

Title	テレビドラマの構造化と評価要因分析：自由回答文のテキスト解析による分析と解釈
Sub Title	Structuring and evaluation of TV dramas : analysis and interpretation of free answers using text mining
Author	妹尾, 紗恵(Seno, Sae)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2007
Jtitle	Keio SFC journal Vol.7, No.2 (2007.) ,p.110- 125
JaLC DOI	10.14991/003.00070002-0110
Abstract	<p>従来では考えられなかった規模のテキストをネットワークから獲得することが可能になった。本稿では、テレビドラマに関する大量の自由回答文を、分析的、かつ解釈的に解析するための解析フローを提案する。第一にtf-idfを用いて自由回答文を自動分類し、ドラマをクラスタリングすることにより、ドラマを分析的に構造化した。第二にTextlmiとキーグラフを使ってドラマの構造を比較分析することにより、視聴者がドラマに求める要素を解釈的に抽出した。その結果から、視聴者のドラマに対する多様なニーズを明らかにした。</p> <p>It becomes realistic to get an amount of information which could not be possible until recently. This article presents a quantitative and interpretive analytical method for large amount of free answers from the Internet about TV dramas. Firstly TV dramas are structured quantitatively by the automatic classification of free answers and clustering of TV dramas. Secondly factors expected from audience are picked out interpretively by comparative analysis of the structure of TV drama using Textlmi and Key Graph. From the above results, we evaluate TV dramas based on various needs of audiences.</p>
Notes	自由論題 研究論文
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=0402-0702-0900

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

◆研究論文◆

テレビドラマの構造化と 評価要因分析

自由回答文のテキスト解析による分析と解釈

Structuring and Evaluation of TV Dramas
Analysis and interpretation of free answers using text mining

妹尾 紗恵

慶應義塾大学大学院政策・メディア研究科後期博士課程

Sae Seno

Doctoral Program, Graduate School of Media and Governance, Keio University

従来では考えられなかった規模のテキストをネットワークから獲得することが可能になった。本稿では、テレビドラマに関する大量の自由回答文を、分析的、かつ解釈的に解析するための解析フローを提案する。第一に tf-idf を用いて自由回答文を自動分類し、ドラマをクラスタリングすることにより、ドラマを分析的に構造化した。第二に TextImi とキーグラフを使ってドラマの構造を比較分析することにより、視聴者がドラマに求める要素を解釈的に抽出した。その結果から、視聴者のドラマに対する多様なニーズを明らかにした。

It becomes realistic to get an amount of information which could not be possible until recently. This article presents a quantitative and interpretive analytical method for large amount of free answers from the Internet about TV dramas. Firstly TV dramas are structured quantitatively by the automatic classification of free answers and clustering of TV dramas. Secondly factors expected from audience are picked out interpretively by comparative analysis of the structure of TV drama using TextImi and Key Graph. From the above results, we evaluate TV dramas based on various needs of audiences.

Keywords: テレビ視聴質、テキストマイニング、自動分類、キーグラフ、構造分析

1 目的

1日1世帯当たりのテレビ平均視聴時間に変化はないが、視聴率30%以上の大ヒット番組が25年前の1%以下に激減しているという現状がある¹。つまり、視聴率の分散化が起きており、その要因としてライフスタイルの多様化に伴う番組選択の多様化が挙げられる。かつてのように視聴者がみな同様に大ヒット番組を見る時代であれば「どの番組がどれだけ(量)見られたか」という指標である“視聴率”のみでも十分であったが、近年の多様化した現状を捉えるためには「どの番組がどのように(質)見られたか」という指標である“視聴質”からの評価が必要とされている。通常、視聴率に対する視聴質のスコアとしては“満足度²”があるが、それらスコアのみからでは、多様化した視聴者の具体的なニーズを捉えることができないのが実状である。

本稿の目的は、視聴率及び満足度を踏まえた上で、意味内容を含んだ視聴質データである自由回答文に基づいて、視聴者が番組をどのような視点からどのように受け止めているのかを明らかにすることである。さらに、視聴率や満足度の決定要因、すなわちドラマの評価要因がどこにあるかを自由回答文から抽出することである。制作者の想いに偏重しがちであった制作の現場に、視聴者の声である自由回答文を持ち込むことで、制作者及び視聴者の双方向からのドラマ制作が可能となる。その結果もたらされる、広告的価値の向上や、ターゲット層の特定化などを通して、テレビ文化が活性化するような戦略的な番組制作を行うための第一歩として本研究は位置づけられる。

以上の目的を達成するため、インターネット調査により、従来の質問紙調査では考えられなかった規模の自由回答文を獲得し、そこから有意な知見を導き出すための、解析フローの構築を行った。第一に、自由回答文から、テレビドラマの構造化を分析的に行う。自由回答文に基づいてドラマ間の類似性を求めることにより、ドラマの相関図を描き出す。第二に、その相関図から得られた5つのドラマ類型間の比較を解釈的に行う。各ドラマ類型に対する自

由回答文のテキスト解析を行うことにより、それぞれの特性を明らかにする。この二つの分析を通して、ドラマ類型ごとにドラマの評価要因を明らかにする。

2 分析データ

対象とするテレビドラマは、2003年4月から2005年12月までに放送された121の連続ドラマに限定した³。自由回答文は、テレビ視聴質調査サイト『リサーチQ』⁴において各番組に寄せられた「この番組の良かった点・悪かった点をなるべく具体的にお願いします。」という設問に対するものを使用した。対象番組への全回答者数は9,030名であり、全自由回答数は129,082件だった。回答者の特性としては、女性が全体の85%を占め、平均年齢は男性が38.4歳、女性が35.2歳となっている。

3 解析フロー

本稿の解析フローは、次の3段階から構成される。第1段階の「ドラマの構造化分析」では、自由回答文の自動分類とKohonenネットワークによるクラスタリングの結果からテレビドラマ相関図を作成することで、ドラマを5つのドラマ類型に振り分けた。第2段階の「ドラマ類型間の比較分析」では、「基礎意味チャンクキーグラフ」を用いてテレビドラマ相関図を自由回答文から掘り下げることにより、各ドラマ類型の特性を明らかにした。第3段階の「ドラマ類型別の評価要因分析」では視聴率の高低を決定付ける要因、すなわち評価要因を明らかにするとともに、視聴率と視聴質の双方の視点からドラマの再評価を行った。

4 ドラマの構造化分析

自由回答文をもとに121のテレビドラマの構造化を行った。この構造化分析ではテキストの持つ意味内容は考慮せず、定量化された形態素のみを解析対象とする。テキストマイニングを使ったコメントからのドラマ(映画)の分類としては(阿部他、2002)の例がある。

構造化分析は二段階から構成される。第一段階と

して、全自由回答文を設定した分類軸に沿って振り分けた。その際、約 13 万件の自由回答文を手作業で分類することは困難であるため、自動分類の手法を用いて機械的に分類を行った。第二段階として、自由回答文の分類結果をもとに、テレビドラマの相関図を作成した。テレビドラマ相関図を水平方向に区切るものとして視聴率という量のデータ、垂直方向に区切るものとして自由回答文という質のデータを用いた。

4.1 自由回答文の自動分類

4.1.1 分類軸の決定

分類には『キャスト』『ストーリー』『感情移入』の分類軸⁵を設定した(表1)。「どのような評価がされているか」という評価の意味内容を一度放棄し、「何が評価されているか」という評価の対象のみを分類軸とすることで、テキストを定量的に扱うことが可能となる(鈴木、2006)。

