

# 共起ネットワーク分析を用いた訳あり市場の考察

## －「カニ」と「ミカン」のユーザーレビューを題材として－

### Consideration of WAKEARI Market by Co-occurrence Network Analysis

#### － In Cases of User Reviews about “Crab” and “Mandarin Orange” －

吉見 憲二<sup>1</sup>, 樋口 清秀<sup>2</sup>  
Kenji YOSHIMI, Kiyohide HIGUCHI

## 1 はじめに

近年、「訳あり商品<sup>1</sup>」と呼ばれる規格外商品を扱う市場が拡大し、一種のブームを形成している。Googleで「訳あり商品」という単語を検索すると約15,700,000件ものページがヒットするだけでなく、上位には多くの専用サイトが表示される<sup>2</sup>。このように「訳あり商品」は現在では私たちの生活に深く関与している。しかし、「訳あり商品」が一般のユーザーにとって身近になったのはごく最近のことである。「Google Insights for Search<sup>3</sup>」を用いて「訳あり、わけあり、ワケあり」といった単語の検索数の変遷を調べてみると、2005年の中頃から一定数は存在したものの、2008年の後半になって急激に伸びていることが確認できる（図1）。

「訳あり商品」の流通の特徴として、当初はインターネット通販の形態を取り、規格外であることを明示して大々的に販売していたことが挙げられる。「訳あり商品」はその特性から市場規模を特定することは困難であるが、株式会社富士経済が発表した「食品ダイレクトセールス市場トレンドデータ2011」では、食品のみを対象としたインターネット通販の2010年の市場

規模を2,645億円と推計し、前年比20.2%増となった要因の1つとして「訳あり商品」の存在に言及している。現在では扱う商品は食品だけでなく、衣料品や家電製品、旅行プラン<sup>4</sup>など多岐にわたっている。加えて、インターネット通販に留まらず、大手スーパーや百貨店等による「訳あり商品」の直売、飲食店における「訳あり商品」の提供など販売網についても拡大している。楽天が運営する仮想商店街「楽天市場」では、2008年7月に約6万点だった「訳あり商品」が2010年1月には約48万点にのぼっている<sup>5</sup>。

このように「訳あり商品」の市場規模は順調に拡大しているが、こうした「訳あり商品」が流通網に乗ることは、生産者側にとっては「廃棄ロスの削減」という点で、流通側にとっては「新たな販売機会の獲得」という点で、消費者側にとっては「安価な質の高い商品の入手」という点でメリットが存在し、win-win-winの関係性が築かれていると言える。ただし、その前提として「訳あり商品」がその定義の通り、「見た目は悪いが質は通常の商品と変わらない」ものであることが必須となる。当然のことであるが、単純に質の悪い商品を販売することは消費者にとってはデメリットにしかならず、win-win-winの関係性は崩れてしま

ウェブ検索の人気度: 訳あり, わけあり, ワケあり

日本, 2004年 - 現在

カテゴリ: ニュースとメディア, 食品と飲料, エンターテインメント, ショッピング, 地域, 芸術, 人文科学

平均値 ⑦  
訳あり 16  
わけあり 22  
ワケあり 3

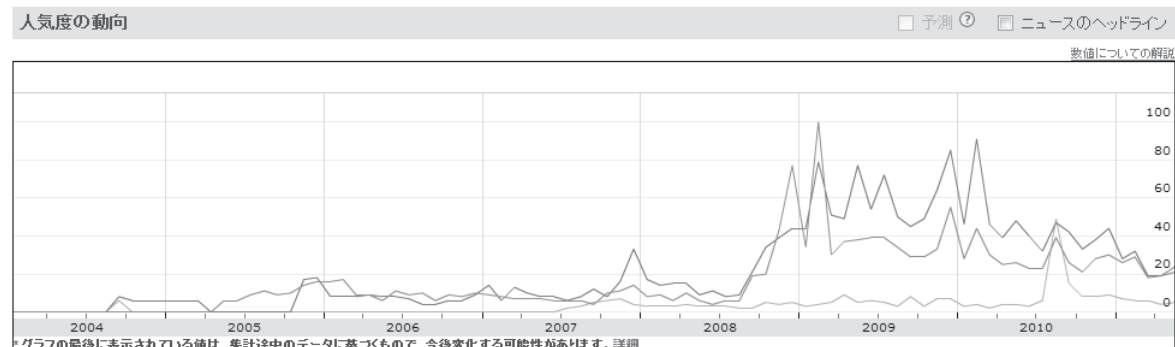


図1 ウェブ検索の人気度

2011年5月13日受付, 2011年11月21日再受付, 2011年11月24日採録

<sup>1</sup> 早稲田大学大学院国際情報通信研究科博士課程修了, 早稲田大学本庄高等学院非常勤講師

<sup>2</sup> 早稲田大学大学院国際情報通信研究科兼任教授

う。「訳あり商品」のブームの背景には、各主体における信頼関係が存在していることが示唆される。

一方で、近年では、不正会計操作や耐震偽装の問題と並んで食品偽装の問題が大きく報じられている。こうした社会問題に共通するのは、ネガティブな情報を隠すことによって消費者の利益を損なっているという点である。そのため、多発する企業不祥事に対しては、規制の強化により情報開示を義務付けるケースが増えている。例えば、不正会計操作の問題ではいわゆる日本版SOX法が制定され、耐震偽装の問題では建築基準法が改正されている。しかし、このような規制強化はコンプライアンス不況と呼ばれる状況を招いているとの批判もある。このように、規制の強化は企業におけるコストの増加要因となるだけでなく、消費者にとっても利便性の低下や価格への転嫁などマイナスの影響が少なくない。これに対して「訳あり市場」の事例では、規格外であることを明示することが人気につながっており、消費者の利益にも適合しているという点で対照的な結果を導いている。

