

## 連続的測定方法による音と映像の印象の一致に基づく調和感に関する研究

藤山, 沙紀

<https://doi.org/10.15017/1398380>

---

出版情報：九州大学, 2013, 博士（芸術工学）, 課程博士  
バージョン：  
権利関係：全文ファイル公表済

連続的測定方法による音と映像の印象  
の  
一致に基づく調和感に関する研究

**Study on perceived congruence based on  
similarity of affective impressions between sound and  
motion picture using continuous measurement**

2013 年 8 月

藤山 沙紀

Saki FUJIYAMA

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 はじめに.....	1
1.2 視覚と聴覚の相互作用に関する基礎的研究.....	2
1.3 映画における音（あるいは音楽）に関する評論.....	3
1.3.1 diegetic sound（物語世界の音）と non-diegetic sound（非物語世界の音）.....	4
1.3.2 音（あるいは音楽）と映像のさまざまな関係.....	4
1.4 音楽と映像の相互作用に関する心理学的研究.....	5
1.5 音（あるいは音楽）と映像の調和.....	7
1.5.1 音と映像の構造的調和.....	7
1.5.2 音楽と映像の意味的調和.....	8
1.5.3 映像メディア作者の意図.....	9
1.6 連続的な測定方法を用いた実験心理学的研究.....	9
1.7 本研究の目的と位置付け.....	11
1.8 本論文の構成.....	12
<b>第2章 音と映像の調和感の形成に及ぼす要因 —単純な刺激を用いた場合—</b>	<b>14</b>
2.1 はじめに.....	14
2.2 予備実験：連続測定におけるマウス操作による 遅延時間の測定.....	14
2.2.1 実験方法.....	15
2.2.1.1 実験参加者	
2.2.1.2 実験刺激	

2.2.1.3	実験装置	
2.2.1.4	実験手続き	
2.2.2	結果.....	16
2.3	実験1：音と映像のタイミングが同期した刺激に おける音と映像の調和感の連続測定.....	16
2.3.1	実験方法.....	16
2.3.1.1	実験参加者	
2.3.1.2	実験刺激	
2.3.1.3	実験装置	
2.3.1.4	実験手続き	
2.3.2	結果と考察.....	18
2.4	実験2：音と映像の印象が類似した刺激における 音と映像の調和感の連続測定.....	22
2.4.1	刺激選定のための予備実験.....	22
2.4.1.1	実験方法	
2.4.1.1.1	実験参加者	
2.4.1.1.2	実験刺激	
2.4.1.1.3	実験装置	
2.4.1.1.4	実験手続き	
2.4.1.2	結果	
2.4.2	音と映像の印象が類似した刺激における 音と映像の調和感の連続測定.....	25
2.4.2.1	実験方法	
2.4.2.1.1	実験刺激	
2.4.2.1.2	実験装置	
2.4.2.1.3	実験手続き	
2.4.2.1.4	実験参加者	
2.4.2.2	結果と考察	

2.4.3	構造的調和と意味的調和の形成過程の比較.....	31
2.5	本章のまとめ.....	33

### 第3章 市販の映像作品を用いた音楽と映像の 調和感の連続測定 35

3.1	はじめに.....	35
3.2	実験3：市販の作品を用いた視聴覚刺激の 音楽と映像の調和感の連続測定.....	35
3.2.1	実験方法.....	36
3.2.1.1	実験参加者	
3.2.1.2	実験刺激	
3.2.1.3	実験装置	
3.2.1.4	実験手続き	
3.2.2	結果と考察.....	36
3.3	実験4：市販の作品を用いた視聴覚刺激の各時点での 音楽と映像の印象の連続記述選択実験.....	39
3.3.1	実験方法.....	39
3.3.1.1	実験参加者	
3.3.1.2	実験刺激	
3.3.1.3	実験装置	
3.3.1.4	実験手続き	
3.3.2	結果と考察.....	41
3.3.3	音楽と映像の調和度と音楽と映像の 印象の一致度の関係.....	46
3.4	実験5：意味的調和の形成過程に関する実験.....	48
3.4.1	実験方法.....	48
3.4.1.1	実験参加者	