4.1.2 形態素スコア表の作成(自動分類の手法1)

自由回答文の自動分類を行うために、その前提となる形態素スコア表を作成した。形態素スコア表は学習用データである約1万件⁶の自由回答文に基づく。作成の流れを以下に述べる。

1. 学習用データを手作業で『キャスト』『ストーリー』『感情移入』『その他⁷』の4つに分類する。複数の分類軸に分類される場合もある。
2. 『キャスト』『ストーリー』『感情移入』に分類された自由回答文を形態素解析⁸する。
3. 各ドラマ固有の形態素を除去するために、二番組以上の自由回答文で利用されている形態素(一般

語)を抽出する。

4. 各分類軸と一般語の連関度を求める。連関度の算出には $tf \cdot idf$ ⁹ というスコアリング手法を利用した¹⁰。

作成された形態素スコア表は、約 3600 語の形態素と各分類軸との連関度が示されており、この連関度をもとに自由回答の自動分類を行う。

4.1.3 形態素スコア表に基づく自動分類(自動分類の手法2)

学習用データから作成された形態素スコア表をもとに、ある自由回答文を分類軸のいずれかに自動で分類する方法を述べる。

1. 自由回答文を形態素解析する。
2. 1で得られた形態素全てにおいて、形態素スコア表のキャスト・ストーリー・感情移入との連関度($tf \cdot idf$)を参照する。
3. 形態素と分類軸の連関度を分類軸ごとに合計することにより、自由回答文と分類軸の連関度を求める。
4. 自由回答文を、連関度が最も大きい分類軸へ分類する。

1～4の作業を全自由回答文に対して行うことによって、いずれかの分類軸に分類した。その結果、キャストに関する自由回答文が30,863件、ストーリーに関する自由回答文が64,707件、感情移入した自由回答文が33,471件、分類不可能が41件となり、キャスト・ストーリー・感情移入の割合は、それぞれ24%・50%・26%となった。また、分類精度は80.0%

表1 自由回答の分類軸

分類	説明	例(「anego・アネゴ」より)
キャスト	キャストに関する評価	「篠原さんの演技がとても良い」「黒澤くんかっこいい!」
ストーリー	ストーリーに関する評価	「これからの展開が楽しみです」「ドキドキするくらい面白かった」
感情移入	感情移入した感想	「自分のことをみているようで怖かった」「アネゴと黒澤くんにくっついて欲しい。」

であった¹¹。これをドラマごとに集計するとドラマの特色が見えてくる(表2)。

表2 ドラマごとの自動分類結果

番組名	キャスト	ストーリー	感情移入
鬼嫁日記	24.6%	49.1%	26.3%
オレンジデイズ	24.2%	50.3%	25.5%
anego・アネゴ	23.7%	49.0%	27.3%
大奥 第一章	23.0%	51.5%	25.5%
元カレ	22.6%	50.6%	26.8%
花より男子	29.2%	55.7%	15.1%
女王の教室	14.8%	49.7%	35.5%
ビギナー	25.5%	54.4%	20.1%
砂の器	25.1%	53.4%	21.6%
ウォーターボーイズ2	25.0%	53.2%	21.7%
ウォーターボーイズ	23.7%	54.2%	22.1%
恋におちたら	22.5%	54.6%	22.9%
：	：	：	：



図1 クラスタリング結果(レイヤー2)

4.2 テレビドラマ相関図の作成

4.2.1 レイヤー分割

テレビドラマを評価する量の指標であり、テレビ局にとって最も重要な数値である平均視聴率¹²でドラマを4分割した。レイヤー1を平均視聴率20%以上、レイヤー2を15～19.9%、レイヤー3を12～14.9%、レイヤー4を12%未満と定めた。

4.2.2 レイヤー内でのクラスタリング

テレビドラマを評価する質の指標である自由回答文から、レイヤーごとにクラスタを作成した¹³(図1)。自由回答文の自動分類から得られたドラマごとのキャスト・ストーリー・感情移入率(表2)を変数とし、Kohonen ネットワーク(Kohonen, 1989)を用いてクラスタリングを行うことにより¹⁴、視聴者の視点(評価の対象)が類似したドラマがクラスタを形成し、121の番組が46のクラスタに縮約された。

4.2.3 レイヤー間のクラスタ連結

視聴率は異なるが視聴者の視点は似ているクラスタのつながりを明らかにするために、異なるレイヤーにあるクラスタ同士の類似性を算出した。類似

性の数値化には一致度というスコアを設定した。一致度は、クラスタごとにキャスト・ストーリー・感情移入率の平均値を求め、分類軸ごとに割合の値を標準化した、その標準化得点の2クラスタ間での差の総和である¹⁵。この値は3つの分類軸からみた時の、2クラスタ間の差異の大きさを表しており、一致度が小さい値であるほどクラスタ同士は類似している。あるクラスタから上位レイヤーのクラスタ群を見たときに、その中で一致度が最も小さいクラスタと線で結んだ。

4.3 テレビドラマ相関図とドラマ5類型

ドラマの構造化分析によって得られるのがテレビドラマ相関図である(図2)。テレビドラマ相関図の上部から順に視聴率の高いレイヤーが並んでおり、各レイヤー内は自由回答文の分類結果をもとにクラスタ化されている。また、クラスタ同士をつなぐ実線上にあるスコアは一致度である。

テレビドラマ相関図、すなわち3分類軸それぞれの割合に基づくドラマ分類の結果から、5つのドラマ類型を取り出した。ドラマ類型ごとにキャスト率・ストーリー率・感情移入率の平均を求め、その値をもとに、感情移入1、感情移入2、複合的、キャ