両者の違いは、企業と消費者間における信頼が形成されているか否かであると考えられる。本研究では、こうした企業と消費者間における信頼関係に着目し、「訳あり市場」の拡大において消費者の声（ユーザーレビュー）が果たした役割について取り上げる。

## 2 先行研究の考察

先行研究として、情報開示政策と情報費用の問題を取り上げ、その解決策としてWebマイニングやテキストマイニングといったデータマイニング手法が活用されていることについて言及する。

### 2.1. 情報開示政策と情報費用についての先行研究

Akerlof (1970) は米国のレモン市場を例に、買い手と売り手の間の情報の非対称性の問題について取り上げている。レモンとは質の悪い中古車のことを指し、中古車市場では売り手は自分の車の情報をもっているが、買い手は当該の車の情報を十分にもっていないため、売り手が質の悪い中古車（レモン）を質の良い中古車として騙す危険性がある。このような行為が横行すると、結果として質の悪い中古車だけが市場に流通するようになってしまう。情報の非対称性への対策としては、シグナリング<sup>6</sup>やスクリーニング<sup>7</sup>などが知られており、企業への情報開示強化はスクリーニングの役割を果たしていると考えられる。

Davis and Botkin (1994) は機械工の話为例に、情報費用の増大について取り上げている。1965年の機械工は500ページの修理マニュアルを覚えればほとんどすべての自動車を修理することができたが、1994年時点ではその修理マニュアルの分量は50万ページにも達するという。このような事例から、現代では扱う情報が飛躍的に増えていることについて触れており、「情

報を知識に変える方法を知っている企業こそが最も成功する」としている。

一方で、WINSTON (2008) は米国の情報開示政策の効果を疑問視し、理論と実証研究が一致していないと主張している。例えば、本来情報を開示することによって企業価値を守るはずのSOX法は、中小企業の会計報告と監査報告書のコストを上げて、企業価値に否定的な影響を与えているとしている。このような理論と実証研究との齟齬については、Davis and Botkin (1994) が触れているような情報費用の問題があると考えられる。

### 2.2. データマイニングに関する先行研究

こうした日々増大する情報の海から意味のある情報を抽出する手法がデータマイニングである。データマイニングには、Webから役立つ情報を抽出するWebマイニングや、テキストデータから必要な情報を抽出するテキストマイニング等がある。

松尾、石塚 (2002) では、語の共起の統計情報を用いて、単一の文書からキーワードを抽出する方法を提案している。また、共起関係をもつ一般的なデータから特徴的なアイテムが発見できることを示している。松尾ほか (2005) では、Webマイニングの手法を用いて、Web上の情報のみから学会というコミュニティにおける人間関係を再現している。高橋ほか (2010) では、Webと修飾表現の適合度を判定する手法を提案し、投稿型レシピサイト「クックパッド」と旅行ツアーサイト「Yahoo!トラベル」に掲載されている情報を用いて実験を行っている。新谷ほか (2011) では、マイクロブログから特定の話題について抽出する手法を提案し、共起語と投稿間隔を併用した抽出手法が最も精度が高かったと結論づけている。

### 2.3. 先行研究における課題

先行研究では、極端な情報開示が情報費用の増大を招いていること、情報費用を減少させるためにWebマイニングやテキストマイニングといったデータマイニング手法が活用されていることが示された。しかし、「訳あり市場」にフォーカスした研究は著者の知る限りまだ存在しない。

本研究では、「訳あり市場」において「消費者からの評価」を示すユーザーレビューに着目し、ユーザーレビューが情報費用を減少させるはたらきを示しているかについて検討する。

### 2.4. 仮説

本研究における仮説は以下の通りである。

#### 【仮説1】

「訳あり商品」と「通常商品<sup>8</sup>」においては、使用単語に違いある。

## 【仮説2】

「訳あり商品」と「通常商品」においては、ユーザーレビューの構造が異なる。

## 【仮説3】

「訳あり商品」に関するユーザーレビューでは、消費者の不安を和らげるような特性を有している。

これらの仮説を検証する手段として、テキストマイニングによる分析を行った。なお、テキストマイニングには、樋口耕一氏によって提供されているフリーのテキストマイニングソフトであるKH Coder<sup>9</sup>を使用した。

### 3 ユーザーレビュー分析の手順

本研究では、分析の対象であるユーザーレビューについて、楽天市場の「カニ」ジャンルと「ミカン」ジャンルを採用した。以下では、その理由と分析までの手順について説明する。

#### 3.1. 楽天市場を選んだ理由

図2は、「Google Insights for Search」を用いて代表的な「訳あり商品」の取扱いサイトである「楽天」と「ぐるなび」の検索結果を比較したものである。

