後
FF-tt	ー	朝起きて;しっとり	日中;粉吹きなく
TT-tt	付けたて;サラサラ	馴染ん;ツヤ	夕方;化粧崩れ汚くならない

(A)パターン FF-tt-tt

パターン FF-tt-tt は、「朝起きてても肌はしっとりです。日中の粉拭きもなくこの冬は特に助けてもらってます。」という文章がある。キーワードは、「朝起きて、しっとり、日中、粉吹きなく」となる。この文章は、スキンケア商品でのレビュー文であり、この文の前に、「乾燥肌のわたしはスキンケアの最後に2滴。」という文が記述されていたため、寝る前にコスメ商品を使用したことが分かる。そのため、「朝起きて、しっとり、日中、粉吹きなく」は全て時間経過後に分類される。

(B)パターン TT-tt-tt

パターン TT-tt-tt は、「付けたてはサラサラ、時間が経つと馴染んでツヤがでます。夕方の化粧崩れが汚くならないのも嬉しい♪」という文章がある。この文章のキーワードは、「付けたて、サラサラ、馴染ん、ツヤ、夕方、化粧崩れ汚くならない」となる。さらに、「付けたて、サラサラ」は使用直後に分類、「馴染ん、ツヤ、夕方、化粧崩れ汚くならない」は時間経過後に分類される。さらに見ていくと、「夕方、化粧崩れ汚くならない」この二つは、馴染んでツヤがでた後に起こった現象であるため、さらなる時間経過後に分類される。このように、コスメレビュー文では、使用直後と時間経過後の二つの軸のみならず、さらなる時間経過後のコスメ評価が書かれているものも多く見られた。また、スキンケア商品のレビュー文においては、商品を使用する前の肌状態についても書かれている場合が多かったため、今後は使用以前から使用後の変化について着目して分類する必要性がある。

4. パターン分類手法

本章では、コスメレビュー文を使用直後と時間経過後の2つを軸としてパターン分類をする際に用いた方法について述べていく。

はじめに、@cosme からデータを取る際、「レビューに時間的要素が含まれているか」を基準とした。そして、時間的要素が含まれているレビューのみをデータとして保存した。ここでの“時間的要素”とは、ユーザがコスメ商品を使用してから経過時間を表すものである。例えば、「使用した時もちりりしていた肌が、時間が経つとかさかさしてきた。」という文章のように、時間経過によるユーザ使用感の変化を表しているもののことを指す。

このようにして、コスメのカテゴリ別商品アイテム3種類(クリーム、化粧下地、口紅)のロコミから、該当する文章が含まれるロコミを抽出したところ、このような結果が得られた(表4)。

表4 時間表現を含むレビュー

商品	調べたレビュー	該当レビュー	%
クリーム	154 件	73 件	47.4 %
化粧下地	100 件	48 件	48%
口紅	80 件	34 件	42.5 %

3 種類とも、全体の 40%以上の口コミに時間的要素を含むユーザ使用感が書かれていたことが分かり、この観点を利用することは、分析する際に有用であることが言える。

特に、クリームや化粧下地は 50%近い数字であったため、今回は商品カテゴリをスキンケアとベースメイクに絞ってレビューを抜き出し、保存した。

次に、保存したレビューを利用し、形態素解析 (MeCab) を行った。そしてレビュー文を品詞ごとに分類した (表 5)。

表5 コスメレビューの品詞分類

商品ID	レビューID	文番号	語句番号	語句	品詞A	品詞B	品詞C
A-3	505151609	1	1	カバー	名詞	サ変接続	*
A-3	505151609	1	2	力	名詞	接尾	一般
A-3	505151609	1	3	とても	副詞	助詞類接続	*
A-3	505151609	1	4	高い	形容詞	自立	*
A-3	505151609	1	5	!	記号	一般	*
A-3	505151609	2	1	綺麗	名詞	形容動詞	*
A-3	505151609	2	2	に	助詞	副詞化	*
A-3	505151609	2	3	つき	動詞	自立	*
A-3	505151609	2	4	ます	助動詞	*	*
A-3	505151609	3	1	伸び	名詞	一般	*
A-3	505151609	3	2	も	助詞	係助詞	*
A-3	505151609	3	3	よい	形容詞	非自立	*
A-3	505151609	3	4	の	名詞	非自立	一般
A-3	505151609	3	5	で	助動詞	*	*
A-3	505151609	3	6	良かった	形容詞	非自立	*
A-3	505151609	3	7	た	助動詞	*	*
A-3	505151609	3	8	です	助動詞	*	*
A-3	505151609	4	1	時間	名詞	副詞可能	*