類型 1

類型 2

類型 3

類型 4

類型 5

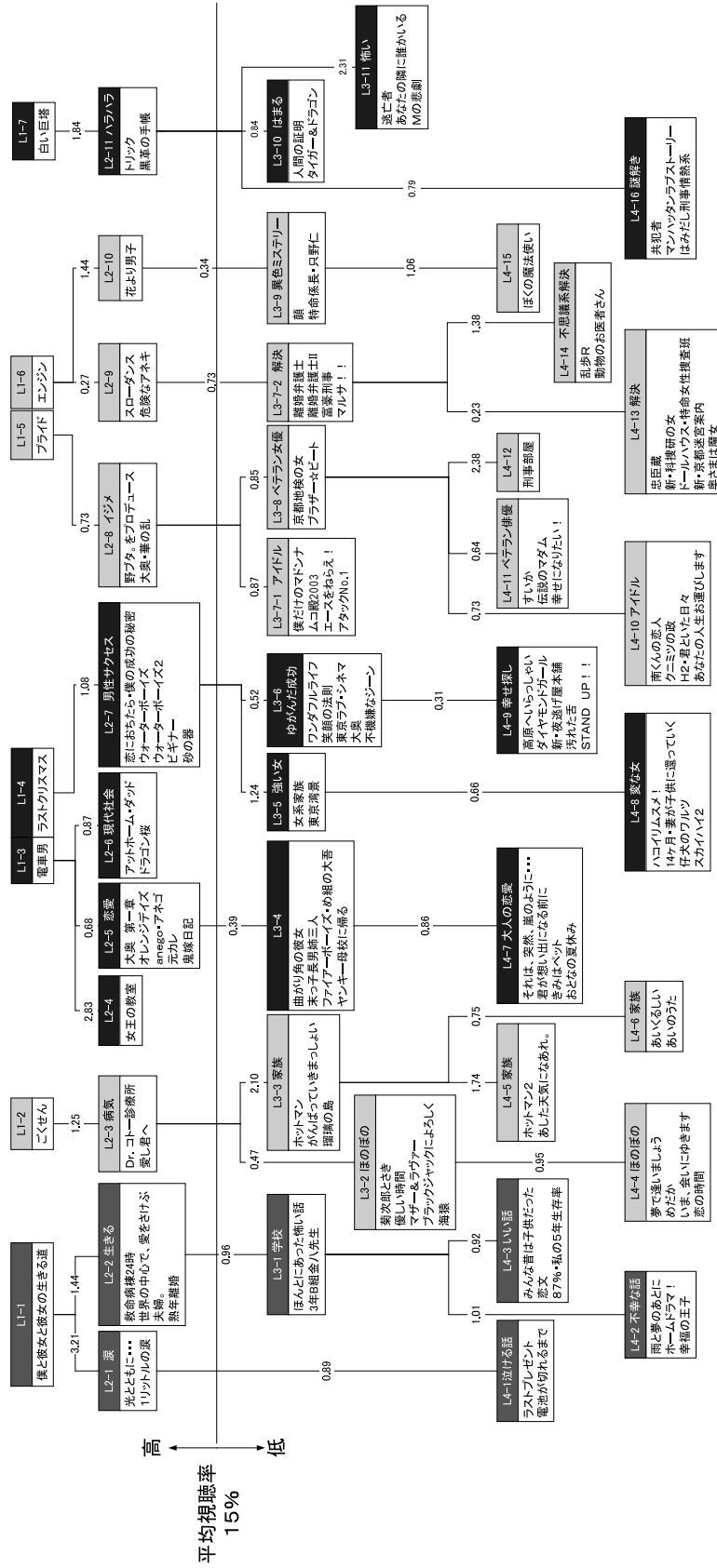


図 2 テレビドラマ相関図

表 3 ドラマ 5 類型の特性

ドラマ類型	類型 1	類型 2	類型 3	類型 4	類型 5	
キャスト率	18.0%	26.7%	22.8%	29.5%	18.6%	
ストーリー率	41.8%	42.4%	52.9%	50.1%	61.6%	
感情移入率	40.1%	30.9%	24.3%	20.4%	19.8%	
類型名	感情移入 1	感情移入 2	複合的	キャスト	ストーリー	
ドラマ数	17	19	39	35	11	
回答総数	17,975	17,772	42,386	36,354	14,549	
平均視聴率の平均	13.6	13.6	13.7	12.8	13.0	
性別	男性	2526(14.1%)	2861(16.1%)	5969(14.1%)	6373(17.5%)	2260(15.5%)
	女性	15449(85.9%)	14911(83.9%)	36417(85.9%)	29981(82.5%)	12289(84.5%)
平均年齢	37.2	37.2	36.4	37.3	36.7	

スト、ストーリー、という類型名を設定した(表3)。ドラマ類型と視聴者属性には強い関連は見られなかった。

5 ドラマ類型間の比較分析

ドラマの構造化分析では、テキストの持つ意味内容を放棄し、定量化された形態素のデータからドラマ 5 類型を求めたのに対し、本章では再びテキストに戻り、その意味内容からの 5 類型間の比較分析を行う。テレビドラマ制作者側の視点からのドラマの内容分析はこれまでも行われてきたが(村松、1978; 岩男、1998; Hicks, 1999)、我々は自由回答文という視聴者側の視点からドラマの内容に踏み込むことで、その視聴質(どのように見たか)を明らかにする。

比較分析には、基本的な意味のまとまりである「基

礎意味チャンク」(深谷、2003)を単位としたキーグラフ(大澤、1999)である「基礎意味チャンクキーグラフ」という独自の手法を用いた。さらに、その結果を補完するという位置づけで、特化係数を用いて、形態素の特徴語を抽出した。また、解釈者の手によって意味的に同一であると判断された言葉をまとめるために、同義語辞書¹⁶を作成・反映した上で分析を行った。

5.1 ドラマ類型ごとの特徴語抽出

形態素解析の結果から、特化係数¹⁷を用いて求めた各類型の特徴語(各ドラマ類型における出現確率が0.5%以上で特化係数の上位10語)を表4に示す。基礎意味チャンクキーグラフだけでなく、この分析を行った理由は、形態素単位で集計した時に高頻度であった言葉が、基礎意味チャンク単位でも高頻度

表 4 ドラマ類型ごとの特徴語

感情移入 1			感情移入 2			複合的			キャスト			ストーリー		
特徴語	頻度	特化	特徴語	頻度	特化	特徴語	頻度	特化	特徴語	頻度	特化	特徴語	頻度	特化
病気	608	4.72	暖かい	232	3.23	ハッピーエンド	478	1.70	アニメ	258	3.23	犯人	224	3.54
生きる	243	3.43	癒す	98	3.10	男	742	1.59	マンガ	597	2.56	ハラハラする	91	2.89
涙	515	3.30	ほのぼのする	95	2.94	恋	267	1.43	カッコいい	710	1.76	ドキドキする	203	2.77
辛い	185	3.24	親子	174	2.80	女	1057	1.40	イメージ	403	1.73	まだまだ	90	2.62
命	99	3.07	優しい	148	2.54	仕事	257	1.28	原作	802	1.72	ますます	202	2.62
学校	132	2.97	生徒	115	2.36	映画	432	1.24	犯人	255	1.60	医者	96	2.46
夫婦	132	2.95	家族	364	2.24	いよいよ	254	1.23	キャラ	315	1.53	1時間	77	2.31
考える	837	2.93	息子	103	2.15	二人	384	1.22	違和感	191	1.47	怖い	530	2.00
泣く	913	2.84	母親	541	2.12	先生	396	1.22	テンポ	519	1.47	目	270	1.98
重い	162	2.79	父親	393	2.06	つまらない	473	1.20	前回	357	1.42	いよいよ	137	1.93

ある。家族を中心とし、[景色がきれい(12)・景色が良い(6)]、[子供がかわいい(60)・女の子がかわいい(14)]ということも含めて、暖かくほのぼのと癒されるドラマの世界を“感受”するという視聴スタイルがとられる。このジャンル名を、“癒し系ドラマ”とする。

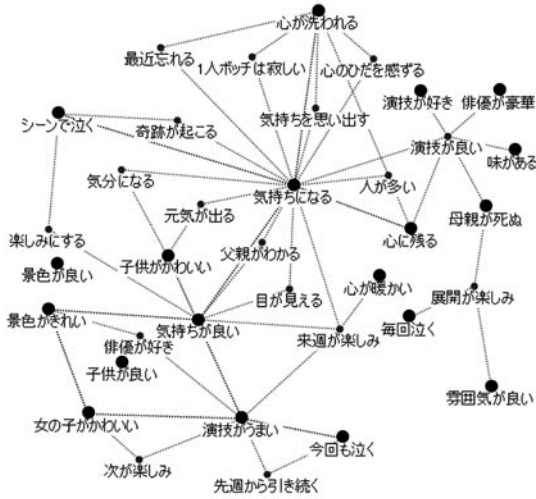


図4 癒し系ドラマのキーグラフ

(3) 複合的ドラマ

キーグラフ(図5)の中心には[ハッピーエンドが良い(141)]があり、このジャンルは“仕事と恋に揺れる男と女のハッピーエンドストーリー”である。結末があらかじめ定められているため、[先の読める(46)]ものではあるが、その過程で[勉強になる(52)][元気になる(30)][女は怖い(33)]というように、幸福物語のドタバタを[結構面白い(175)]と“共感”するという視聴スタイルをとる。このジャンル名を“トレンド系ドラマ”とする。