上記の2つのサイトにおいては、ほとんどの時期において楽天が検索数において優位に立っており、より使用されているプラットフォームであると考えられる<sup>10</sup>。そのため、ユーザーレビューの分析の対象として妥当であると判断した。

#### 3.2. 対象商品を選んだ理由

対象となる商品として、楽天市場の海産物市場における「カニ」ジャンルとフルーツ市場における「ミカン」ジャンルを選択した。この2つの商品は、海産物市場とフルーツ市場のそれぞれにおいて商品数が最多となっており、ユーザーレビュー<sup>11</sup>が商品決定に与える影響が大きいと考えられる。加えて、「訳あり商品」と「通常商品」のそれぞれがユーザーレビュー数の上

位に十分存在していることから、分析に適していると判断した。

#### 3.3. ユーザーレビュー選定の手順

分析にあたっては、「訳あり商品」と「通常商品」のそれぞれにおいて、ユーザーレビュー数の上位10件の商品から、2011年4月末時点で新着のユーザーレビュー100件ずつを採用した。これにより、「カニ通常」「カニ訳あり」「ミカン通常」「ミカン訳あり」の4つの類型において1,000件ずつ、合計4,000件のユーザーレビューが分析対象となる。

なお、今回採用した商品の感想（ユーザーレビュー）の数、価格、評価（5点満点）及び各類型の平均値をまとめたものが表1から表4である。

表1から表4を比べてみると、「カニ通常」において商品1の感想（ユーザーレビュー）の数が著しく多いことから平均値が高くなっているが、他に際立った違いは見られない。また、「ミカン」ジャンルにおいては、感想（ユーザーレビュー）の数、価格において「訳あり商品」の方が「通常商品」よりも多くなっている。

表1 「カニ通常」の対象商品

カニ通常			
名称	感想	価格	評価
商品1	13821	5,250	4.47
商品2	4204	3,980	4.21
商品3	3223	13,000	4.59
商品4	3203	8,980	4.65
商品5	2502	3,333	4.35
商品6	2095	3,780	4.47
商品7	1665	5,980	4.35
商品8	1304	8,980	4.21
商品9	1172	4,980	4.61
商品10	1001	3,980	4.57
平均	3419	6,224	4.45

ウェブ検索の人気度: 楽天訳あり, 楽天わけあり, ぐるなび訳あり, ぐるなびわけあり

日本, 2004年 - 現在

カテゴリ: ニュースとメディア, ショッピング, 食品と飲料, 地域

平均値

楽天訳あり 7  
 楽天わけあり 11  
 ぐるなび訳あり 10  
 ぐるなびわけあり 2

人気度の動向

☐ 予測 ☐ ニュースのヘッドライン



図2 「訳あり商品」取扱いサイトの比較

表2 「カニ訳あり」の対象商品

カニ訳あり			
名称	感想	価格	評価
商品 1	4649	5,250	4.25
商品 2	2577	4,179	3.94
商品 3	2359	5,250	4.44
商品 4	2211	10,500	4.58
商品 5	1781	3,129	4.16
商品 6	1462	3,980	4.13
商品 7	1343	6,780	4.17
商品 8	1106	4,980	4.09
商品 9	1046	5,980	4.31
商品10	1020	6,980	4.7
平均	1955	5,701	4.28

表3 「ミカン通常」の対象商品

カニ訳あり			
名称	感想	価格	評価
商品 1	5589	1,580	3.64
商品 2	2376	500	4.55
商品 3	2152	1,680	4.42
商品 4	1224	1,000	4.41
商品 5	1162	2,980	4.25
商品 6	689	999	4.39
商品 7	555	3,150	4.27
商品 8	544	2,980	4.37
商品 9	530	1,470	4.15
商品10	472	2,580	4.53
平均	1529	1,892	4.30

表4 「ミカン訳あり」の対象商品

カニ訳あり			
名称	感想	価格	評価
商品 1	4042	2,480	4.25
商品 2	3905	3,280	4.59
商品 3	2178	1,500	4.4
商品 4	1900	1,980	4.53
商品 5	1494	2,480	4.25
商品 6	1338	1,500	4.41
商品 7	1233	2,380	4.42
商品 8	1183	2,680	4.27
商品 9	985	1,333	3.89
商品10	733	1,480	3.49
平均	1899	2,109	4.25

## 4 分析結果及び仮説の検証

本研究では、通常商品のユーザーレビューと訳あり商品のユーザーレビューとの間に何らかの差異が生じていることを仮定している。ここでは、カイ二乗検定と共起ネットワーク分析の結果から、仮説の検証を行う。

### 4.1. 抽出単語数の比較

それぞれのジャンルにおいて、出現数の上位20件を抽出したものが表5と表6である<sup>12</sup>。背景色の付いている単語に関しては、通常商品と訳あり商品のそれぞれに登場していることを意味している。

まず、「カニ」のジャンルについては、上位20件のうち17件の単語が重なる結果となった。全体のユーザーレビュー数に占める比率も7%強から40%前後と近似している。次に、「ミカン」のジャンルでは、上位20件のうち15件の単語が重なった。比率は7%前後から高いもので60%近くにまで達したが、特に上位の語句では通常商品と訳あり商品の間で大きな違いは見られなかった。