表5 左から、{商品 ID, レビューID, 文番号, 語句番号, 語句, 品詞 A, 品詞 B...}というように保存。

品詞ごとに分類することで、分析する際に品詞ごとの特徴を抽出することを考慮に入れていた。しかし、今回の研究では品詞による特徴を捉えることは困難であった。

そのため次に、レビュー文から時間的要素を含むコスメ評価記述を抽出した(表 6)。

表6 時間的要素と関連するコスメ評価表現

レビューID	文番号	語句番号	コスメ評価	時間的要素
505167949	2	1		付け
505167949	2	2		たて
505167949	2	4	サラサラ	
505167949	2	10		馴染ん
505167949	2	12	ツヤ	
505167949	3	1		夕方
505167949	3	3	化粧	
505167949	3	4	崩れ	
505167949	3	6	汚く	
505167949	3	7	なら	
505167949	3	8	ない	
505167671	5	9		叩い
505167671	5	10		た
505167671	5	11		後

表 6 左から{レビューID, 文番号, 語句番号, コスメ評価, 時間的要素}というように保存。

そしてここで時間的要素を、特に“使用直後”と“時間経過後”の2つに軸を絞った。

最後に、図のデータと 3.1 で作成したパターン表 (図) を照らし合わせ、パターン分類を行った。

5. データ分析

本章では、実際のデータを数値化し、それぞれパターンごとの傾向について述べていく。

5.1 抽出データ

本節では、今回の研究で実際に利用したデータについて述べる。

今回の研究では、@cosme からデータを手動で抽出した。ユーザがコスメ商品を使用してからの時間経過を表すレビューが多く含まれているカテゴリは“ベースメイク”と“スキンケア”であったため、この二つに着目した。

ベースメイク・スキンケアカテゴリからそれぞれ 5 商品選び、最新 30 件のレビューデータを使用した (表 7)。

表7 @cosme から利用したデータ数

	調べたデータ	合計
ベースメイク	5商品 × 30件	150件
スキンケア	5商品 × 30件	150件
合計	10商品 × 30件	300件

次に、レビューの中から時間的要素を含むレビューのみを抽出した。その結果、ベースメイクアイテムのレビューからは 76 件、スキンケアアイテムのレビューから 87 件抽出することができた(表 8)。

表 8 時間的要素を含むレビュー抽出データ

ベースメイク	時間的要素含むレビュー	スキンケア	時間的要素含むレビュー
商品A	14件	商品F	24件
商品B	16件	商品G	17件
商品C	12件	商品H	18件
商品D	22件	商品I	10件
商品E	12件	商品J	18件
合計	76件	合計	87件

それぞれ商品によってバラつきはあるが、ベースメイク・スキンケア共に時間的要素を含むレビューを半分以上抽出することができた。

次に、抽出したデータを利用し、時間的要素を含むコスメ評価記述を抽出した(4章 表 6)。

そしてこのデータを利用し、レビューID ごとではなく、文章ごと、または時間の流れにおける文章のつながりごとに新たに系列番号を付けた。

例えば、「ランピッド(商品名)は真っ白なパウダーなので、叩いた後は顔面も真っ白な感じになりますが、時間が経って馴染んでくると明るい透明感だけが残るようなイメージです。」という文章がある。この文章は1文で一連の時間経過を表しているため、一つの系列として保存しておく。

また、「初日はすごくいいかも!と思いましたが、その翌日からはニキビが出来始めます。首は、何故か乾燥してカサカサになってしまいます。」この文章では、2文に渡って、時間の流れとともに変化したコスメ評価が表記されているため、2文で一つの系列として保存しておく。

このような手順で系列番号を振り分けるところ、全部で213個の系列に分けることができた。

これら213個に分けた系列を3章で分類したパターンに沿って分けた(3章 表 1)。

今回、10パターンに分類したところ、213個中164個をこのパターン表で分けることができた。パターン表に当てはまらなかった49個中21個は、3.2節で述べたさらなる時間経過後のパターンに当てはめることができる。また残りの28個の系列の中には、時間的要素を含むコスメ評価が“使用直後”や“時間経過後”ではなく、“使用する前”のユーザ自身の肌状態などを記述しているものが見られた。