(4) キャストドラマ

キーグラフ(図6)の中心には[俳優がかっこいい(61)]が据えられ、ほとんどのノードがそこにつながることから、キャストのかっこよさを軸として物語が展開するジャンルである。かっこいいの[俳優]なのかキャラなのかは次章で明らかになる



図5 トレンジ系ドラマのキーグラフ

が、いずれにせよ、アニメやまんがを原作とする傾向があることからわかるように、主人公の漫画的なほどの強い個性によって物語が展開する。このジャンルはかっこいい俳優を見て“充足”するという視聴スタイルがとられ、その後には何も残らないという意味で、ジャンル名を“消費系ドラマ”とする。

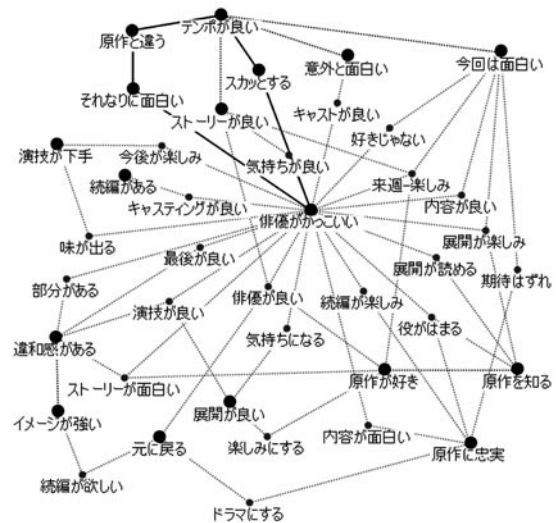


図6 消費系ドラマのキーグラフ

(5) ストーリードラマ

キーグラフ(図7)の中心には[見応えがある(19)]と、[来週が待ち遠しい(51)]の2つがある。見応えとは、[誰が犯人(26)・犯人がわかる(10)]、[謎だらけ(25)・謎が解ける(22)・謎がある(9)]といった複雑性や、ハラハラ・ドキドキ、[スリルがある(13)]といった緊張感からもたらされる。次回につながる[謎]が[来週が待ち遠しい]という強い継続性を生む、という繰り返しで物語が展開していくのがこのジャンルの特性である。ストーリーに引き込まれる一方で感情移入する対象としての登場人物は存在せず、ドラマと視聴者の距離が保たれるこのドラマジャンルの視聴スタイルは“鑑賞”である。

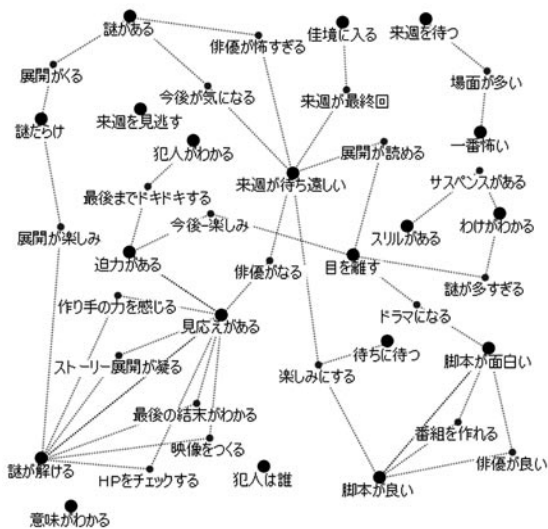


図7 濃密系ドラマのキーグラフ

また、複雑性・緊張感・継続性で表せるこのドラマジャンルを“濃密系ドラマ”とする。

6 ドラマ類型別の評価要因分析

前章まででドラマ類型の比較分析を通して5つのドラマジャンルを解釈的に浮かび上がらせた。ドラマジャンルごとに視聴者の視点は異なるため、当然、ドラマの評価要因も異なる。本章では、ドラマジャンルごとに、その内部での定量的な評価の差異が引き起こされた要因を、自由回答文から明らかにする。

6.1 視聴率と視聴質の連関分析

定量的な評価の指標として、視聴率と満足度を用いた。視聴されない限りドラマ制作は無意味であるし視聴者がその内容を評価することも不可能であることから、テレビドラマを評価する第一の指標を視聴率とする。一方で、視聴率が高いからと言って視聴者が満足しているとは限らないし、逆に視聴率が低くても特定のターゲットに高い満足を提供している場合もある。そこでドラマ評価の第二の指標として満足度のスコアを加味した。

ドラマ類型別に平均視聴率15%以上の高視聴率ドラマと15%未満の低視聴率ドラマの2つに分け、それぞれのグループごとに満足度の平均値を求めた。そして、ドラマを視聴率の高低で分けたとき、それに相関して満足度にも有意な差²²が見られるジャンルを“量質連動型”、逆に視聴率と満足度に有意な差がない、もしくは比較的差が小さいジャンルを“量

表5 ドラマ類型と視聴率からみた満足度

ドラマ類型		感情移入1	感情移入2	複合的	キャスト	ストーリー
ドラマジャンル		イタイ系	アットホーム系	トレンドィ系	消費系	濃密系
満足度 平均値	高視聴率 15%以上	4.57 (n=10019)	4.56 (n=4647)	4.42 (n=23880)	4.26 (n=11280)	4.56 (n=6721)
	低視聴率 15%未満	4.36 (n=8633)	4.37 (n=13925)	4.24 (n=20388)	4.27 (n=26904)	4.44 (n=8421)
視聴率高低間の 満足度の差		0.21	0.19	0.18	-0.01	0.12
信頼区間 95%	上限	0.24	0.22	0.2	0.01	0.15
	下限	0.19	0.17	0.16	-0.03	0.1
				量質連動型	量質非連動型	

質非連動型”とした(表5)。ここでの“量”とは視聴率を、“質”とは満足度を指す。量質連動型に属す低視聴率のドラマは満足度も低いことを意味し、視聴率と満足度のスコアを上昇させるべく改善を行う必要がある。一方、量質非連動型に属す低視聴率のドラマは、満足度においては高視聴率ドラマと遜色なく、低視聴率であることをドラマの内容の持つ特性として許容することができる可能性がある。

6.2 自由回答文からの評価要因分析

ドラマの視聴率および満足度の差異が生じた要因を探るために、各ドラマ類型内において、高視聴率ドラマ(平均視聴率15%以上)と低視聴率ドラマ(同15%未満)の自由回答文の比較分析を行った。5.1で抽出したドラマ類型ごとの特徴語(特化係数1.1以上)に限定し、その中で視聴率高低それぞれに特化した形態素のリストを作成し、視聴率高低の決定要因とした(表6~10)。加えて、形態素を説明するものとして、その形態素を含む基礎意味チャンクの頻度の高い上位3チャンクを添えた。本章では、視聴率高低で分けて基礎意味チャンクキーグラフを作成するには元データの分量が不足したため、キーグラフは使用しなかった。以下の解釈では、「」内は特徴語やそれに関連する基礎意味チャンクを示す。