表5 「カニ」ジャンルにおける抽出単語数の比較

カニ通常			カニ訳あり		
単語	出現数	比率	単語	出現数	比率
美味しい	425	42.5%	美味しい	379	37.9%
食べる	357	35.7%	食べる	359	35.9%
カニ	289	28.9%	カニ	279	27.9%
購入	244	24.4%	購入	235	23.5%
思う	243	24.3%	思う	227	22.7%
身	193	19.3%	満足	211	21.1%
味	185	18.5%	家族	196	19.6%
満足	170	17.0%	味	183	18.3%
大きい	145	14.5%	身	166	16.6%
価格	119	11.9%	量	161	16.1%
家族	119	11.9%	価格	146	14.6%
注文	107	10.7%	良い	131	13.1%
喜ぶ	106	10.6%	入る	107	10.7%
良い	104	10.4%	注文	105	10.5%
量	85	8.5%	買う	85	8.5%
届く	84	8.4%	喜ぶ	84	8.4%
買う	84	8.4%	正月	80	8.0%
今回	82	8.2%	大変	74	7.4%
商品	76	7.6%	大きい	72	7.2%
送る	74	7.4%	商品	71	7.1%



表6 「ミカン」ジャンルにおける抽出単語数の比較

ミカン通常			ミカン訳あり		
単語	出現数	比率	単語	出現数	比率
美味しい	578	57.8%	美味しい	574	57.4%
ミカン	436	43.6%	ミカン	351	35.1%
甘い	387	38.7%	甘い	344	34.4%
食べる	267	26.7%	購入	264	26.4%
味	239	23.9%	食べる	263	26.3%
購入	217	21.7%	味	251	25.1%
思う	194	19.4%	思う	190	19.0%
買う	118	11.8%	買う	129	12.9%
リピート	117	11.7%	大きい	122	12.2%
皮	97	9.7%	リピート	112	11.2%
大きい	95	9.5%	注文	103	10.3%
注文	94	9.4%	入る	96	9.6%
小さい	93	9.3%	届く	94	9.4%
届く	89	8.9%	良い	85	8.5%
価格	88	8.8%	家族	84	8.4%
小粒	86	8.6%	小さい	83	8.3%
良い	84	8.4%	訳あり	81	8.1%
甘み	79	7.9%	価格	80	8.0%
満足	75	7.5%	見た目	78	7.8%
今年	69	6.9%	安い	74	7.4%

しかし、単純な出現数だけでは両方の商品において差異がないとは言いきれない。そこで、表5と表6に登場した単語についてカイ二乗検定を行った。それぞれの結果については、表7と表8の通りである（太字で示しているものは有意に登場数が多かったジャンル）。結果より、「カニ」ジャンルでは6単語が該当し、出現数上位20件で重複した単語において出現の傾向が異なることが読み取れた。一方で、「ミカン」ジャンルでは8単語が該当し、出現数上位20件において重複していない単語が目立った。このように、「訳あり商品」と「通常商品」においては、一部の使用単語に違いあることが確認できたが、その傾向は異なるものであった。

ただ、この結果からだけでは、ユーザーレビューの構造までもが異なっているとは断言することはできない。そこで、それぞれのユーザーレビューに登場する単語から共起ネットワーク分析を行った。

表7 カイ二乗検定（カニ）

単語	カニ普通	カニ訳あり	合計	カイ2乗値
美味しい	425 (42.50%)	379 (37.90%)	804 (40.20%)	4.212 *
満足	170 (17.00%)	211 (21.10%)	381 (19.05%)	5.188 *
大きい	145 (14.50%)	72 (7.20%)	217 (10.85%)	26.797 **
家族	119 (11.90%)	196 (19.60%)	315 (15.75%)	21.764 **
量	85 (8.50%)	161 (16.10%)	246 (12.30%)	26.073 **
入る	69 (6.90%)	107 (10.70%)	176 (8.80%)	8.529 **
ケース数	1000	1000	2000	

\*\* 1% 有意, \* 5% 有意

表8 カイ二乗検定（ミカン）

単語	ミカン普通	ミカン訳あり	合計	カイ2乗値
ミカン	436 (43.60%)	351 (35.10%)	787 (39.35%)	14.783 **
購入	217 (21.70%)	264 (26.40%)	481 (24.05%)	5.792 *
皮	97 (9.70%)	65 (6.50%)	162 (8.10%)	6.455 *
小粒	86 (8.60%)	62 (6.20%)	148 (7.40%)	3.860 *
甘み	79 (7.90%)	55 (5.50%)	134 (6.70%)	4.231 *
入る	60 (6.00%)	96 (9.60%)	156 (7.80%)	8.517 **
訳あり	8 (0.80%)	81 (8.10%)	89 (4.45%)	60.960 **
見た目	35 (3.50%)	78 (7.80%)	113 (5.65%)	16.545 **
ケース数	1000	1000	2000	

\*\* 1% 有意, \* 5% 有意

## 4.2. 共起ネットワーク分析

共起ネットワークとは、テキストの中で用いられた語と語の関係性を示したネットワークであり、ユーザーレビュー1件の中で用いられている単語間の関係を意味している。本研究では、ノードを登場数50以上の単語（全体の5%以上のレビューに登場）、リンクをJaccard係数<sup>13</sup>0.1以上<sup>14</sup>の共起関係とし、分析を行った。Jaccard係数は類似性の指標であり、単語間の共起関係を表すものとして広く使用されている。

