

演劇評価尺度の構築 ——Twitterによる語彙収集ならびに妥当性・信頼性の検証——*

Development of the Theatrical Performance Evaluation Scale: Collection of the Adjectives via Twitter as well as Confirmation of its Validity and Reliability

正田 悠[†], 山下 瑛司[‡]

[†]Haruka Shoda, [‡]Eiji Yamashita

[†]立命館大学スポーツ健康科学部

[‡]神戸大学国際人間科学部

[†]College of Sport and Health Science, Ritsumeikan University

[‡]Faculty of Global Human Sciences, Kobe University

[†]shoda@fc.ritsumei.ac.jp

概要

演劇は紀元前から現在に至るまで人々に広く受け入れられている舞台芸術の一つである。本研究では演劇作品を鑑賞することによって生じる経験や演劇作品に対する評価を定量的に評価するための尺度を構築することを目的とした。Twitterによる語彙収集、インターネット調査による尺度構築、および演劇の動画視聴による尺度の妥当性と内的一貫性の検証を行った。その結果、5因子26項目からなる演劇評価尺度(Theatrical Performance Evaluation Scale; TPES)を構築した。今後、実際の舞台作品を鑑賞直後にTPESを評価してもらうことで、TPESの舞台鑑賞への適用可能性を探る必要がある。

キーワード: 演劇 (theatrical performance), 評価 (evaluation), 尺度構築 (scale development), Twitter

1. はじめに

演劇は紀元前から現在に至るまで広く受け入れられてきた舞台芸術である。近年、演劇は人々にとって疎遠なものとなりつつあったが、COVID-19の影響によって映像配信を導入する劇団が増え、観劇そのものに対する敷居はそれほど高くなってきた。従来の認知科学研究において演劇の観客の心理的反応を得るためには、観客が実際に舞台に足を運び(あるいは研究者自らが舞台芸術の場を設定し)、アンケート票を手渡しその場で回収する等、多大な労力が必要であった。演劇作品を身近に鑑賞できるようになってきた現在においては、むしろオンライン上でさまざま

な演劇を鑑賞し、鑑賞中の心理的体験が測定しやすくなってきていると考えられる。本研究は、将来の演劇の定量的研究に資するため、演劇の鑑賞者の心理的反応を測定する心理尺度を構築することを目的として行った。

演劇と同じ実演芸術である芸人の「ネタ」に関しては日本語版の印象評価尺度が開発されている(野村, 2010)。ヒュース・岡田(2016)は野村(2010)の芸人評価/位置づけ尺度(GELOS)を用いて即興演劇と台本演劇における観客が受ける印象の違いを分析しており、演劇の要素のうち少なくとも一部はGELOSによって弁別可能であることを示している。その一方で、演劇の印象を評価するためには、演技や演出の技術性や独自性への評価、観客が演者や演出、物語に対して抱く情動である「タスク情動」、観客が物語に没入し、登場人物に共感することによって生じる「共感性情動」といった、観客が演劇から受ける幅広い印象を含めることが必要である(Konjin, 1999)。

本研究では、マイクロブログであるTwitterを用いて、幅広い人々からの記述に基づいて演劇の印象評価語の収集を行った。Twitterでは作品全体の総評を書く人もいれば、役者の演技や演出面に言及する人もいる。観劇を終えてすぐに記述する人が多いため、鑑賞時に抱いた情動がそのまま反映されていると考えられる。研究1ではTwitterによる語彙収集および語の取捨選択を行い、研究2ではインターネット調査により過去5年間に演劇の観劇経験のある人を対象に、探索的因子分析を用いた尺度構築を行った。研究3では、実際に映像を通して演劇を鑑賞してもらい、研究2で構築した尺度を検証した。

*本研究は、第二著者が神戸大学国際人間科学部に提出した令和2年度卒業論文に基づく。

2. 研究1: Twitter による語彙収集

研究1では、歌声の印象評価語を収集した金礪・中野・後藤・菊池(2016)を参考に、Twitter上の演劇に関する投稿に基づいて演劇評価尺度の候補となる語彙を収集し、演劇評価尺度のプロトタイプとなる評価語を選定した。