6.2.1 量質連動型のドラマジャンルにおける評価要因

視聴者の期待に応え満足度を上昇させることが高視聴率獲得に直結するのが量質連動型のドラマジャンルである。ここでは低視聴率ドラマが高視聴率と

高視聴率を獲得し得なかった要因を視聴率から明らかにする。

(1) イタイ系ドラマ

高視聴率ドラマと低視聴率ドラマの対比から見えてくるものは、第一に、高視聴率ドラマが「病気」(克服できること)であるのに対して、低視聴率ドラマが「命・死ぬ」(克服できないこと)である。第二に、高視聴率ドラマは「夫婦・父親・母親・親子・親・娘」というように家族に関する言及が多く、それらが「考える」に係っているのに対して、低視聴率ドラマにはそれが無い。ここから読み取れることは、視聴者はただ主人公と“痛み”を共有したいのではなく、それを克服する過程を通して考えるきっかけを求めているのだということである。家族との絆のないところで孤独に「命を考える」ことは、「怖い・暗い・重い・悲しい」ことでしかない。

視聴者はただ不幸な話に涙したいのではない。不幸な出来事に直面した主人公と現実世界での自分と重ね合わせ、考えさせることが出来るかどうか、視聴率の決定要因となる。

(2) 癒し系ドラマ

癒し系ドラマにおける視聴率高低による差異は、“暖かくほのぼのとした心洗われる感動”の舞台が、高視聴率ドラマは「島」や「学校(先生・生徒)」であるのに対して、低視聴率ドラマは「家族(親子・母親)」である点である。後者の家族を舞台とした癒し系ドラマとは、いわゆるホームドラマと呼ばれてきたものであるが、既に飽きられたのか、家族の

表6 イタイ系ドラマ視聴率決定要因

特徴語	頻度	特化	高視聴率			低視聴率					
			上位意味チャンク(数字は頻度)			上位意味チャンク(数字は頻度)					
夫婦	121	1.70	*.考える(4)	元*.戻る(2)	*.役-合う(2)	怖い	696	1.88	話*(15)	心霊写真*(13)	結構*(10)
病気	487	1.49	*.いう(23)	*.ある(14)	*.なる(12)	命	64	1.40	*.考える(3)	*.大切さ-実感する(3)	*.尊さ-考える(2)
父親	398	1.42	*.なる(9)	*.変わる(9)	*.いく(8)	死ぬ	132	1.31	子供*(4)	私*(4)	母親*(3)
母親	315	1.35	*.気持ち-わかる(7)	*.良い(5)	*.暮らす(5)	時間	91	1.28	*.遅い(7)	*.ずれる(4)	*.欲しい(2)
感動	204	1.32	姿*(3)	場面*(2)	演技*(2)	暗い	73	1.28	内容*(4)	画面*(2)	話*(2)
親子	104	1.32	*.考える(5)	*.関係-良くなる(3)	*.良い(3)	みんな	153	1.26	*.揃う(2)	*.が-集まってくる(2)	
親	191	1.30	*.考える(4)	*.気持ち-伝わる(3)	*.気持ち-わかる(2)	重い	87	1.17	内容*(3)		
伝わる	111	1.29	気持ち*(6)	思い*(4)	親-気持ち*(3)	学校	70	1.15	*.いく(2)		
娘	120	1.29	*.父親-絆-強くなる(2)	*.良い(2)	*.対する(2)	悲しい	187	1.15	姿*(2)		
かわいそう	79	1.28	子供*(8)	俳優*(5)	父親*(3)	暖かい	60	1.14	心*(7)	最後*(3)	

あり方が変わりつつあるのか、「家族愛を感じる」だけでは、今や「感動」を呼べず視聴率は獲得できない。家族の関係性を超えたコミュニティを舞台とする新しい“ホーム”を提示し、「感動」させることができるかが、視聴率獲得のための重要な要素となる。

(3) トレンディ系ドラマ

トレンディ系ドラマで視聴率高低の比較を行うと“仕事と恋に揺れる男と女”における男女の対比が見えてくる。高視聴率ドラマに特徴語として「頑張る男」が出てくるのに対して、低視聴率ドラマの特徴語には「怖い・戦う女」が出現する。従来のトレンディドラマは「優しい・戦う男」と「頑張る女」の物語であったはずが、それが振れた形で逆転せざるをえなかったときトレンディドラマは終焉を迎え、次に表れたのがこのジャンルである。「頑張る男」と「怖い・戦う女」のラブストーリーは成立しないため、「頑張る男」物語と「怖い・戦う女」物語が別々に作られる。そして、マスに受け入れられて視聴率

を獲得できるのは「頑張る男」物語であり、「怖い・戦う女」物語は、基本的に「つまらない」ようである。

6.2.2 量質非連動型のドラマジャンルにおける評価要因

量質非連動型のドラマジャンルにおいては満足度に差異が見られないことから、低視聴率ドラマが高視聴率ドラマを見習うべきであるという単純な図式は成立しない。視聴率の高低を特性としてのみ捉えて、その差異が生じた要因を探る。

(1) 消費系ドラマ

“カッコいい俳優”を中心に構成される消費系ドラマは、俳優がいかにかっこいいかによって視聴率が決定することは容易に想像できるし、実際に高視聴率ドラマでは、特徴語として「(良い・かっこいい・かわいい)俳優」が抽出された。しかしその他は「悪い・飽きる・いまいち・つまらない」といった否定的な言葉がならんでおり、“カッコいい俳優”

表7 癒し系ドラマ視聴率決定要因

高視聴率					低視聴率						
特徴語	頻度	特化	上位基礎意味チャック(数字は頻度)			特徴語	頻度	特化	上位基礎意味チャック(数字は頻度)		
生徒	84	2.92	*-かっこいい(5)	*-良い(3)	*-信ずる(2)	女の子	106	1.26	*-かわいい(20)	*-出る(2)	演技*うまい(2)
島	134	2.16	*-戻ってくる(5)	*-くる(4)	*-出て行く(3)	かわいい	686	1.24	子供*(76)	俳優*(64)	女の子*(20)
病気	63	1.76	*-怖い(3)	*-ある(2)	*-ない(2)	家族	337	1.23	*愛-感じる(10)	*愛-良い(6)	*暖かさ-感じる(4)
続編	65	1.71	*-楽しみ(9)	*-ある(5)	*-欲しい(5)	死ぬ	205	1.23	母親*(20)	お兄ちゃん*(6)	父親*(6)
先生	114	1.67	*-居る(4)	*-良い(2)	*-憧れる(2)	描く	106	1.23	三年間*(3)	丁寧*(2)	
涙	81	1.51	*-出る(9)	*-止まる(6)	*-流す(2)	母親	498	1.23	*-死ぬ(29)	*-なる(9)	*-いう(6)
久しぶり	44	1.51	*-泣く(3)	*-する(2)	*-ドラマ(2)	優しい	136	1.22	俳優*(4)	人*(3)	みんな*(3)
言葉	41	1.49	*-良い(3)			映画	324	1.22	*-違う(14)	*-面白い(8)	*-良い(7)
泣く	223	1.45	今回*(12)	毎回*(8)	久しぶり*(3)	親子	158	1.21	*-いう(3)	*-楽しい(3)	*-なる(2)
感動	76	1.29	最終回*する(4)	*-ある(3)	今*する(2)	時間	157	1.20	*-流れる(10)	*-ずれる(5)	*-過ぎる(3)

