

脚本執筆における概念構造の抽出

—映画・演劇脚本術の計量的比較—

高橋 椋一 村井 源 猪原 健弘
東京工業大学大学院 社会理工学研究科

本研究では、映画・演劇脚本の執筆プロセスを背後で支える概念構造の、共通点と相違点を明らかにすることを試みた。具体的には、「脚本術」と呼ばれる映画・演劇脚本の参考書のうち、米国で出版されているものを対象とした計量テキスト分析を行った。ネットワーク分析を行った結果、映画・演劇脚本の執筆において共通に中心的役割を果たしている概念が抽出された。また、係り受け解析の結果、それらの概念に対する意味づけには、映画・演劇脚本家の間で違いが存在することが明らかになった。この結果は、脚本執筆のストラテジーが表現媒体によって異なることを示唆するものである。

Extracting the Conceptual Structure of Script Writing: Quantitative Comparison of Screenwriting and Playwriting

Ryoichi Takahashi Hajime Murai Takehiro Inohara
Graduate School of Decision Science and Technology
Tokyo Institute of Technology

This paper attempts to investigate the similarities and differences of scriptwriter's conceptual knowledge between writing movie script and writing play script. We numerically analyze the textbooks about how to write a movie script and play script, that are published in the United States. As a result of network analysis, some common key concepts of scriptwriting were extracted. Moreover, utilizing dependency analysis, it is clarified that screenwriter and playwright consider those key concepts from different point of view. This result means that there are different strategies between writing movie script and writing play script.

1. 問題と目的

心理学・認知科学の分野において物語研究は主に、文章理解研究の領域で行われてきた。文章理解研究の主眼は、人間が物語作品を読んだり鑑賞したりする際の認知メカニズムの解明であり、物語作品の受け手が関心の対象となっている。しかし、近年では物語作品の受け手だけでなく作り手にも関心が向けられはじめ、クリエイティブ・ライティング[1]という新たな研究領域においては、作家の物語創作を支える認知メカニズムが検討されはじめている。しかし、文章理解研究と比較するとその規模は小さく、作家が物語を創り出す際の思考・認知のプロセスに関して具体的な知見はあまり得られておらず、未だ明らかになっていない点が多いのが現状である。

従来の人文学的な物語論が扱っているように、物語に関わる要素は、主題、モチーフ、世界観、物語構造、登場人物、時間関係、視点、語り手など、実に多種多様である。しかし、物語に関わるこうした諸概念のうち、どういったものが創作プロセスにおいて中心を占めているのかは明らか

でない。また、作家が種々の概念にどのような関連づけや意味づけを行うことで、物語を効果的に創り出すためのストラテジーを組み立てているかについても、現状では不明な点が多い。本研究ではこうした背景のもと、作家の物語創作プロセスの背後に存在する概念構造を探索的に検討する。

そこで、本研究では「創作論」と呼ばれる、作家が物語創作の技術・ノウハウを記述したテキストを対象として分析を行うことにした。創作論に明確な学術的定義は存在しないが、以下の3点が主な特徴としてあげられる。すなわち、創作論とは、(1)物語創作の初心者向けに、(2)著者が自身の持つ創作の技術を解説した、(3)説明的文章である、ということができる。創作論は、いわば物語創作の「参考書」であり、作家の創作過程における認知や思考のすべてを直接的に反映したテキストであるとはいえない。しかし、物語の創作過程において作家がどのような点を重要視しているかが明示的に記述されているという点で、創作の核となる概念や作家のストラテジーを検討するには、適当なテキストであるということができる。

2. 分析方法

2.1 計量テキスト分析

本研究では、「創作論」に対して計量テキスト分析を行うことで、物語創作の背後にある概念構造を検討する。計量テキスト分析とは、テキスト型データに対して数学的・統計学的手法を用いた定量的解析を行うことで、客観性・科学性を担保した形でテキストから情報を得る分析方法である。コンピュータを用いた計量テキスト分析には、大量のテキストを一様の基準で分析することができるという大きなメリットが存在する。したがって、計量的手法を用いて多くの創作論を分析することで、作家の有する概念的知識の全体像を実証的に検討することが可能であると考えられる。