2.1 Twitter による評価語の収集

舞台の口コミサイトである ENJOY THEATER (<https://theater.enjoycinema.net>) に掲載されている、2019年(COVID19の影響前)に上演された153公演に関するツイートを対象に、印象評価語の収集を行った。Python 2.7上でGetOldTweets-pythonパッケージ(Henrique, 2018)を使用して、上記の153公演のタイトルを含む過去1年間のツイートをすべて収集した。その結果抽出された276,584ツイートを対象に、KHCoder(樋口, 2014)によって形態素解析を行ったところ、75,403語の形態素が抽出された。この際、名詞・サ変名詞・形容詞・形容詞B・形容動詞・副詞・動詞を評価語の収集対象の品詞とした。固有名詞や英語、未知語(「ファンサ(ファンサービスのこと)」のような若者言葉等)、感動詞(「ありがとう」等)は除外した。その結果、評価語の候補として18,393語を抽出した。この18,393語から、出現数が100回未満の評価語を除外し、さらに、明らかに作品の印象とは判断できない語(「チケット」等)や単体では意味が不明確な語(「めっちゃくちゃ」、「頑張る」)など、演劇評価尺度として不適切と考えられる抽出語を除外し、114語を選出した。さらに、先行研究(Konjin, 1999)で使用されている16語のうち、これら114語と重複のない10語を加えた124語を評価語の候補とした。これらの10語は、英語圏への留学経験がある大学生2人とともに英語を日本語に訳した。

2.2 了解性調査

了解性調査は、収集した印象評価語124語の中から、演劇の印象を連想させやすい評価語を選定することを目的に行った。18—23歳($M = 20.82$, $SD = 1.26$)の学部生20人を対象に、ウェブ上のアンケートを用いて調査を行った(男性11名、女性9名)。20人のうち18名が演劇経験者で、経験年数は1—9年間であった($M = 3.44$, $SD = 2.16$)。金礪他(2016)にしたがって、20名に「評価語が表す演劇(演技、芝居)をどのくらい想像できるか」について、0(想像できな

い)から10(想像できる)の11段階で回答を求めた。20名の回答の平均値を了解性得点の代表値とし、了解性が11件法の中央である6より小さい18語を同義性調査から除外した。

2.3 同義性調査

同義性調査では、了解性調査によって選定された評価語106語の中から、連想させる演劇の印象が類似する評価語どうしを統合することで、互いに独立した意味を持つ評価語のみを残し、仮尺度を構成することを目的とした。これ以降、評価語対が表す演劇の印象が類似することを「似ている」と表現する。

2.3.1 総当たり表の調査

まず、19—23歳($M = 20.51$, $SD = 1.11$)の学部生10人を対象に調査を行った(男性5名、女性5名)。金礪他(2016)の調査手順にしたがって、選定された106語に対応する縦横106マスの表を用い、「対象の評価語と似たような演劇(演技、芝居)を表す評価語」を選択するよう求めた。全5,565($= 107 \times (107 - 1) / 2$)通りの評価語対のうち、3人以上の評価者が「似ている」と回答したのは621対であった。この段階でどの評価語とも「似ていない」と判断された評価語12語は、他の語が表す演劇の印象とは類似せず、独立した印象を表す語であるため、仮尺度の評価語として確定させた。

2.3.2 「作品の内容」と「鑑賞者の体験」を表す語の分類

「似ている」と判断された評価語の内容を見ると、評価語が物語や演技、演出といった内容に対するもの、および観客自身の情動に対するものの2種類があるように見受けられたため、この2つを基準として評価語を2つに分類した。この分類は、演劇および演出経験のある3人の大学生が行った。まず、選定された評価語106語について「作品の内容への評価を表す言葉である」、「作品の内容への評価を表す言葉ではない」の二択で回答を求め、次に「鑑賞者の体験(情動)を表す言葉である」、「鑑賞者の体験(情動)を表す言葉ではない」の二択で回答を求めた。

その結果、3人とも「作品の内容への評価を表す言葉である」と答えた評価語を「内容への評価語」、3人とも「鑑賞者の体験(情動)を表す言葉」と答えた

評価語を「情動を表す評価語」、それ以外の評価語を「(内容と情動の)両方を表す評価語」とコーディングした。3人の分類について、Cohenの κ 係数を算出したところ、 $\kappa = .55-.70$ と中程度以上の一致を示した。これを踏まえ、39語を「内容への評価語」、29語を「情動を表す評価語」、26語を「両方を表す評価語」として分類した。この分類に基づいて、同じ分類の評価語対を「似ている」と判断し、次の調査対象とした。一方、異なる分類の評価語対(例えば、「内容への評価語」と「両方を表す評価語」)は「似ていない」と判断し除外した。その結果残った503対を対象にさらに類似度調査を行った。