表8 トレンディ系ドラマ視聴率決定要因

高視聴率					低視聴率						
特徴語	頻度	特化	上位基礎意味チャック(数字は頻度)			特徴語	頻度	特化	上位基礎意味チャック(数字は頻度)		
映画	360	1.54	*-面白い(32)	*-違う(13)	*-良い(13)	つまらない	261	1.20	話*(7)	内容*(5)	段々*(4)
男	553	1.38	*-なる(6)	*-良い(5)	*-頑張る(5)	結構	332	1.20	*-面白い(103)	*-良い(18)	*-楽しい(8)
先生	282	1.32	*-怖い(9)	*-居る(7)	*-いて欲しい(3)	恋	141	1.15	*行方-気になる(7)	*愛-する(4)	*-する(4)
いよいよ	170	1.24	来週*(5)			初回	130	1.12	*-面白い(10)	*-見逃す(3)	*見逃してしまう(3)
頑張る	185	1.18	男*(5)	私*(3)	みんな*(3)	二人	198	1.12	*-良い(5)	*-なる(3)	*-かわいい(2)
テンポ	279	1.11	*-良い(57)	*-面白い(23)	*-悪い(11)	女	540	1.11	*-怖い(20)	*-戦う(6)	*-いく(6)
楽しみ	2214	1.11	来週*(141)	今後-展開*(75)	展開*(47)	魅力	143	1.10	*-ある(8)	*-ない(5)	*-感じる(5)
主人公	526	1.07	*-ない(7)	*-する(7)	*-良い(6)	面白くなる	244	1.10	展開*(9)	段々*(8)	ドラマ*(3)
面白い	4102	1.06	ストーリー*(106)	結構*(72)	ドラマ*(60)	はまる	284	1.07	結構*(8)	役*(7)	俳優*(5)
展開	1417	1.06	*-楽しみ(168)	今後*楽しみ(75)	*-面白い(52)	ハッピーエンド	226	1.03	*-良い(78)	*-終わる(6)	最終回*(6)

表9 消費系ドラマ視聴率決定要因

高視聴率					低視聴率						
特徴語	頻度	特化	上位意味チャンク (数字は頻度)			特徴語	頻度	特化	上位基礎意味チャンク (数字は頻度)		
魅力	130	1.60	*-ある (10)	*-出る (3)	登場人物*ない(2)	アニメ	241	1.33	*-違う (13)	*-良い (4)	*-好き (3)
悪い	204	1.47	カッコ* (4)	イメージ* (3)	俳優* (3)	犯人	231	1.29	人* (6)	*-わかってしまう (6)	*-わかる (6)
飽きる	79	1.41	俳優* (3)	最後* (2)	内容* (2)	続編	273	1.23	*-楽しみ (42)	*-ある (36)	*-欲しい (17)
いまいち	204	1.40	演技* (8)	内容* (7)	*-面白い (6)	前回	304	1.21	*-面白い (12)	*-違う (7)	*-する (5)
セリフ	91	1.36	*-ある (4)	*-良い (3)	*-ちょっと (2)	昔	253	1.19	*-読む (5)	*-思い出す (5)	*-ある (4)
ストーリー	485	1.33	*-面白い (67)	*-良い (30)	*-つまらない (8)	イメージ	338	1.19	*-強い (28)	*-違う (18)	*-良い (6)
シーン	269	1.26	*-良い (24)	*-ある (13)	*-多い (9)	原作	670	1.19	*-違う (46)	*-忠実 (40)	*-知る (36)
俳優	3643	1.25	*-良い (137)	*-カッコいい (80)	*-かわいい (35)	配役	168	1.16	*-良い (21)	*-面白い (5)	*-合う (3)
演じる	106	1.25	俳優* (30)	俳優-ドラマ* (2)	役* (2)	残念	710	1.14	最終回* (28)	非常* (7)	終わり* (6)
つまらない	152	1.22	ストーリー* (8)	内容* (5)	最終回* (3)	笑う	415	1.13	今回* (10)	姿* (4)	結構* (4)

表10 濃密系ドラマ視聴率決定要因

高視聴率					低視聴率						
特徴語	頻度	特化	上位基礎意味チャンク (数字は頻度)			特徴語	頻度	特化	上位基礎意味チャンク (数字は頻度)		
1時間	49	1.43	*-いう (14)	*-短い (3)	*-早い (3)	犯人	212	1.70	誰* (26)	*-誰 (16)	*-わかる (12)
昔	91	1.42	*-ある (2)	*-ドラマ-思い出す (2)		怖い	444	1.51	俳優* (9)	今回* (5)	シーン* (3)
期待	73	1.19	*-ずれ (11)	*-裏切る (10)	来週* (9)	脚本	104	1.44	*-面白い (17)	*-良い (13)	*-すごい (2)
続く	113	1.14	半年* (2)	話* (2)		意外	80	1.41	*-面白い (8)	*-性-狙う (2)	
現実	40	1.13	*-ある (5)	*-離れる (3)	*-有る (2)	テンポ	122	1.40	*-良い (22)	*-面白い (9)	*-ある (3)
印象	43	1.11	*-強い (6)	*-受ける (3)	言葉* (3)	ハラハラする	69	1.36			
毎週	138	1.10	*-欠かす (11)	*-楽しい (7)	*-見逃す (3)	わかる	509	1.34	わけ* (41)	意味* (14)	犯人* (12)
死ぬ	59	1.05	俳優* (6)	ガン* (5)	人* (2)	最初	120	1.28	*-楽しみ (6)	*-怖い (3)	
残念	152	1.05	最終回* (6)	非常* (4)	ないの* (2)	何	190	1.27	*-わかる (6)	*-起こる (3)	*-面白い (3)
楽しみ	428	1.04	来週* (59)	展開* (40)	*-する (26)	笑う	156	1.23	今回* (6)	毎回* (4)	俳優* (2)

に依存しきったドラマであることが明らかである。一方、低視聴率ドラマでは「アニメ」などを「原作」としたもので、そのキャラクターと俳優との比較においてキャストの「イメージが違う」「配役が良い」「原作と違う」という言及が多くなることから、役柄としての“キャラ”が重視されていることがわかる。よって、高視聴率ドラマと低視聴率ドラマの差異は、俳優重視であるかキャラ重視であるかで説明できるが、このようにキャストにばかりコメントが付くようなドラマは総じて満足度が低く、俳優やキャラクターに頼ったドラマ制作の限界を示唆している。

(2) 濃密系ドラマ

“見応えと継続性”で表される濃密系ドラマにおいて、低視聴率のドラマとは、「誰が犯人」、「テンポが良い」、「ハラハラする」、「怖い」というように、典型的な推理ドラマである。見えない「犯人」を追いかけることで、わけがわからない状態から、少しずつ「わけわかる」状態に展開していく。一方、高視

聴率ドラマは、「誰が犯人」という“謎”が存在しないにも関わらず、「1時間が短い」「来週が楽しみ」というように継続性を維持できている。“謎”に変わってそれをもたらす要素が何なのかは、残念ながらこの分析からは明らかにすることができなかった。ただ、この領域には非常に話題性の高いドラマが集まっており、高視聴率の決定要因があったというよりは、ひとつの作品としての評価を獲得した名作ドラマであり、ドラマ制作者が目指すべき到達点はここにあるのではないかとすら思える。いずれにしても他のジャンルと比較した時に濃密系ドラマの満足度は非常に高く、脚本を重視した作りこまれたドラマは視聴者を強く引き込む魅力を持っている。