Jaccard係数のデメリットとして、単独での出現数が多い単語ほど他の単語との関係が薄くなる傾向があることが指摘されているが、本研究における出現数が多い単語は両方のレビューに登場し、一般的に用いられるものであると考えられるため、問題は少ないと判断した。

以下では、各ジャンルの結果について、図を示しながら説明する。なお、図における円の大きさは単語の出現数を、リンクの太さは共起関係の強さを、円の色の濃さはネットワークにおける中心であることをそれぞれ意味している。ただし、円同士の距離は意味をもたないことも付言しておく。

### 4.2.1. 「カニ」ジャンルの通常商品

「カニ」ジャンルの通常商品では、図3のような結果となった。共起ネットワークの中心は、「身」とい

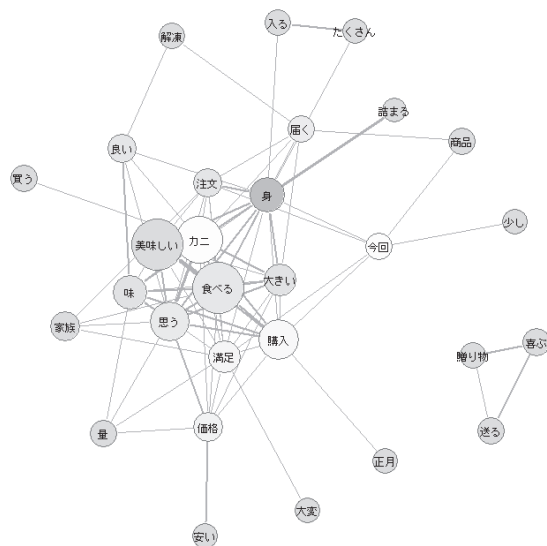


図3 「カニ」ジャンル通常商品のネットワーク

う単語であり、「大きい」「詰まる」「入る」「食べる」など多くの単語とのネットワークを有している。また、「贈り物」「喜ぶ」「送る」という一連の単語がつながっており、贈答用に多く利用されていることも窺える。

#### 4.2.2. 「カニ」ジャンルの訳あり商品

これに対して、訳あり商品では図4のような結果になった。こちらは「食べる」という単語がネットワークの中心になっていて、「身」「美味しい」「家族」など多くの単語とつながっていた。また、「購入」という単語もネットワークの中心に位置しており、実際に「食べた」「購入した」ということが意識して記載されていると言える。

特筆すべきは、「価格」という単語が「安い」「量」という単語だけでなく「味」という単語にもつながっていることで、通常商品には見られない点であった。「期待」と「残念」という単語間に共起関係が見られたのも訳あり商品に特有の現象であった。

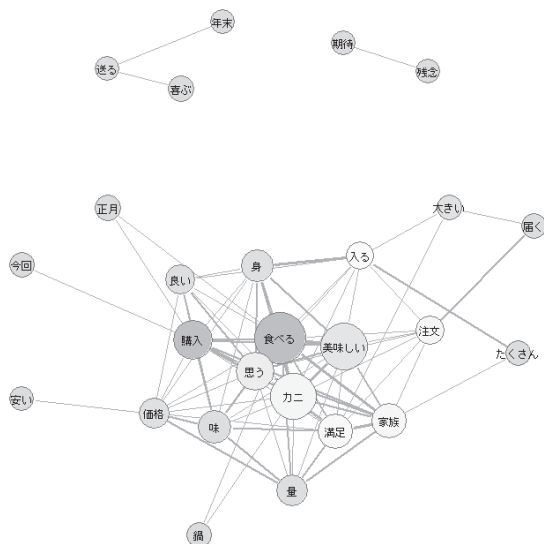


図4 「カニ」ジャンル訳あり商品のネットワーク

#### 4.2.3. 「ミカン」ジャンルの通常商品

「ミカン」ジャンルの通常商品の結果は、図5のようになった。こちらは「味」という単語がネットワークの中心で、「濃い」「美味しい」といった単語と強いつながりを有していた。また、「満足」という単語が「価格」に対して強いつながりを有している。「カニ」ジャンルでは訳あり商品だけに現れた「期待」と「残念」という単語間の共起関係も見られた。