2.3.3 類似度評定

類似度評定は、これまでの2つの調査を経て「似ている」と判断された評価語対がそれぞれどの程度「似ている」のかを確認し、類似度が一定の基準を超えた評価語対を統合することで、仮尺度の評価語を確定することを目的として行った。19—23歳($M = 19.88$, $SD = 0.83$)の大学生男女25名(男性5名, 女性20名)を対象にウェブ上のアンケートページを用いて調査を行った。評価者25名はいずれも演劇未経験者であった。25名に、選定された503対について、2つの評価語が表す演劇(演技・芝居)がどの程度似ているか、1(似ていない)—7(似ている)の7段階評価で回答を求めた。

各回答者の回答を平均した値を同義性とし、同義性が4.5以上であった二対の語は、互いに同義性が高いと判断した。この際、3語以上にわたって同義性が高かった場合は、その中から了解性が最も高い評価語を選択した。この結果、同義性が4.5以上の評価語対は「内容への評価語」では48個、「情動を表す評価語」では33個、「両方を表す評価語」では22個残った。これらの評価語対を統合し、了解性が最も高い評価語のみを残したところ、「内容への評価語」では12語、「情動を表す評価語」では7語、「両方を表す評価語」では6語の、計25語の評価語が選定された。この25語に、総当たり表で除外された12語を加えた37語を仮尺度とした。

3. 研究2：オンライン調査による尺度構築

研究2は、研究1で得られた37語の仮尺度から因子を抽出し、それらの信頼性と妥当性を検証することを目的とした。

3.1 方法

3.1.1 スクリーニング調査

調査はオンライン調査会社に依頼して実施した。まず、過去5年間に舞台演劇の観劇経験がある者を抽出し、かつSatisfice(目的を達成するために必要最小限を満たす手順を決定し、追求する行動)を除外するためにスクリーニング調査を行った。スクリーニング調査では、過去5年間に劇場で観劇した経験を持つことを調査協力者の条件として、その作品名を挙げることができるか否か、映像で演劇を観る頻度、演劇の重要性(「全く重要でない」、「重要でない」、「やや重要でない」、「どちらとも言えない」、「やや重要である」、「重要である」、「非常に重要である」の7件法)についての項目を尋ねた。さらにSatisficeが含まれるデータを除外するためのダミー項目を2問を尋ねた。調査協力の条件を満たした5000人のうち、ダミー項目で除外されなかった回答者は991人であった。

3.1.2 本調査

調査参加者 スクリーニング調査で残った991人のうち、本調査への回答に応じたのは527人であった。そのうちダミー項目に対し適切に回答した362人を分析対象とした(男性150人, 女性212人, 18—50歳, $M = 37.75$, $SD = 8.50$)。

調査項目 調査協力者の性別や年齢(年代)については、調査会社に事前登録されているため質問項目を設けなかった。観劇した作品名とそれを観た西暦年、作品の印象の評価、演劇経験、過去10年間で観劇した回数について尋ねた。

作品の印象評価について、仮尺度と芸人評価/位置づけ尺度(GELOS, 野村, 2010)からなる57語の評価語のそれぞれにどの程度当てはまるかを「1:全く当てはまらない」—「7:非常に当てはまる」の7件法で回答するよう求めた。また、スクリーニング調査と同様、Sacrificeを含むデータを除くため2つの「ダミー項目」を設けた。

分析 TPESの仮尺度37項目に対して、因子の単純構造を目的とする直接オブリミン解による探索的因子分析(最尤法)を行った。このとき因子数は平行分析に基づいて決定した。独自性が極端に高い語、因子負荷量がどの因子においても.35以下である評価語は演劇の印象を構成する因子には寄与しない項目であると判断し、除外した。