3.4.1.2	実験刺激	
3.4.1.3	実験装置	
3.4.1.4	実験手続き	
3.4.2	結果と考察.....	49
3.5	本章のまとめ.....	51

## 第4章 黒澤明の作品における音と画を対比させる

	手法の効果	<b>52</b>
4.1	はじめに.....	52
4.2	実験6：視聴覚刺激の音楽と映像の調和感の連続測定.....	53
4.2.1	実験方法.....	53
4.2.1.1	実験刺激	
4.2.1.2	実験参加者	
4.2.1.3	実験装置	
4.2.1.4	実験手続き	
4.2.2	結果と考察.....	55
4.3	実験7：視聴覚刺激の音楽と映像の印象の連続記述選択実験....	57
4.3.1	実験方法.....	57
4.3.1.1	実験刺激	
4.3.1.2	実験参加者	
4.3.1.3	実験装置	
4.3.1.4	実験手続き	
4.3.2	結果と考察.....	57
4.4	実験8：視聴覚刺激全体の印象評定.....	71
4.4.1	実験方法.....	71
4.4.1.1	実験刺激	
4.4.1.2	実験参加者	

4.4.1.3	実験装置	
4.4.1.4	実験手続き	
4.4.2	結果と考察.....	72
4.5	実験9：音楽の音源が画面に映っていない	
	視聴覚刺激全体の印象評定.....	75
4.5.1	実験方法.....	76
4.5.1.1	実験刺激	
4.5.1.2	実験装置	
4.5.1.3	実験手続き	
4.5.1.4	実験参加者	
4.5.2	結果と考察.....	76
4.6	本章のまとめ.....	79
<b>第5章</b>	<b>全体的考察</b>	<b>80</b>
5.1	本研究の概要.....	80
5.2	音と映像の調和感の形成過程のモデル化.....	82
5.3	視聴覚融合過程のモデル化.....	89
<b>第6章</b>	<b>結論</b>	<b>91</b>
	参考文献	<b>93</b>
	謝辞	<b>99</b>

# 第1章 序論

## 1.1 はじめに

映画やテレビ番組のような映像作品は、「映像」と「音」で構成されている。映像作品では映像が重視されがちであるが、音を出さずにテレビドラマを見てもつまらないように、音は映像作品において重要な役割を果たしている。映像表現における音の役割は多岐にわたるが、通常「脇役」扱いされている。しかし、主役の「映像」が引き立つのは、脇役の「音」がうまく機能してこそである。

映像に加えられるのは、役者のせりふや足音のように映像に表現された対象から発せられる音（diegetic sound：物語世界の音）だけではない。映像で表現された世界には存在しない特殊な効果音や音楽（non-diegetic sound：非物語世界の音）が、映像の効果を高めるために用いられている。1927年にトーキー映画が公開された当初は、登場人物の話し声があれば、音楽は必要ないと考えられていた。しかし、音楽なしでは何かが欠けていると感じられたそうである（Kracauer, 1960）。効果音や音楽は、映像作品にはなくてはならないものなのである。

映画やテレビドラマなどにおいて効果音や音楽は、場面を強調したり、登場人物の気持ちを表したり、場面のムードを伝えたりと、各種の演出効果を担っている。特に、音楽は人物の心情や場面の状況を伝えるのに効果的である。恋愛映画などでの恋人同士の場面では、ロマンティックな音楽を流すことでムードを盛り上げることができる。ホラー映画では不気味な音を強調することで恐怖感を増大させることもできる。

効果音や音楽は、映像に表現された対象から発せられた音ではないために映像作品に自由に加えることが可能である。ただし、どんな効果音やどんな音楽でも映像に付加すれば効果的かという点、そうではない。うまく組み合わせた効果音や音楽は作品の質を高めるが、組み合わせに失敗すると映像作品が台無しになる。作品としての質を高めるためには、音と映像の調和をうまく図らなければならない。