7 結論

自由回答文を解析対象に据える場合、一方では、自由回答文をある一定レベルまで計量化することで「分析と実証」としての正当性を確保することが可能になり、他方では、自由回答文のもつ意味内容を

柔軟に取り入れることで「解釈と政策（制作）」としての間主観的妥当性を実現することが可能になる。しかしながら、計量化するとその意味内容が失われ、また意味内容を確保すると計量的に扱うことができないというジレンマが、長年分析者を悩ませてきた。本稿では、ややもすると対立関係にある「分析と解釈」あるいは「実証と政策（制作）」を相補的・同時に活用することで、両者の統合の可能性を具体的に示したという点で、1つの方法論上の貢献を果たしている。本論文がもつ、上記のようなメタレベルでの学術的な価値を前提として、以下、テレビドラマの構造化と類型別の評価要因分析というテーマに沿った個別具体的な結論を明確にする。

第一に、ドラマ類型は量質連動型と量質非連動型に分けられた。量質連動型は視聴質（満足度）が視聴率に直結し両者の関係性は2パターンに限られるため、視聴率という量の指標のみで評価可能である（図8左。番号は表11参照）。視聴率が高いことは、視聴質も高いことを意味しており、そのようなドラマが成功と評価される。一方、量質非連動型においては視聴率と視聴質の間に対応関係はなく、両者の関係性には⑦～⑩の4パターンが考えられる。このジャンルのドラマは視聴率だけでなく、視聴質という視点からの再評価が必要だと言える（図8右。番号は表11参照）。両方とも高ければ、成功ドラマであるが、視聴率だけ高い場合には、マスに受けるという意味での量的な成功ドラマであり、視聴質だけ高い場合には、ターゲットが絞られたという意味での質的な成功ドラマといえる。

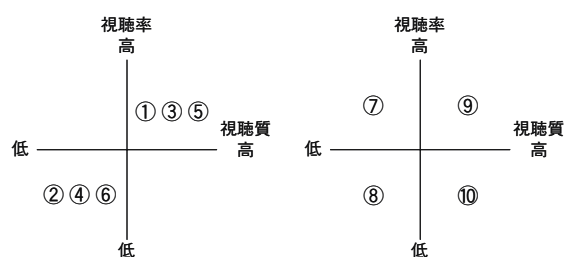


図8 量質連動型（左）と量質非連動型（右）における位置関係

第二に、ドラマジャンルごとに視聴率の決定要因を表11にまとめた。量質連動型のドラマジャンルは、高視聴率ドラマの特性を制作に反映させれば視聴質も上昇する。つまり、視聴率及び視聴質の評価要因は同一である。他方、量質非連動型のドラマジャンルでは、視聴率と視聴質は独立した評価基準である。したがって、表11にあるのは視聴率の決定要因であり、視聴質の決定要因とは異なる。

第三に、量質非連動型ドラマにおける視聴質の決定要因は、視聴者の視点が「キャスト」にあるか「ストーリー」にあるかの違いである。⑦は人気キャストの出演による広い認知により視聴率が高いが、キャスト依存の単純なストーリーであるため視聴質は低い。だからこそ、マスを対象とした消費的なドラマであるかぎり、量的には成功ドラマと評価できる。逆に、⑩はストーリーの複雑さゆえにターゲット層は限られ視聴率は低いが、そのターゲット層に対しては高い視聴質を提供することに成功している。まさに、質的には成功しているドラマである。

表11 視聴率と視聴質の連関とその要因

ドラマ類型		感情移入1	感情移入2	複合的	キャスト	ストーリー
ドラマジャンル		イタい系	癒し系	トレンド系	消費系	濃密系
高視聴率	番号	①	③	⑤	⑦	⑨
	要因	考えさせる	非家族ホームドラマ	頑張る男	俳優中心	(名作)
低視聴率	番号	②	④	⑥	⑧	⑩
	要因	イタいだけ	ホームドラマ	怖い・戦う女	キャラ中心	推理(謎)
評価		成功	成功	成功	量的成功	成功
ドラマタイプ		量質連動型			量質非連動型	
		満足(4.57)	満足(4.56)	満足(4.42)	不満(4.26)	満足(4.56)
		不満(4.36)	不満(4.37)	不満(4.24)	不満(4.27)	満足(4.44)
		失敗	失敗	失敗	失敗	質的成功

かつてのように、視聴者がみな同様に大ヒット番組を見るという時代が過ぎ去って久しい今、視聴者のニーズの多様化を踏まえ、それに合ったドラマを制作することが求められている。ドラマジャンルごとにドラマ制作の手法はまったく異なる。さらに、視聴率と視聴質が連動しない場合には、ターゲットをどこに設定するかによって、視聴率を優先させるか、それとも視聴質を優先させるのか、の判断が求められる。ドラマは、制作者の想いだけではなく視聴者の多様なニーズを把握する視点を導入しないかぎり成功ドラマにはなりえない、というのが現在のドラマ制作の現場である。だからこそ、今回のようなインターネットを介した大量の自由回答文の分析と解釈から、ドラマを再評価していくことが意義をもつのである。

8 今後の課題

第一に視聴率と自由回答文をもとに、キャスト・ストーリー・感情移入の分類軸からドラマの構造化を行い、ドラマを5類型に分類した。自動分類の手法の精査による精度の向上を行うと同時に、今後も継続して構造化の安定性を検証していく。第二に基礎意味チャンクでキーグラフを描くという新しい手法でドラマ類型の比較上の特性を明らかにし、そこからドラマジャンルを作成した。形態素ではなく基礎意味チャンク単位で分析を行うことで妥当な意味内容を把握することが可能となった。一方で、より深く掘り下げていくためには膨大な量のテキストが必要であることが明らかになり、データ取得についても改善策を練らなければならない。第三にドラマの視聴率と視聴質の連関分析とその高低を弁別する要因を明らかにした。今後は本研究で構築した解析フローをベースとして、自由回答文からの個別のドラマの評価や、視聴者の属性に基づく具体的なターゲットの絞り込みを行うことで、実際のドラマ制作現場に役立つ解析を行っていきたい。

注

- 1 朝日新聞 (2004年5月2日付)
- 2 テレビ番組に対する質の評価の指標。リサーチ Q では「この番組を見て、よかったですか?」という問いへの5段階評価。
- 3 3クール以上放送されたドラマ、回答数400以下のドラマ、海外ドラマは除外した。
- 4 登録ユーザが前日の19時から26時(翌日の午前2時)までに放送されたNHK関東圏および民放キー局5チャンネルの地上波放送テレビ番組表の中から、視聴した番組について、期待度・満足度・集中度および自由回答文を回答するものである。
- 5 ドラマは「誰が(キャスト)何をしたか(ストーリー)」から構成されている。そして、そのドラマを見る視聴者がいる限り、そのドラマに対して「何を感じたか(感情移入)」という要素がさらに追加される。この3要素がドラマ制作における最も基本となる要素であり、テレビ朝日との話し合いの結果、分類軸とした。
- 6 2005年4月クールに放送された8ドラマに関する自由回答文を対象とした。分類者の主観の影響が強くなりすぎることを防ぐために8名で作業を分担して行った。
- 7 野球中継延長に関する自由回答などの、ドラマの内容とは関係のない自由回答文。
- 8 本稿での形態素解析は、すべてText Mining for Clementineを使用した。
- 9
$$tfidf_{ik} = tf_{jk} \times idf_i \quad tf_{jk} = w_{jk} / \sum_j w_{jk} \quad K \quad (1)$$