表1 TPES の各下位尺度の Cronbach の α 係数, 尺度得点の平均値および標準偏差, ならびに尺度得点間相関

	Cronbach's α	M(SD)	尺度得点間相関				
			F1	F2	F3	F4	F5
F1: クオリティ	.93	5.50 (1.01)	—	< .001	< .001	< .001	< .001
F2: 悲劇	.81	2.96 (1.19)	-.30	—	.14	< .001	.01
F3: 感動	.82	4.80 (0.99)	.71	.08	—	< .001	< .001
F4: ユーモア	.78	4.58 (1.23)	.57	-.41	.32	—	< .001
F5: 迫力	.89	5.19 (1.00)	.85	-.14	.67	.45	—

注: 尺度得点間相関は下三角が相関係数, 上三角が p 値を示す。

表2 GELOS と TPES の各尺度得点とのピアソンの相関係数

TPES	GELOS		
	独自性	技術安定性	リズムにぎやかさ
F1	.52***	.68***	.74***
F2	.02	-.04	-.48***
F3	.45***	.64***	.42***
F4	.48***	.47***	.77***
F5	.61***	.75***	.67***

*** $p < .001$

3.2 結果

3.2.1 探索的因子分析および内的一貫性の分析

平行分析で因子数を設定した後, 探索的因子分析を実施し, 因子パターンの絶対値の最大値が .35 を上回らない項目があった場合にはその項目を削除するという手続きを繰り返し, 最終的に 37 語のうち 26 語を尺度の項目として採択した。これらの因子に対し, 因子パターン値が高い語を参考に, F1 を「クオリティ」, F2 を「悲劇」, F3 を「感動」, F4 を「ユーモア」, F5 を「迫力」と命名した。これらの因子の尺度得点間相関に関しては, 悲劇因子以外の 4 因子については互いに有意な正の相関関係が認められた。「悲劇」因子に関しては, 「感動」因子との相関は有意ではなく, クオリティ, ユーモア, 迫力の各因子とは有意な負の相関を示した。

各因子について, 信頼性係数である Cronbach の α 係数を算出した (表 1)。最も小さい値でも「ユーモア」因子の $\alpha = .78$ であり, 十分な内的一貫性が確認されたといえる。

3.2.2 基準関連妥当性

基準関連妥当性の尺度として用いた GELOS の各尺度得点と TPES の各尺度得点との相関係数を表 2 に示す。GELOS の独自性因子と技術安定性因子については, 「悲劇」因子を除いて有意な正の相関が示された。リズムにぎやかさ因子については, いずれの因子についても有意な相関があることが示され, 「悲劇」因子のみ負の相関を示した。したがって, 本研究で抽出された 5 つの因子のうち「悲劇」因子を除くすべての項目で, GELOS との関連がみいだされ, これらの因子は実演芸能としての演劇の側面を表す収束的妥当性を示すものと考えられる。また, 「悲劇」はリズムにぎやかさ因子とは負の相関を示し, 他の項目とも有意な相関が示されなかったことから, 「悲劇」因子は, 本研究で構築した演劇評価尺度独自の因子であるといえ, 従来の尺度との弁別性を確保する項目であったといえる。

3.3 考察

本調査では, 研究 1 で作成した仮尺度から演劇に特有の因子を抽出することで演劇評価尺度を作成し, その信頼性と妥当性を検証した。仮尺度からは 5 つの因子が抽出され, 「クオリティ」は「悲劇」因子以外の 3 因子と正の相関を示した。これは「クオリティ」因子が具体的な印象を評価する因子ではなく, 「良い」や「面白い」など, 作品への包括的なポジティブ感情を表す項目から構成されるためであると考えられる。

基準関連妥当性に関して, GELOS の各因子は演劇評価尺度の「悲劇」因子以外の 4 因子と正の相関を示した。「悲劇」因子が GELOS のどの因子とも正の相関を示さなかったのは, 演劇と「ネタ」の指向性における根本的な差異によるものであると考えられる。芸人の話芸においては, 全ての「ネタ」にとって最も重要なことが観客を笑わせることであるのに対して, 演