2.2 ネットワーク分析

現在の計量テキスト分析では、テキスト中に相対的に高い頻度で登場する語句はそのテキストにおいて重要な位置を占めるという頻度論的仮定に基づき、形態素解析を中心とした語彙レベルでの分析が行われることが多い。その際にしばしば用いられるのが、単語をノードとし、単語間の共起関係をエッジとしたネットワーク分析である。

ネットワーク分析とは、複数のノードとそれらを繋ぐエッジによって構成されたネットワークの性質を分析する手法であり、数学におけるグラフ理論に基礎を置いている。ネットワーク分析は社会学や情報学などをはじめとした多くの分野で用いられており、近年ではこうした領域で培われた分析手法を計量テキスト分析にも応用しようとする試みが存在する。たとえば、河瀬・村井・往住[2]や村井・往住[3]は、批評文・評論文に含まれる概念や思想の関係を記述する際に、ネットワーク分析を用いている。これらの研究では、文章中で同時に用いられる単語には何らかの意味的連関があることを仮定し、エッジで結ばれた単語間の関係を分析することで、書き手の思考傾向や価値観をテキストから抽出している。

3. 分析対象

本研究では、「脚本術」と呼ばれる、映画および演劇脚本の創作論を対象として分析を行った。脚本術を分析の対象として選択した理由は、以下の二点である。一点目は、現在、市場においては多くの映画・演劇脚本術が流通しているということである。そのため、脚本術を研究対象とすることで、一定以上のデータ量を確保することが可能である。二点目は、映画脚本は演劇脚本をルーツに持つ[4]ため、両媒体の脚本術は比較可能性が高いということである。物語を生み出すプロセスにおいては、各表現媒体やジャンルの違いに関わらず共通している側面と、それぞれの媒体に固有の側面という、二つの側面が存在していると考えら

れる。そこで、映画と演劇という二つの表現媒体における創作論を比較することで、領域一般的な知見と領域依存的な知見を同時に得ることが可能である。

また、本研究では脚本術のうち、米国で出版されているものを分析対象とする。これは、ハリウッドやブロードウェイをはじめとした米国映画・演劇は商業的成功を収めていることから、その脚本術に記述されている物語創作のノウハウは、一定の質的水準を満たしていると考えられるためである。

分析対象となる書籍の選定にあたっては、米国のオンライン書店 Amazon.com および Burns & Nobles の売上ランキングを参考にした。具体的には、映画脚本 (screenwriting) カテゴリーおよび演劇脚本 (playwriting) カテゴリーにおいて両書店で共に上位 100 位以内にランクインしている書籍の中から分析対象を選定した。また、著者の偏りを避けるため、同一著者による書籍は 2 冊以上選出しないようにした。選出したテキストの概要を表 1 に示す。

表 1 テキストデータの概要

| | 映画 | 演劇 |
|-------------|---------|--------|
| テキスト数(冊) | 13 | 10 |
| テキストサイズ(MB) | 6.8 | 5.4 |
| 総単語数(語) | 1089144 | 877375 |
| 平均単語長(文字) | 4.269 | 4.306 |
| 平均単語出現頻度(回) | 50.242 | 43.066 |

4. 脚本執筆における中心的概念の分析

本研究ではまず、脚本の執筆における中心的概念を明らかにするため、名詞を対象とした分析を行う。そこでまず、電子化されたテキストに対して TreeTagger[5]を用いた形態素解析を行った。つづいて、頻出する名詞の共起ネットワークを作成した。その際、共起の強さの指標としてはシンブソン係数を用いた。また、出現頻度が低い単語はテキストを代表するキーワードであるとは考えにくいため、本研究では閾値を設定し、出現頻度が閾値以下の単語を分析から除外した。閾値設定の基準としては、単語の累積出現頻度を用いた。具体的には、映画・演劇脚本術のそれぞれにおいて全単語の出現頻度を大きい方から累積加算し、累積出現頻度が 90%にあたる単語の出現頻度を求めた。その結果、90%に相当する出現頻度は、映画が 47 回、演劇が 39 回であり、これらの値を閾値として設定した。

図 1 は、グラフ描画パッケージである Graphviz を用いて、映画脚本術における名詞の共起関係のネットワークを視覚化したものである。ノードのフォントサイズは単語の出現頻度の二乗根に、エッジの長さはシンブソン係数の二乗に比例している。可視性のため、出現頻度が上位 45 までの