映像作品においては音と映像の調和を図るのに、制作者の経験や勘、感性によるところが大きい。しかし、そこには共通する心理的な法則も存在し、科学的な手法で音と映像の調和を図る手法に迫ることも可能だと考えられ、実験心理学的なアプローチが開始



された。かつては、映像作品における音や音楽の効果に関しては、それほど多くの研究はなかったが、認知心理学の発展や、デジタル技術の発達などによって音や映像の加工が比較的容易に行われるようになったことにより、多くの研究がなされるようになってきた。

本研究も、映像作品における音の役割に科学的な手法で迫ることを目指して行ったものである。本研究では、音と映像の印象を一致させることによって得られる意味的調和と呼ばれる調和感に焦点をあて、その形成過程とその効果について探る。

本章では、音と映像の相互作用の基礎としての視聴覚融合に関する基礎的知見、映像作品における音、音楽に関する評論的検討、音と映像の相互作用に関する研究等過去の知見を概観するとともに、本研究で用いる連続測定法などの有効性を示し、最後に本研究の目的と特徴およびその立場について述べる。本章で概観する過去の知見において、音楽が用いられた場合は「音楽」、効果音や単音のように音楽と呼べないような音が用いられた場合は「音」と称する。

## 1.2 視覚と聴覚の相互作用に関する基礎的研究

人間の感覚は、視覚、聴覚、味覚、嗅覚などといった異なる種類に分かれている。この感覚の種類は、様相 (modality) と呼ばれている (丸山, 1969)。また、視覚には光、聴覚には音波というように、どの感覚にも受容される妥当な刺激 (適刺激, adequate stimulus) が定まっている。諸感覚はそれぞれが特徴的な役割を果たしており、視覚は空間情報の処理に、聴覚は時間情報の処理によく特性を発揮すると言われている (和田, 1967)。基礎的な心理学研究の分野において、ある 1 つの感覚の機能だけでなく、感覚間関係の解明の重要性が主張されており (Ryan, 1940 ; 北村, 1967)、視聴覚間のさまざまな相互作用の研究が古くから行われている。

感覚のある種の心理質が様相間に共通して認められる現象は、通様相性現象 (intermodality phenomenon) と呼ばれている (丸山, 1969)。たとえば、「明るい」「暗い」「澄んだ」「濁った」のように形容詞で表現できる印象は、視覚だけではなく聴覚にも共通するものが多い。「明るい」印象の色に「明るい」印象の音を加わると、色の明るさが増すように感じられるといった現象は、共鳴 (consonance) と呼ばれている (Ryan, 1940)。共鳴現象とは、視覚と聴覚の共通する心理質が、同方向に変わることを意味する。また、複数の感覚が単なる寄せ集めとしてではなく、独自の経験を生じるといった「協合現象」の存在も指摘されている (丸山, 1969)。「協合現象」の例として「マガーク効果」があげられる (MacGurk and MacDonald, 1976)。「マガーク効果」とは、「バ

という声に合わせて「ガ」と発音をしている人の映像を見ると「ダ」と聴こえるというものである。

### 1.3 映画における音（あるいは音楽）に関する評論

映像メディアの始まりは映画であり、映画の中の「音」に関しては、「音楽」が最も多く議論されている。初めて映画が公開されたのは 1895 年である。この公開の年から約 30 年間の映画は、無声（サイレント）映画であった。この間、映写機からの雑音を遮蔽するために、また、映画の中の動きを説明するために音楽が利用されていた (Palmer, 1980)。その後、1927 年に初めてのトーキー映画「ジャズシンガー」が公開された。トーキー映画が公開された当初は、登場人物の話し声と効果音があれば、音楽が映画のムードや感情内容を作り上げる必要はないと考えられた。しかし、音楽のない、登場人物の話し声だけのトーキー映画の寿命は短く、音楽はすぐに再び映画に用いられるようになった (Kalinak, 1992)。多くの映画評論家や研究者が、映画における音楽が、映像と深く関わっていることを指摘している。