劇は作品によって観客に与える印象が異なる。本研究で演劇の感想や口コミを収集した 153 公演の感想を見ると、コメディ要素が終始強いと思われる作品は少なく、「胸が苦しくなった」、「涙が止まらなかった」など、感動的な物語を連想させるような作品が多かった。またリズムにぎやかさ因子については、「悲劇」因子と負の相関を示した。演劇の悲劇的な場面では、舞台上の役者同士でセリフがあまり交わされなかったり、もし交わされたとしても会話の間や沈黙が効果的に使われていたりするため、そのシーンが観客に与える印象は重く、暗いものであるといえる。つまり演劇評価尺度の「悲劇」因子と GELOS のリズムにぎやかさ因子は対極の関係にあるため、負の相関を示したと考えられる。「悲劇」因子以外の因子と GELOS のリズムにぎやかさ因子は正の相関を示したものの、「クオリティ」因子、「ユーモア」因子、「迫力」因子に比べると、「感動」因子との相関はやや小さかった。芸人の「ネタ」においてテンポ感やそれに伴うにぎやかさが重要であるのと同じで、同じ舞台芸術である演劇においてもセリフの掛け合いのテンポ感は重視される。演劇では、自然な会話を演出する上で、良いテンポ感を生むために、自分の前の人が台詞を言い終わるのを待たずに自分の台詞を入れる発話重複が多く（後安，2006），この技術を演劇では「ケツカミ」ということがある。この「ケツカミ」が生むテンポ感によって場面の疾走感や迫力を表現したり、良いテンポ感の中であえて間を作って台詞を入れることによって笑いを誘ったりするなど、演劇においてテンポ感が果たす役割は非常に大きい。そのため、「クオリティ」因子、「ユーモア」因子、「迫力」因子において高い正の相関が示されたと考えられる。以上の点から、「悲劇」因子を除いた 4 因子と GELOS の各尺度との正の相関は説明可能であり、演劇評価尺度の基準関連妥当性が示されたといっていよいであろう。

本調査によって演劇評価尺度の収束的・弁別的妥当性ならびに各因子の内的一貫性が検証された。この尺度を構築するにあたっては、ウェブ調査において「過去に観劇した演劇作品」の印象についての評価を求めたため、あくまで記憶の中の演劇が評価されており、作品を鑑賞した当時の評価がすべて反映されているとは限らない。印象に残っている部分のみが強められている可能性もある。基準関連妥当性は検証されたものの因子的妥当性は検証されておらず、演劇評価尺度の再現性を検証する必要がある。そこで第 4 章では、実際に演劇を映像で視聴してもらい、演劇評価尺度を用いた印象評定実験を通して、本尺度の因子的妥当性を

確認する。

4. 研究 3：演劇の映像視聴実験

4.1 目的

研究 2 で探索的因子分析によって抽出した 5 因子 26 項目の因子構造の信頼性と妥当性を検証することを目的とする。具体的な手順としては、まず大学生を対象に演劇の映像を視聴してもらい、その作品を演劇評価尺度によって評価するよう求めた。その結果をもとに、研究 2 の探索的因子分析に基づくモデルを構築し、確認的因子分析により、尺度の再現性を確認した。また演劇評価尺度で評価された演劇の尺度得点を比較することで、本尺度に基づいて、異なる映像作品を弁別することができるかを調べた。また、同一作品を 3 つの場面（序盤、中盤、終盤）に分け、作品の中での印象の推移についても調べた。

4.2 方法

4.2.1 実験参加者

18—23 歳 ($M = 19.86$, $SD = 1.64$) の学部生の男女 96 名（男性 25 名、女性 70 名、未回答 1 名）を対象に印象評定実験を行った。異なる印象を持つと考えられる作品 A～作品 C の 3 本の演劇の映像刺激を用意し、それぞれの映像を異なる実験参加者群がオンライン上で視聴した（作品 A： $n = 49$ ，作品 B： $n = 28$ ，作品 C： $n = 19$ ）。

4.2.2 使用した演劇作品

作品 A と作品 B は短編演劇（作品 A：27 分間，作品 B：28 分間），作品 C は長編演劇（作品 C：149 分間）であった。長編演劇である作品 C は、観客が受ける印象の推移を調べるため、物語の展開に合わせて 40 分間，39 分間，69 分間の 3 場面に分割した。以後、分割した映像刺激を、作品の始まりから終わりの順にそれぞれ C1，C2，C3 と表記する。以上の作品の選定にあたっては、世界観や物語の展開，演出手法などがそれぞれ異なり、喚起される印象が多様となるように選定することを考慮した。