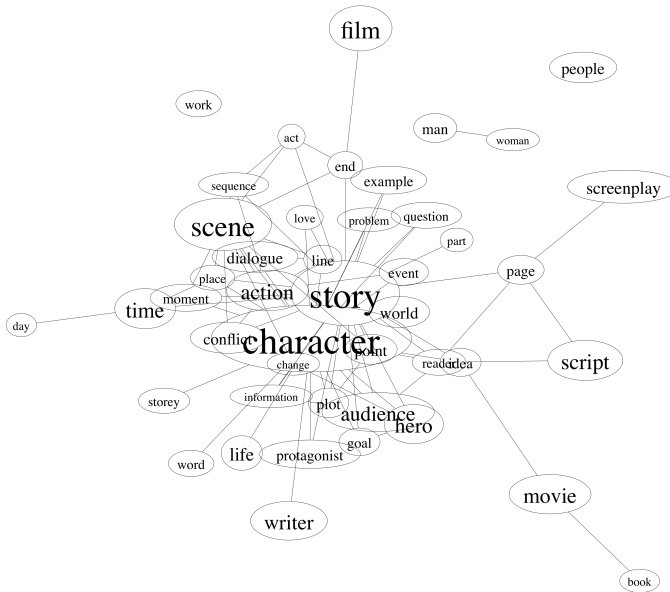


図 1 映画脚本術における名詞の共起ネットワーク



図 2 演劇脚本術における名詞の共起ネットワーク

表 2 映画・演劇脚本術における名詞の中心性

| Degree | | Closeness | | | | Betweenness | | | | Eigenvector | | | | | |
|-----------|-----|------------|-----|-----------|-------|-------------|-------|-----------|-------|-------------|-------|-----------|-------|------------|-------|
| 映画 | | 演劇 | | 映画 | | 演劇 | | 映画 | | 演劇 | | 映画 | | 演劇 | |
| story | 686 | play | 644 | story | 0.879 | play | 0.932 | story | 0.038 | play | 0.063 | character | 0.376 | character | 0.405 |
| character | 681 | character | 614 | character | 0.875 | character | 0.896 | character | 0.038 | character | 0.056 | story | 0.315 | play | 0.370 |
| scene | 626 | playwright | 471 | scene | 0.825 | playwright | 0.756 | scene | 0.031 | time | 0.025 | scene | 0.179 | story | 0.142 |
| film | 603 | story | 470 | film | 0.805 | story | 0.755 | film | 0.027 | playwright | 0.025 | film | 0.149 | scene | 0.113 |
| time | 564 | time | 469 | time | 0.775 | time | 0.755 | time | 0.024 | story | 0.024 | action | 0.132 | action | 0.110 |
| writer | 554 | audience | 445 | writer | 0.767 | audience | 0.735 | writer | 0.021 | audience | 0.021 | audience | 0.103 | playwright | 0.106 |
| movie | 516 | scene | 435 | movie | 0.740 | scene | 0.728 | movie | 0.018 | scene | 0.019 | time | 0.103 | audience | 0.105 |
| audience | 507 | life | 397 | audience | 0.734 | life | 0.700 | audience | 0.017 | life | 0.017 | hero | 0.103 | time | 0.101 |
| action | 506 | action | 387 | action | 0.733 | action | 0.693 | action | 0.016 | action | 0.016 | writer | 0.101 | three-act | 0.091 |
| script | 497 | people | 383 | script | 0.727 | people | 0.690 | life | 0.016 | people | 0.015 | script | 0.088 | stage | 0.073 |

名詞をノードとして、語彙間の共起の強さが上位 90 までのものをエッジとして描画している。図 2 は演劇脚本術のネットワークであり、図 1 と同様の手順で作図した。

つづいて、図 1 および図 2 のネットワークに対して中心性分析を行った。中心性分析とは、ネットワークにおける重要ノードを数学的に抽出する手法であり、ネットワーク分析においては主要な分析手法のひとつとなっている。中心性の算出方法は中心性概念の捉え方に依存しており、これまでに数々の指標が提案されている。本研究では、Freeman[6]がまとめた 3 つの中心性である次数中心性、近接中心性、媒介中心性と、Bonacich[7] によって提案された固有ベクトル中心性の 4 種類を代表的なものとして選択した。各中心性の計算には、NetworkX[8]を用いた。表 2 は、映画およ