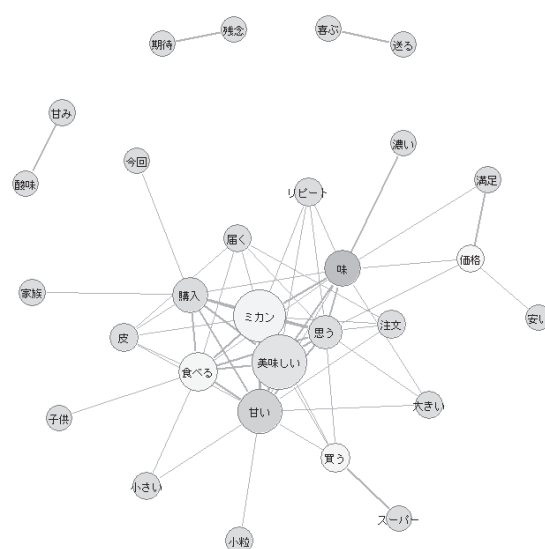


図5 「ミカン」ジャンル通常商品のネットワーク

#### 4.2.4. 「ミカン」ジャンルの訳あり商品

「ミカン」ジャンルの訳あり商品では、図6のような結果になった。訳あり商品でも「味」という単語がネットワークの中心にあるが、「訳あり」「見た目」といった訳あり商品に特有の単語との共起関係が見られる。また、「見た目」と「悪い」という単語間には強い共起関係が見られるが、「味」については「良い」という単語としか共起関係を有していないことが特徴的である。なお、「カニ」ジャンルと同様に「食べる」という単語もネットワークの中心に位置している。

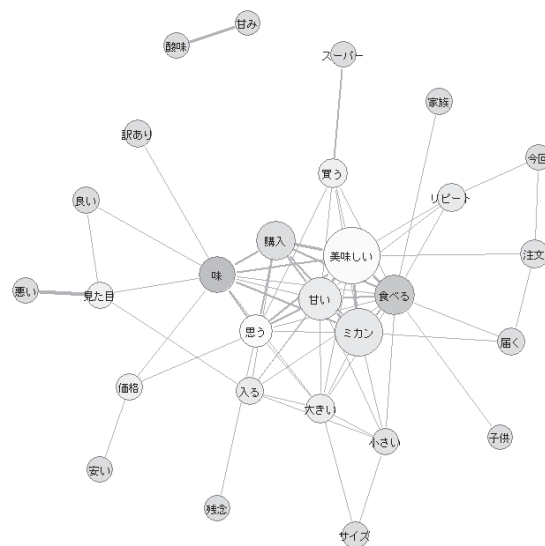


図6 「ミカン」ジャンル訳あり商品のネットワーク

#### 4.3. 共起ネットワークの検討

各類型における共起ネットワーク分析の結果から、ユーザーレビューの構造の差異が示唆された。特に、「カニ」ジャンルの訳あり商品においては「価格-味-量」という単語間のクラスタが、「ミカン」ジャンルの訳あり商品においては「悪い-見た目-味-良い」という一連の単語のつながりが導かれたことは興味深い結果であった。

しかし、共起ネットワーク図だけでは統計的に有意な差が生じているかは判別できない。そこで、共起関係の強い単語についてそれぞれ抽出し、カイ二乗検定を行った。共起関係の強い単語の上位20件の組み合わせは表9と表10に、その中でも有意な差が見られたものについては表11と表12に示している。

「カニ」ジャンルの訳あり商品においては、「家族」と「満足」という2つの語に対する共起関係において有意差が多く見られた。これは表7の結果とも整合的である。「ミカン」ジャンルの訳あり商品においては、「見た目-悪い」と「購入-味」という組み合わせにおいて有意差が見られたが、これも表8の結果と整合的である。

表9 共起関係の強い単語の抽出（カニ）<sup>15</sup>

カニ普通			カニ訳あり		
単語1	単語2	weight	単語1	単語2	weight
食べる	美味しい	0.290429	食べる	美味しい	0.272414
カニ	食べる	0.266667	実家	送る	0.267442
カニ	美味しい	0.261484	カニ	食べる	0.23166
身	詰まる	0.21393	家族	食べる	0.230599
カニ	思う	0.211845	カニ	美味しい	0.211786
美味しい	身	0.207031	入る	身	0.207965
贈り物	喜ぶ	0.201342	家族	満足	0.204748
食べる	思う	0.2	食べる	思う	0.203285
思う	美味しい	0.197133	カニ	身	0.196237
購入	美味しい	0.19678	購入	思う	0.193798
喜ぶ	送る	0.192053	満足	量	0.185304
たくさん	入る	0.185185	思う	美味しい	0.181287
思う	身	0.184783	カニ	購入	0.178899
カニ	購入	0.184444	家族	美味しい	0.178279
カニ	味	0.179104	家族	量	0.178218
購入	思う	0.176329	満足	食べる	0.173196
食べる	大きい	0.175644	満足	味	0.169643
購入	味	0.168937	美味しい	身	0.169528
思う	大きい	0.168675	カニ	満足	0.167064
カニ	身	0.16707	カニ	家族	0.161369

表10 共起関係の強い単語の抽出（ミカン）

ミカン普通			ミカン訳あり		
単語1	単語2	weight	単語1	単語2	weight
ミカン	美味しい	0.352	見た目	悪い	0.343434
美味しい	甘い	0.305819	ミカン	美味しい	0.31392
ミカン	甘い	0.281931	美味しい	甘い	0.294781
食べる	美味しい	0.246313	ミカン	甘い	0.261343
食べる	甘い	0.238636	購入	美味しい	0.247024
買う	スーパー	0.232394	酸味	甘み	0.244444
甘み	酸味	0.230088	食べる	美味しい	0.243685
ミカン	味	0.222826	買う	スーパー	0.21875
ミカン	食べる	0.222609	食べる	甘い	0.216433
ミカン	購入	0.209259	ミカン	食べる	0.211045