4.2.3 質問紙の構成

人口統計学的変数 性別，年齢，学校教育以外の演劇経験の有無，および演劇経験がある場合はその年数を回答するよう求めた。

表 3 5 因子モデル, 1 因子モデル, 2 因子モデル, 双因子モデル, 二次因子分析モデルにおける確認的因子分析モデルの情報量基準および適合度指標。

指標	5 因子 モデル	1 因子 モデル	2 因子 モデル	双因子 モデル	二次因子分析 モデル
自由度	289	299	298	264	294
χ^2	760.5	1662.04	1271.94	410.12	822.58
AIC	10642	11524	11136	8809	10694
BIC	10897	11750	11364	9046	10935

作品の内容に関する質問 適切な回答行動をとったかどうかを確認するため, 各映像の視聴後, 作品の内容に関する 3 つの質問を設け, それぞれ正しい説明がなされている選択肢を選ぶよう求めた。

作品の印象の評価 研究 2 と同様, 仮尺度と GELOS からなる 57 語の評価語のそれぞれにどの程度当てはまるかを「1: 全く当てはまらない」—「7: 非常に当てはまる」の 7 件法で回答するよう求めた。また, ここでも Sacrifice を含むデータを除くため 2 つの「ダミー項目」を設けた。

4.2.4 手続き

ウェブ上のアンケートを用いて質問票を作成し, 映像視聴による実験を行った。実験参加者に対して質問票の中に記載した動画のリンクから映像の視聴ページに移動し, 途中で再生を止めずに視聴するよう指示した。視聴後, TPES の 24 項目を用いてその作品を評価するよう求めた。作品 C に関しては, C1, C2, C3 を視聴するたびにその印象を評価してもらい, 3 つの映像の視聴には間隔を空けず, 各映像の評価が終わり次第, 次の映像を視聴するよう指示した。

4.3 結果

作品の内容に関する質問や作品の印象評価に含まれたダミー項目において, Satisfice を疑うデータは認められなかったため, 実験参加者全員の回答をもとに分析を行った。

4.3.1 因子モデルの検証

TPES の 26 項目について, 複数の因子モデルを設定し確認的因子分析を行った。作成したモデルは (1) 探索的因子分析から得られた 5 因子を想定した 5 因子モデル, (2) 1 つの一般因子が各項目に寄与する 1 因子モデル, (3) 「クオリティ」因子と「感動」因子を作

品の全体的評価を表す 1 因子と想定し, その他の 3 因子をより具体的な印象を表す 2 因子と想定した二因子分析モデル, (4) 5 因子とは別に全項目の背景に 1 つの一般因子を想定する双因子モデル, (5) 5 因子の背景に 1 つの高次因子があることを想定した二次因子分析モデルの計 5 種類のモデルの情報量基準を比較した (表 3)。この結果, AIC および BIC から, 双因子モデルが採択された¹。

このモデルの適合度は $\chi^2(264, N = 134) = 410.12, p < .001, CFI = .94, TLI = 0.92, RMSEA = .064$ であった。 χ^2 値による適合度検定では有意であったが, 他の適合度基準が Hu & Bentler (1999) や West, Taylor, & Wu (2012) の示したカットオフ値 (CFI と TLI が .95 以上, RMSEA が .06 以下) を概ね満たしていることから, モデルにデータが適合していると判断した。

4.3.2 内的一貫性

このモデルにおける各因子の信頼性係数とその 95% 信頼区間は, 「クオリティ」因子: $\alpha = .91$ [.88, .93], 「悲劇」因子: $\alpha = .88$ [.85, .91], 「感動」因子: $\alpha = .81$ [.75, .86], 「ユーモア」因子: $\alpha = .76$ [.69, .83], 「迫力」因子: $\alpha = .81$ [.76, .86] であった。研究 2 と同様, 「ユーモア」因子のみやや低かったものの, 十分な内的一貫性が確認された。

4.3.3 作品 A と作品 B の各尺度得点の比較

作品 A と作品 B における各尺度得点の平均値の差を対応のない t 検定によって分析した (表 4)。その結果から, 「クオリティ」因子を除く 4 因子で有意に差があることが示された。このことから, 「悲劇」因子, 「感動」因子, 「迫力」因子において作品 A の方が作品 B よりも有意に尺度得点が高く, 「ユーモア」因子におい