び演劇の脚本術に登場する各名詞の中心性の値を、上位 10 位まで示したものである。

図 1、図 2 および表 2 を見てわかるように、映画においては **character** と **story** という 2 つの概念が脚本の執筆において最も中心的な位置を占めている。一方、演劇において中心となっているのは **play** である。これは、演劇においては物語の構成要素よりも、脚本がどう演じられるかということが関心の対象となっていることを示すと考えられる。

また、表 2 からは、脚本執筆において中心となる概念の多くは映画・演劇に共通であることがわかる。特に、**character**、**story**、**scene**、**time**、**action**、**audience** という 6 語は、全ての中心性指標において映画・演劇で共通して 10 位以内に含まれている。したがって、これら 6 つの概念は表現媒体

に依存せず、脚本の執筆において常に主要な関心の対象になっていると考えられる。

5. 脚本の執筆ストラテジーの分析

本節では、脚本執筆における中心的概念に対する映画・演劇脚本家のアプローチを検討することで、脚本執筆におけるストラテジーの共通点・相違点を分析することを試みる。

手法としては、具体的には、前節で抽出した6つの名詞と係り受け関係にある動詞を検討する。動詞は行為や動作を表す品詞であるため、物語に関わる概念を表す名詞にどのような動詞が結びついているかを分析することで、それらの概念に対する脚本家の働きかけや操作といった様相を分析することが可能であると考えられる。また、

分析に係り受け解析を用いる理由は、単語間の関係性を検討する場合には、単純な共起情報よりも係り受け情報を用いた方がよりノイズを低減できると一般に考えられているためである。

分析ではまず、TreeTaggerによってタグ付けされたテキストに対して、MaltParser[9]を用いた係り受け解析を行った。つづいて、名詞と動詞の係り受け共起の頻度をカウントし、その頻度からシンプソン係数を算出した。表3には、分析の結果それぞれの名詞と強く結びついていた動詞を、シンプソン係数が大きい順にそれぞれ上位10位まで示した。以下では、表3で示した名詞と動詞の係り受け関係に基づき、脚本家の執筆ストラテジーを考察する。

表3 名詞と係り受け共起関係にある動詞

| character | | | | story | | | | scene | | | |
|-----------|-----------|-------|-----------|--------|-----------|-------|------------|----------|------------|-------|-----------|
| 映画 | | 演劇 | | 映画 | | 演劇 | | 映画 | | 演劇 | |
| 0.146 | introduce | 0.245 | activate | 0.389 | progress | 0.174 | center | 0.131 | advance | 0.065 | center |
| 0.145 | react | 0.138 | prevent | 0.213 | advance | 0.118 | tell | 0.100 | craft | 0.056 | trigger |
| 0.100 | undergo | 0.124 | affect | 0.183 | unfold | 0.104 | sum | 0.085 | construct | 0.056 | revise |
| 0.091 | propel | 0.123 | construct | 0.132 | enhance | 0.091 | unfold | 0.057 | arrange | 0.055 | unfold |
| 0.091 | prevent | 0.104 | motivate | 0.127 | propel | 0.073 | evolve | 0.053 | enter | 0.053 | construct |
| 0.089 | define | 0.098 | struggle | 0.127 | tell | 0.059 | structure | 0.052 | rewrite | 0.051 | mark |
| 0.085 | depict | 0.094 | introduce | 0.105 | slow | 0.057 | begin | 0.050 | analyze | 0.050 | count |
| 0.083 | support | 0.091 | flesh | 0.105 | push | 0.034 | drive | 0.048 | belong | 0.049 | separate |
| 0.079 | evoke | 0.088 | emerge | 0.102 | reduce | 0.034 | present | 0.046 | end | 0.049 | enter |
| 0.073 | speak | 0.079 | address | 0.081 | structure | 0.028 | plan | 0.043 | plan | 0.042 | sustain |
| time | | | | action | | | | audience | | | |
| 映画 | | 演劇 | | 映画 | | 演劇 | | 映画 | | 演劇 | |
| 0.123 | spend | 0.075 | spend | 0.115 | advance | 0.113 | interrupt | 0.118 | confuse | 0.127 | engage |
| 0.115 | waste | 0.051 | list | 0.086 | surround | 0.110 | pursue | 0.094 | hook | 0.111 | grab |
| 0.058 | limit | 0.040 | invest | 0.073 | propel | 0.098 | justify | 0.085 | guess | 0.106 | confuse |
| 0.052 | revise | 0.036 | block | 0.070 | slow | 0.078 | motivate | 0.085 | sense | 0.105 | address |
| 0.031 | divide | 0.026 | cry | 0.063 | view | 0.056 | enable | 0.081 | remind | 0.095 | guess |
| 0.029 | shift | 0.024 | mention | 0.060 | eliminate | 0.051 | describe | 0.080 | surprise | 0.085 | entertain |
| 0.021 | stick | 0.024 | take | 0.052 | reverse | 0.048 | express | 0.080 | anticipate | 0.083 | remind |
| 0.021 | marry | 0.023 | flesh | 0.051 | imply | 0.041 | complicate | 0.080 | satisfy | 0.078 | invite |
| 0.021 | recall | 0.022 | center | 0.046 | indicate | 0.040 | influence | 0.067 | respond | 0.070 | laugh |
| 0.021 | mark | 0.020 | structure | 0.044 | alter | 0.036 | unfold | 0.066 | identify | 0.065 | surprise |