購入	美味しい	0.206373	ミカン	購入	0.196498
喜ぶ	送る	0.2	購入	味	0.186636
美味しい	味	0.190962	美味しい	味	0.185345
味	濃い	0.1893	思う	味	0.182306
ミカン	思う	0.186441	購入	甘い	0.180583
価格	満足	0.181159	思う	美味しい	0.169985
購入	思う	0.181034	ミカン	味	0.16441
思う	甘い	0.180894	購入	思う	0.164103
思う	味	0.17663	思う	甘い	0.155844
食べる	思う	0.17602	購入	食べる	0.155702

表11 共起関係の単語のカイ二乗検定（カニ）

単語1	単語2	カニ普通	カニ訳あり	合計	カイ2乗値
カニ	美味しい	148 (14.80%)	115 (11.50%)	263 (13.15%)	4.483 *
カニ	思う	93 (9.30%)	63 (6.30%)	156 (7.80%)	5.847 *
美味しい	身	106 (10.60%)	79 (7.90%)	185 (9.25%)	4.027 *
贈り物	喜ぶ	30 (3.00%)	9 (0.90%)	39 (1.95%)	10.460 **
食べる	大きい	75 (7.50%)	23 (2.30%)	98 (4.90%)	27.908 **
思う	大きい	56 (5.60%)	23 (2.30%)	79 (3.95%)	13.495 **
家族	食べる	50 (5.00%)	104 (10.40%)	154 (7.70%)	19.762 **
家族	満足	32 (3.20%)	70 (7.00%)	102 (5.10%)	14.143 **
満足	量	24 (2.40%)	59 (5.90%)	83 (4.15%)	14.531 **
家族	美味しい	47 (4.70%)	87 (8.70%)	134 (6.70%)	12.166 **
家族	量	17 (1.70%)	54 (5.40%)	71 (3.55%)	18.925 **
満足	味	38 (3.80%)	58 (5.80%)	96 (4.80%)	3.950 *
カニ	満足	45 (4.50%)	71 (7.10%)	116 (5.80%)	5.720 *
カニ	家族	37 (3.70%)	66 (6.60%)	103 (5.15%)	8.025 **
ケース		1000	1000	2000	

\*\* 1%有意, \* 5%有意

表12 共起関係の単語のカイ二乗検定（ミカン）

単語1	単語2	ミカン普通	ミカン訳あり	合計	カイ2乗値
ミカン	美味しい	264 (26.40%)	221 (22.10%)	485 (24.25%)	4.801 *
ミカン	甘い	181 (18.10%)	144 (14.40%)	325 (16.25%)	4.761 *
ミカン	味	123 (12.30%)	85 (8.50%)	208 (10.40%)	7.346 **
喜ぶ	送る	19 (1.90%)	5 (0.50%)	24 (1.20%)	7.127 **
ミカン	思う	99 (9.90%)	70 (7.00%)	169 (8.45%)	5.067 *
価格	満足	25 (2.50%)	12 (1.20%)	37 (1.85%)	3.965 *
見た目	悪い	7 (0.70%)	34 (3.40%)	41 (2.05%)	16.833 **
購入	味	52 (5.20%)	81 (8.10%)	133 (6.65%)	6.315 *
ケース		1000	1000	2000	

\*\* 1%有意, \* 5%有意

#### 4.4. 考察

上記の共起ネットワーク分析の結果からは、

- ・ 通常商品では、商品の内容（「カニ」ジャンルでは「身」について、「ミカン」ジャンルでは「味」について）が中心に取り上げられている。
  - ・ 訳あり商品では、消費者自身が不安に感じている点（「カニ」ジャンルでは「価格」に対しての「味」と「量」について、「ミカン」ジャンルでは「見た目（の悪さ）」に対しての「味」について）が通常商品との比較において詳細に触れられている
- という2点が確認できた。

カイ二乗検定の結果からは、それぞれの商品におい

て有意差が生じている語が、共起ネットワークでも有意差を生み出していることが確認できた。ただし、上位20件の組み合わせに限定したせいか、共起ネットワーク分析によって得られた示唆とは必ずしも一致しなかった。共起ネットワークにおいては、Jaccard係数0.1以上がリンクの基準となったため、カイ二乗検定の対象範囲を拡大することで新たな知見が得られるかもしれない。

#### 4.5. 仮説検定

本研究では、3つの仮説を提示した。仮説1については、一部の単語において出現傾向の有意な差が見られた。仮説2と仮説3については、共起ネットワーク分析から「訳あり商品」と「通常商品」のユーザーレビューの構造の違いが観察された。特に、「訳あり商品」のユーザーレビューにおいては、消費者が不安に感じるポイントがユーザーレビューに反映されていることが確認できた。

## 5 まとめ

本研究では、信頼の低下が過大な情報費用の負担を招いている状況を背景に、テキストマイニングの手法を用いて、「訳あり市場」におけるユーザーレビューが果たしている役割について分析した。結果からは、「訳あり商品」に対して通常の商品とは違った配慮がユーザーレビュー上でなされていることが示唆された。

ただし、本研究においては、留意すべき点が3点ある。

1点目は、テキストマイニングソフトに対する信頼性である。KH Coderは広く使用されているソフトウェアではあるが、より正確を期すのであれば単語の抽出を人力で再確認するなどの措置が必要であると考えられる。