¹TPES の項目の詳細ならびに因子負荷量等の情報については著者にお問い合わせください。

表 4 作品 A と作品 B における演劇評価尺度の各尺度得点の平均値の差の分析

演劇評価尺度	<i>M (SD)</i>		対応のない <i>t</i> 検定		
	作品 A (<i>n</i> = 49)	作品 B (<i>n</i> = 28)	<i>t</i>	<i>p</i>	<i>d</i>
一般因子	4.77 (0.31)	3.65 (0.41)	12.42	< .001***	3.19
クオリティ	4.94 (1.01)	5.04 (0.91)	0.46	.647	0.11
悲劇	5.13 (0.86)	2.58 (1.02)	11.24	< .001***	2.79
感動	3.69 (0.76)	2.92 (1.02)	3.51	.001***	0.90
ユーモア	3.22 (1.13)	5.00 (0.84)	7.83	< .001***	1.71
迫力	5.56 (0.77)	5.04 (0.73)	2.97	.004**	0.69

表 5 作品 C における演劇評価尺度の各尺度得点の平均値の差の分析

演劇評価尺度	<i>M (SD)</i>			分散分析		
	C1	C2	C3	<i>F</i>	<i>p</i>	η_p^2
一般因子	3.85 (0.41)	4.57 (0.83)	4.00 (0.43)	32.23	< .001***	.18
クオリティ	5.09 (0.79)	5.39 (0.86)	5.92 (0.77)	8.82	< .001***	.33
悲劇	2.97 (0.91)	3.49 (0.76)	5.68 (0.61)	59.38	< .001***	.77
感動	3.40 (0.95)	4.33 (1.06)	5.15 (1.00)	19.24	< .001***	.63
ユーモア	4.78 (0.71)	4.78 (0.88)	3.26 (0.77)	30.05	< .001***	.63
迫力	5.11 (0.94)	5.16 (0.91)	6.20 (0.59)	16.01	< .001***	.47

て作品 B の方が作品 A よりも有意に尺度得点が高いことが示された。

4.3.4 作品 C の時系列推移：各尺度得点の比較

作品 C の序盤、中盤、終盤の印象の推移について、C1, C2, C3 における演劇評価尺度の各因子の尺度得点を要因とする一要因参加者内分散分析を行った結果を表 5 に示す。その結果、演劇評価尺度の全ての因子において場面の主効果が有意であった。Bonferroni の補正を用いた事後比較から、「悲劇」因子と「感動」因子においては C1, C2, C3 の順で有意に尺度得点が高く、「クオリティ」因子と「迫力」因子においては、C3 が C1, C2 よりも有意に尺度得点が高いことが示された。「ユーモア」因子においては C1, C2 の方が C3 よりも有意に尺度得点が高いことが示された。この結果は、この作品の序盤ではユーモア性が強調されており、後半に行くにつれそれが減衰するのに対して、悲劇性、迫力、感動、クオリティといった要素は演劇のクライマックスに向けて高まっていることを示している。

4.4 考察

因子分析の結果から、双因子モデルによって調査 2 の探索的因子分析の因子構造が再現された。また Cronbach の α 係数によって各因子の内的一貫性が示され、モデルがデータに適合したことが示された。

作品 A は、終盤には主人公の死が扱われる、暗く、重々しい内容で構成され、その一方で、作品 B はヤクザ映画のチンピラエキストラに集まる人気のない役者をテーマとした、鑑賞者に笑いを誘うようなコメディ性の強い作品であった。作品 A, B の各尺度得点の平均値の差に関して、「悲劇」因子と「ユーモア」因子で特に効果量が大きかった。このことは作品 A が作品 B よりも悲しい印象を与える作品であり、また同時に作品 B は作品 A よりもかなり可笑しい印象を与える作品であることを意味する。この結果は、作品に意図される定性的な内容と良い対応関係を示していると考えられる。このことから、演劇評価尺度は、少なくとも、短編演劇の悲劇性やユーモア性を弁別する尺度として使用可能であることが示されたといえよう。

同様に、作品 C は、童話を再現する場面などを交えながら明るい雰囲気で行進する序盤・中盤と登場人物が生々しく殺されていく終盤との対比が特徴的な作品である。物語の導入部分である C1 では、主人公が童話を無邪気に、毎週近所の子供たちに読み聞かせると