5.1 “character”との係り受け語

character と結びついている動詞には、映画と演劇で異なる傾向が見られた。映画においては introduce, depict, define といった、人物の紹介や人物説明に関わる動詞が多く見られる。ここから、映画脚本の執筆においては、いかに人物像を視聴者に明確に伝えるかということが重要視されていることがわかる。

一方、演劇においては、人物の造形に関わる動詞が多く結びついている。たとえば affect や motivate は、登場人物に影響を及ぼしている事象や動機といった、人物像の形成に関わる動詞である。また、construct や flesh は登場人物の組み立てや肉付けに関わる動詞であり、activate は生き活きとした登場人物の造形に関わる動詞である。これらの動詞が character と強く結びついていることは、演劇脚本家が精緻な人物造形に重きを置いていることを示唆している。

また、prevent や struggle といった障害や苦難に関わる動詞は、映画および演劇で共通に見られる。ここからは、登場人物を逆境に追い込むことでドラマ性を高めようとする、映画・演劇脚本家の共通のストラテジーを読み取ることができる。

5.2 “story”との係り受け語

story と強く結びついている動詞は、映画と演劇で大きく異なっている。映画においては、progress, advance, propel, push, slow と、速度や移動、前進に関わる動詞が強く結びついている。これは、映画脚本家はストーリー構築において、物語の展開性やスピード感といった要素に重きを置くことを示していると考えられる。

一方、演劇においては center と sum といった中心・要点に関わる動詞や、structure や plan といった物語の構造化や設計に関わる動詞が強く結びついていた。ここからは、演劇脚本家が構造的なストーリーを組み立て、物語の主軸を明確に示そうと意識していることが読み取れる。

5.3 “scene”との係り受け語

scene と結びついている動詞においても、映画と演劇の脚本術で異なったパターンが見られた。まず、映画においてのみ見られた特徴的な動詞からは、映画脚本家はシーン間の関係性を重視している可能性が高いことを読み取れる。たとえば、arrange はシーンの配列に関わる動詞であり、advance は次のシーンへの進行を示す動詞である。

一方、演劇においてのみ見られた特徴的な動詞には、中心的なシーンを表す center, 物語の引き金となるシーンを表す trigger, 特徴的なシーン

を表す mark といったものがあげられる。こうした動詞が結びついていることは、演劇脚本家が「見せ場」となるハイライトシーンを強調することに重きを置いている可能性を示唆している。

5.4 “time”との係り受け語

time は、映画においては spend, limit, waste といった時間の消費・制限を表す動詞と強く結びつく傾向が見られた。これは、厳密に定められた上映時間内に収まる脚本を執筆する必要があるという映画制作における制約を、脚本家が強く意識していることを示すものである。