2点目は、今回分析した事例が「カニ」と「ミカン」という2つの限定されたジャンルに留まっていることである。「訳あり市場」の特性として一般化するには事例が不足していることは言うまでもない。

3点目は、「消費者の評価」が決して万能ではないという点である。例えば、今回モデルとして利用した「楽天市場」では、過去にグルメ大賞となった業者が食品偽装で摘発される事件なども実際に起こっている<sup>16</sup>。しかし、ユーザーレビュー上では、偽装を指摘する声が早くから上がっていたとも言われており、ユーザーレビューのポジティブな効果を維持するために必要な要素についてはまだまだ検討の余地が残されている。

上に挙げた留意事項に加えて、今後の研究課題として、「楽天市場」以外のモデルとも比較していくことが必要である。特に、ユーザーレビューがないモデルと比較することにより、ユーザーレビューがどの程度

影響を与えているかということについては明確にしなければならない。

## 謝辞

本論文をまとめるにあたり、匿名の査読者の方には、非常に有益なコメントをいただきました。この場を借りて深く御礼申し上げます。

## 参考文献

- 1) Clifford Winston, "The Efficacy of Information Policy: A Review of Archon Fung, Mary Graham, and David Weil's Full Disclosure: The Perils and Promise of Transparency", *Journal of Economic Literature* 2008, pp.704-717.
- 2) George A. Akerlof, "The market for lemons: quality uncertainty and the market mechanism", *Quarterly Journal of Economics* 1970, pp.488-500.
- 3) Stan Davis and Jim Botkin, "The Coming of Knowledge-Based Business", Harvard Business Review 1994. (西山昭彦訳『知識の時代の企業戦略』ダイヤモンド社, 1995.)
- 4) 李 知修, 吉見憲二, 樋口清秀「ICTの活用による埋め込みネットワークの形成」情報通信学会第26回学会大会予稿, 2009.
- 5) 株式会社富士経済「食品ダイレクトセールズ市場トレンドデータ2011」, 2011.
- 6) 新谷歩生, 関 洋平, 佐藤哲司「投稿間隔に基づくマイクロブログからの話題チャンク抽出に関する一検討」DEIM Forum 2011, 2011.
- 7) 高橋良平, 小山 聡, 大島裕明, 田中克己「Webテキストと修飾表現との適合度判定手法」DEIM Forum 2010, 2010.
- 8) 松尾 豊, 石橋 満「語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム」人工知能学会誌 17 (3), 2002.
- 9) 松尾 豊, 友部博教, 橋田浩一, 中島秀之, 石橋 満「Web上の情報からの人間関係ネットワークの抽出」人工知能学会誌 20 (1), 2005.
- 10) 吉見憲二, 樋口清秀「「埋め込み」アプローチによる訳あり市場の考察」日本社会情報学会 (JSIS & JASI) 合同研究発表大会予稿, 2009.

## 注

- 1 傷がついたり見た目が悪いなどの理由で一般的な店舗への流通ができない商品のこと。「訳あり、わけあり、ワケあり」といった呼称があるが、本研究では、「訳あり」に統一して使用する。
- 2 [http://shop.gnavi.co.jp/Mallgn/special/wakeari\\_sale/](http://shop.gnavi.co.jp/Mallgn/special/wakeari_sale/) (わけありグルメ特集【ぐるなび食市場】)
- 3 <http://www.wakeichi.com/> (北海道わけあり市場) など。
- 4 <http://www.google.com/insights/search/>
- 5 <http://travel.gnavi.co.jp/special/wakeari/> (訳ありプラン特集【ぐるなびトラベル】) など。
- 6 「「訳あって安い」なぜ増える——「正直」な説明消費くすぐる (エコノ探偵団)」(日本経済新聞2010年1月24日朝刊)
- 7 情報優位者が情報劣位者に品質に関する情報を提供すること。
- 8 情報劣位者が情報優位者に品質に関する情報を開示すること。
- 9 本研究では、「訳あり商品」に対抗する概念として便宜的に「通常商品」という呼称を用いる。これは「訳あり商品」ではない商品の総称を意図している。
- 10 <http://khc.sourceforge.net/> (KH Coder Index Page)
- 11 なお、単純なGoogle検索の結果では、「ぐるなび」のサイトが最上位に位置している (2011年5月1日現在)。



- 11 楽天市場では、ユーザーレビューは「感想」という名称になっているが、本研究ではユーザーレビューという呼称を用いる。
- 12 なお、分析にあたって表記ゆれについては修正している。例えば、「値段」と「価格」はそれぞれ同じ概念を示しているため、「価格」という表記で統一している。
- 13 Jaccard 係数は単語 X と単語 Y において、 $|X \cap Y|/|X \cup Y|$  の式によって求められる。類似性の指標としては他に Simpson 係数等が用いられている。
- 14 50語以上の単語における Jaccard 係数0.1以上の共起関係では、2つの単語が1,000件のレビュー中最低でも10件（平均1件/商品）は同時に登場することとなる。
- 15 weight は Jaccard 係数を示している。
- 16 <http://news.livedoor.com/article/detail/3717663/>（「livedoor ニュース－楽天市場グルメ大賞のウナギ、中国産を四万十川産と偽装」）