いう日常が描かれる。次の C2 は、主人公が恋に落ちたことで素敵な作品を書き、それを著名なコンテストに応募することを決意する場面である。最後の C3 では、その恋が終わり絶縁され、さらに親友に裏切られたことに絶望し、友人や親友が殺されていく場面である。同一作品であるにもかかわらず、演劇評価尺度のすべての下位尺度において場面による違いが認められた。このことは、C1, C2 から C3 で雰囲気が一変するという定性的な印象と合致しており、作品の印象の変化を定量的に評価することができた証左であるといえる。

5. 総合考察

本研究では、演劇作品の印象を評価する尺度を定量的に構築した。演劇評価尺度を構成する評価語の中で、観客の情動体験を表す評価語は少なく、Konjin (1999) の感情語はいずれも最終尺度に残らなかった。また、例えば「悲しい」という形容詞が、観客自身の感情を表すのか、作品の印象を表す語彙なのかは明確ではない。演劇の本質は、Gouhier (1943) が述べるように、観客と俳優が同じ時間と空間を共有する現前性にあり、目の前で起こる出来事に対してのみ演劇における情動体験は喚起されると考えられる。これを踏まえると、本研究で構築された尺度が、舞台上の演劇作品に適用可能かどうかはさらなる追究が必要である。

また演劇評価尺度の各評価語が、演者の演技への印象を表すのか、それとも作品のストーリー性への印象を表すのかも明らかではない。演技、ストーリー性への評価を表す因子が抽出されなかったことから、作品の印象を評価する上でこの2項を切り離して考えることは難しいと思われる。Konjin (1999) は調査の結果から、観客の感情を最も多く喚起する対象は物語 (in fiction) であるのに対し、観客の感情を最も強く喚起する対象は演者の演技 (actor's performance) であるとしている。本研究の結果は Konjin の研究結果の一面を示し、演劇の印象は演技やストーリー性をはじめ、舞台照明・音響、演出など無数の要素によって形成されると考えられる。

本研究では、印象評定実験において3つの作品を映像刺激とし、全ての作品の印象を定量的に測定することができたが、演劇作品は作品によって物語や演出が千差万別であるため、より多くの作品の印象評価に演劇評価尺度を用いることによって、その妥当性と汎用性を検証しなければならない。また映像刺激にはない演劇の現前性が、観客の情動体験だけでなく、印象評価にも影響を及ぼすことは十分に考えられる。このこ

とを踏まえて、劇場の演劇作品の印象評価における演劇評価尺度の妥当性を検証するとともに、映像での演劇と劇場での演劇との印象の違いを定量的に示すことができれば、本研究はより学術的に有意義な研究となるであろう。今後、さらなる調査によって尺度の検討、改良を行うことで、将来的には演劇作品を検索するサービスの検索精度の向上や、演劇創作に携わる演出や演者のパフォーマンスの向上、現状ほとんど行われていない演劇の実証的研究の発展などの一助となることが期待される。

謝辞

本研究は科研費基盤研究 (B) 20H04485 の助成を受けて実施されました。

文献

- [1] 後安 美紀 (2006). 演劇と同時的対話 佐々木 正人 (編) アート／表現する身体——アフォーダンスの現場—— (pp. 26–54) 東京大学出版会
- [2] Gouhier, H. G. (1943). *L'essence du théâtre: Précédé de quatre témoignages*. Plon
- [3] Henrique, J. (2018). Get Old Tweets-python [Computer program]. Retrieved from <https://github.com/Jefferson-Henrique>
- [4] 樋口 耕一 (2014). 社会調査のための計量テキスト分析——内容分析の継承と発展を目指して—— ナカニシヤ出版
- [5] Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- [6] ヒュース 由美・岡田 猛 (2016). 即興演劇と台本演劇における観客による印象の違い 日本認知科学会第33回大会発表論文集, 137–139.
- [7] 金礪 愛・中野 倫靖・後藤 真孝・菊池 英明 (2016). 歌声の印象評価尺度の構築に基づく多様な印象の自動推定手法 情報処理学会論文誌, 57(5), 1375 – 1388.
- [8] Konijn, E. A. (1999). Spotlight on spectators: Emotions in the theater. *Discourse Processes*, 28(2), 169–194.
- [9] 野村 亮太 (2010). あなたにはこの芸人がおススメです——芸人评价/位置づけ尺度 (GELOS) の構成と信頼性・妥当性の検討—— 笑いの科学, 2, 57–62.
- [10] Tan, E. S. H. (1996). *Emotion and the narrative structure of film*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [11] West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). The Guilford Press.