$$idf_i = \log \frac{M}{v_i} + 1 \quad (2)$$
 式(1)について、 w_{jk} は分類軸k(キャスト/ストーリー/感情移入)における単語iの総出現回数、 $\sum_j w_{jk}$ は分野kにおける総特徴語数であり、 tf_{jk} とは分野kにおける単語iの出現頻度となる。また式(2)については、Mは記事総数、 v_i は単語iを含む記事数であり、 idf_i とは全記事から見た単語iの重要度となる(緒方他、2005)。
- 10 自動分類のための特徴素抽出手法の比較研究(石田他、2002)の精度評価を参考に、分析後半のキークラフのとの対応から、形態素を単位とし全ての品詞を特徴素とすることを定めた。さらに、様々な自動分類手法(Joachims 1997; McCallum 1998; Rivest 1987)の中でも、日本語テキストを形態素単位として扱い、分類後の解析が簡単に行えるスコアリング手法は、Naive Bayesを使った研究(阿部他、2002)とtf・idfを使った研究(緒方他、2005)があり、その中でも比較的文字数の少ないコメント単位での同分類を行っていたtf・idfによる手法を採用した。
- 11 ランダムに抽出した1000件の自由回答文を目視で分類した結果と自動分類で得られた結果の合致率は80.0%だった。
- 12 その番組の放送時間内の毎分視聴率を足上げて、放送分数で割ったもの。ビデオリサーチ社による関東地区の世帯視聴率を利用した。
- 13 レイヤー1は番組数が少ないためクラスタリングを行わなかった。
- 14 データマイニングのクラスタリング手法のうち、情報損失量を抑えながら多次元情報を二次元平面に視覚的に再現できるKohonenネットワークが最も適していると判断した。使用ソフトウェアはViscovery SOMineで、いずれの試行もノード数2000、張力0.5、正確-イグゼクト・モードで行った。
- 15 一致度は $\sum_{i=0}^k |f_{1i} - f_{2i}|$ で表せる。 f_{1ij} はあるクラスタの分類軸iにおける割合の標準化得点、 f_{2ij} はそれとは異なるクラスタの同じ分類軸iにおける割合の標準化得点、kは分類軸数(ここでは3)を指す。
- 16 「かなしい」は「悲しい」と同一の意味の言葉であるとして、「かなしい」を「悲しい」に変換するための変換表。特徴語に表れるような重要な単語を中心に、1965語を72語に集約した。

- また、全ての人名は「俳優」という文字列と置換した。
- 17 あるドラマ類型に対する自由回答文において形態素Aが全単語中に占める割合を、全自由回答文において形態素Aが全単語中に占める割合で割った値。その値が1より大きければ、形態素Aはそのドラマ類型に特化している。
 - 18 係り受け関係の自由度が高い言葉ほど、組み合わせが多様化することにより、基礎意味チャンクにしたときの頻度が低くなる傾向があるため。
 - 19 本稿では助詞が異なる、または省略されていても、同一の意味として扱った。本稿で分析対象となり得るだけの頻度がある基礎意味チャンクはほとんどが1個の名詞と述語の2者の係り受け関係であり、それらは助詞によって意味が変わることがほとんどないと判断した。一方、キークラフ及び解釈の文中では、わかりやすくするために、最も多かった助詞を含めて記載している。
 - 20 キークラフとは、チャンス発見の情報技術で、意思決定支援に活用されるテキストマイニングのツールである。東京大学大澤幸生助教授が開発したものを使用した。通常のキークラフの計算方法は使用頻度の高い形態素が黒ノードとして設定され、複数の黒ノードと共起する形態素が赤ノードとして表示される(本稿では黒ノードを大きな黒丸、赤ノードを小さな黒丸で示した)。今回の分析ではドラマ類型を比較しその特徴的な要素を抽出することにあるので、使用頻度ではなく特化係数の高い基礎意味チャンクを強制的に黒ノードに設定してグラフ化を行った。各ドラマ類型において特徴的な基礎意味チャンクを取り出し、その形態素をつなぐ弱い紐帯を赤ノードとして発見することにより、各ドラマ類型の持つ特性を抽出した。
 - 21 キークラフの中央付近にあり多くのリンクがある黒ノードは、そのテキストにおいて単独で重要なだけでなく、その他の重要ノードとも共起しているという意味で、そのテキストを象徴する言葉だと言える。
 - 22 表5の95%信頼区間とは、この区間の中に真の差がある確率が95%であることを意味し、この区間に0が含まれる場合は視聴率に基づく2グループの満足度の平均値の差は有意ではない。

参考文献

阿部 倫子・田中 久美子・中川 裕志「コメントを用いた映画の分類」、『情報処理学会研究報告』、2002 卷 66 号、2002 年、pp.105-110。

石田 栄美・辻 慶太「日本語テキストの自動分類のための特徴素抽出手法の比較」、『情報処理学会研究報告』、2002 卷 87 号、2002 年、pp.81-86。

岩男 寿美子「テレビドラマの内容分析—1977～1994」、『メディア・コミュニケーション』、48 卷、1998 年、pp.1-56。

大澤 幸生・Benson E・谷内田 正彦「Key Graph：単語共起グラフの分割統合によるキーワード抽出」、『電子情報通信学会論文誌』、J82-D1 卷 2 号、1999 年、pp.391-400。

大隅 昇・保田 明夫「テキスト型データのマイニング」、『理論と方法』、19 卷 2 号、2004 年、pp.135-159。

緒方 進・池田 真司・牟田 高信・木本 勝敏「Web 上のテキスト情報を用いた人物評価手法」、『情報処理学会研究報告』、2005 卷 1 号、2005 年、pp.9-14。

鈴木 努「二〇〇五年衆議院選挙における三大紙の社説比較—概念ネットワーク分析の適用」、『マス・コミュニケーション研究』、69 卷、2006 年、pp.2-20。

中野 智仁「日本語テキスト解析ツール TextImi の概要」、日本認知言語学会第 7 回全国大会予稿集。

深谷 昌弘「意味づけ論からソシオセマンティクスへの歩み」、『KEIO SFC JOURNAL』、2 卷、2003 年、pp.28-61。

村松 泰子「女性のテレビドラマ受容過程—受け手および番組内容の総合的分析」、『NHK放送文化研究年報』、23巻、1978年、pp.59-109。

Hicks N, *Screenwriting 101: The Essential Craft of Feature Film Writing*, Michael Wiese Productions, 1999.

Kohonen T, *Self-organization and associative memory*, Springer-Verlag New York, Inc., 1989.

Joachims T, “Text Categorization with Support Vector Machines: Learning with Many Relevant Features”, In *Tenth European Conference on Machine Learning*, 1997, pp.137-142.

Mccallum A, Nigam, K, “comparison of event models for naive bayes text classification”, In *AAAI-98 Workshop on Learning for Text Categorization*, 1998, pp.41-48.

Rivest RL, “Learning decision lists”, In *Machine Learning*, 2(3), 1987, pp.229-246.

Salton G, Buckley C, “Term-weighting approaches in automatic text retrieval”, In *Information Processing & Management*, 24(5), 1988, pp.513-52.