5.2 表現の傾向

今回、抽出したデータから表現例が多かったパターン4つ(TT-tt, FT-tt, FF-tt, FF-ft)から傾向を分析した。

まず初めに、TT-tt, FT-tt, FF-ttのこの三つのパターンは、使用直後・時間経過後のいずれかに時間表現が含まれるものである。この三つに関しては、時間表現が含まれているか含まれていないかの違いはあるものの、それ以外での違いはあまり見られなかった。

しかし、時間表現が含まれているという点で、ベースメイク商品の口コミ、スキンケア商品の口コミでの大きな違いが見られた。

ベースメイク商品の口コミは、ほとんどが使用してから1日以内の使用感について述べられていた。例えば、「仕上がりに～お昼頃」「朝～夕方」「一日中」などである。

それに対してスキンケア商品の口コミは、使用してから1週間後や一か月後、また曖昧な表現では「使い続けていくうちに」などの表現が多く見られた。

また、ベースメイク商品・スキンケア商品共に、時間表現に着目することで、「化粧の持ち」や「肌への効果」をより明確にすることができる。

例えば、ベースメイク商品において、コスメ評価で多く記述されていたものは、「崩れるか、使用してからの変化(顔色・テカリ具合など)、肌への負担」についてだった。このようなコスメ評価と時間表現を組み合わせることで、その商品はいつ崩れる場合が多いのか、どのくらい持つのかを特定可能となる。

スキンケア商品においては、「肌への効果(どのように変化したか)」について多く記述があった。このようなコスメ評価と時間表現を組み合わせることで、その商品が肌に即効性があるかなどの情報を得ることが出来る。

また、4つ目のFF-ftパターンは時間表現を含まず、商品を使用した結果をコスメ評価として表記されていた。このパターンからは、「いつ効果が表れたか?」などの情報を得ることは難しい。しかし、商品を使用した結果を表記しているため、このパターンにおけるコスメ表記は商品のテクスチャや使用直後の表現ではないことが分かる。コスメ商品において、「商品を使用した結果どうなったか?」という情報はユーザにとって重要な部分である。そのため、時間経過後のコスメ評価の記述は必要不可欠な情報となる。

6. まとめと今後の課題

今回の研究では、時間的要素を含むコスメ評価について分析を行った。分析する際、3章で述べたパターンに従ってレビュー文を分類したところ、今回利用したデータの多くを10パターンに分けることができた。ベースメイク商品とスキンケア商品において、多くのレビュー文は今回作成したパターンで分類可能であることが分かったため、このパターンを利用することは有用性があると言える。

しかし、今回はデータを手動で@cosmeからデータを取ってきたため、分析できたデータ量が少なかった。そのため、パターンによってはデータ数が少ないものがあり、データ数が少ないものは、パターンごとの傾向を分析することが困難であった。

今後は、今回手動で行った方法をデータベースやプログラミング等を利用して自動で行う必要がある。そして、それぞれのパターンのアプローチ方法を考え、自動化のためのシステ

ムを構築する必要がある。

また、コスメレビュー文において、時間的要素は”使用直後”と”時間経過後”のみならず、“使用前”や“さらなる時間経過後”など、様々な要素を含むものが存在することが見受けられた。そのため、今後はさらに分析するデータ量を増やし、どのような時間的要素を含むコスメレビュー文が存在するかを抽出する必要がある。また、抽出した時間的要素を分類するためのパターンを新たに作成しなければならない。

さらに、今回得られた知見を元に、時間的要素を含むコスメ評価とユーザ属性を組み合わせることによって、「こうゆう肌質の人にはこのようなコスメ商品を推薦することができそうだ」というようなコスメ商品推薦システムを構築することを目標とし、今後も分析を行っていく。

参考文献

- [1]@cosme <http://www.cosme.net/>
- [2]松波 友稀, 上田 真由美, 中島 伸介. コスメアイテム使用感および嗜好度判定を目的としたレビュー分析手法の提案, 情報処理学会 2015
- [3] 安部小百合, 小林一郎. 化粧品レビューデータを用いたレビュー推薦システムの開発への取り組み, 情報処理学会 2015
- [4] 中野裕介, 湯本高行, 新居学, 上浦尚武. 機械学習による商品レビューの属性一意見ペアの抽出, 情報処理学会 2015
- [5] 松尾哉太, 新妻弘崇, 太田学. レビュー解析に基づくユーザ評価の根拠提示の一手法, 情報処理学会 2014
- [6] 松尾哉太, 新妻弘崇, 太田学. レビュー文書における重要文選択の一手法. 情報処理学会 2015