一方、演劇において time と結びついている動詞に一貫した関連性はみられなかった。したがって、形容詞など他の品詞との関連性からも検討を行う必要がある。

5.5 “action”との係り受け語

action と結びついている動詞においてもまた、映画・演劇脚本術で異なった傾向が見られた。映画において action は、story と同じく、advance, propel, slow といった速度に関わる動詞と結びついていた。これは、映画脚本家は登場人物の取る行為をスピード感やテンポといった観点から捉えていることを示すものである。

一方、演劇においては、行為の動機・理由づけを表す動詞 motivate や、行為が結果として持つ影響を表す動詞 influence, interrupt が結びついていた。ここからは、行為そのものというよりも、行為の持つ文脈的要素に重きを置く演劇脚本家のストラテジーを読み取ることができる。

また、action は、映画においては imply や indicate といった暗示に関わる動詞と結びついている一方、演劇においては describe や express など明示的な表現に意味する動詞と結びついたり、これは興味深い特徴であるといえる。

5.6 “audience”との係り受け語

audience と結びついている動詞には、映画と演劇で共通の傾向が見られる。まず、hook, identify, engage, grab など、物語への没頭に関わる動詞はどちらにおいても見られる。したがって、物語の世界へ観客を引き込むことは、どちらの表現媒体においても重要視されているといえる。また、anticipate, guess, surprise など、予測や推測、意外性に関わる動詞も共通に結びついている。これは、観客が期待する事柄を読み取って演出する能力が脚本家には求められることを示唆している。

前節までの分析の結果、映画・演劇の脚本執筆のストラテジーには相違点が多く見られている

にも関わらず、脚本家が聴衆から引き出したい心理的反応には類似点が多いことは、興味深い結果であるといえよう。

6. 結論と今後の課題

本研究では、映画および演劇の脚本術を収集し、それらのテキストに対して計量的分析を行った。ネットワーク分析の結果、映画・演劇脚本の執筆において共通に中心的役割を果たしている6つの概念が抽出された。また、係り受け共起にもとづく分析の結果、それら6つの概念に対する意味づけには、映画・演劇脚本家の間で違いが存在しており、脚本執筆のストラテジーが表現媒体によって異なることが示唆された。本研究で得られた知見は、作家の知識構造に関しての仮説を提供するだけでなく、物語作品の自動生成といった工学的分野への応用にもつながるものであるといえる。

今後はより多くのジャンルや表現媒体の創作論を収集・分析することで、本研究で得られた知見を更に精緻化する必要があるといえる。

謝辞

本研究は科研費「レトリカルデータベースシステムの構築による計量的修辞分析手法の確立」(22700256)の助成を受けた。

参考文献

- 1) Kaufman, S. B. and Kaufman, J. C. (Eds.): *The Psychology of Creative Writing*, Cambridge University Press (2009).
- 2) 河瀬彰宏, 村井 源, 往住彰文: 音楽評論文にみる概念構造の変遷-ネットワーク中心性を用いた音楽概念の抽出, 情報知識学会誌, Vol.19, No.2, pp.138-143 (2009) .
- 3) 村井 源, 往住彰文: 文芸批評の計量解析による批評行為の背景的特徴の抽出, 情報知識学会誌, Vol.20, No.2, pp.117-122 (2010) .
- 4) Nannicelli, T.: *A Philosophy of the Screenplay*, pp.79-110, Routledge (2013).
- 5) Schmid, H.: Probabilistic Part-of-Speech Tagging Using Decision Trees, *Proc. Proceedings of international conference on New Methods in Language Processing*, pp.44-49 (1994).
- 6) Freeman, L.: Centrality in social networks conceptual clarification. *Social networks*, Vol.1, No.3, pp.215-239 (1978).
- 7) Bonacich, P.: Factoring and Weighting Approaches to Status Scores and Clique Identification, *Journal of Mathematical Sociology*, Vol.2, No.1, pp.113-120 (1972).
- 8) Hagberg, A., Swart, P. and Schult, D.: Exploring network structure, dynamics, and function using NetworkX, *Proc. Proceedings of the 7th Python in Science Conference (SciPy2008)*, Varoquaux, G., Vaught, T. and Millman, J. (Eds), pp.11-15 (2008).
- 9) Nivre, J., Hall, J. and Nilsson, J.: MaltParser: A Data-driven Parser-generator for Dependency Parsing, *Proc. Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC2006)*, pp.2216-2219 (2006).