Gorbman (1987) は、映画における音楽は、感情の伝達を助けるもの（記号表現, signifiant) であると述べている。Gorbman は、さらに、映画における音楽には (a) 物語を区切り、エピソードをつなぐ役割 (b) 「暗示的意味を持つ区分点」を生みだし、映画の出来事を解釈し、強調し、際立たせ、指示する役割があることを指摘した。

岩宮 (2011) は、映像作品における映像と音楽の組み合わせの手法について言及している。このような手法の典型的なものとして、映像の動きに合わせて音楽のメロディ・ラインやリズム・パターンをつける「ミッキー・マウシング (Mickey-mousing)」があげられる。この手法は、Disney のアニメーションで多用されており、アニメの主人公の名前 (Mickey Mouse) を借りてこのように呼ばれている。また、登場人物の気持ちを代弁したり、場面の状況を物語るのに、音楽の醸し出すムードや情感を利用する手法もある。次に来る場面が予測できるような音楽を流すことで、予感を持たせながらストーリーを展開する音楽の予告的効果も、この手法を応用したものである。

中村 (2008) は、映像アートにおける映像と音楽の関係について、映像と音楽それぞれの表現イメージを互いに近づけようとする共同関係や、映像と音楽の相互に関係づけが存在しないかのように思わせる無関係などに分類した。さらに、音響素材と映像素材の様々な関係を通してそれらの持続を形成していくような組み合わせを、映像音響詩 (Audio-Visual Poem) と呼ばれるメディアアートの中で実現している。

### 1.3.1 diegetic sound (物語世界の音) と non-diegetic sound (非物語世界の音)

映画理論では、映画の物語の中の世界に属しているものごとを「diegesis」と呼ぶ (Gorbman, 1987)。映画の物語の中に存在する物から出された音は「diegetic sound (物語世界の音)」と呼ばれている。また、映画の物語の世界に存在しない音のことは「non-diegetic sound (非物語世界の音)」と呼ばれている (Percheron, 1980)。

Chion (1993) は、映像に対する音の基本的関係を「三つの輪 (tricycle)」として提案した。三つの輪の1つ目は、「イン」あるいは「同時 (synchrone)」の音と呼ばれ、画面中にその音の音源が見える音のことである。2つ目は、「フレーム外」の音である。これは、画面中にはその音の音源は見えないが、画面が示す空間に隣接する空間にあると想像される音を指す。3つ目は、「オフ」の音である。これは、背景音楽や過去の場面を物語るナレーションの声などのことであり、画面で示される場面とは、別の時間・空間にある音源の発する音を指す。Chion は、各場面がこれらの3つの領域を効果的に行き来することで、映画に時間的・空間的広がりを持たせることができると述べている。

Chion の述べる「イン」の音、「フレーム外」の音は「diegetic sound」, 「オフ」の音は「non-diegetic sound」に対応すると考えられる。制作者は「diegetic sound」と「non-diegetic sound」を組み合わせることで、映画の世界を作り上げている。

### 1.3.2 音 (あるいは音楽) と映像のさまざまな関係

Cook (1998) は、音楽、映像、ことばなど、異なるメディアからマルチメディアを構成するための一般的な理論を立てた。Cook は、音楽と映像のように異種のメディアが結びつく場合、(I) 一致 (conformance), (II) 相補 (complementation), (III) 競争 (contest) の3つの基本モデルがあると述べた。構成メディア間で印象が類似していると判断された場合は (I) 一致である。類似していないと判断された場合は、相違性の判断 (difference test) が行われる。その結果、構成メディア間の意味が矛盾していると判断されると (III) 競争となり、そうでない場合は (II) 相補となる。

Whitaker (1970/1983) は、映画における音楽の効果について、音楽と映像が協調的な関係にある場合と音楽と映像が対立的な関係 (対位法的関係) にある場合に分けて述べている。音楽と映像が対立的関係にある場合、音楽は映像のコントラストによって視覚の雰囲気強調する、または、映像との対立から新しい風刺的メッセージを作り上げ

る効果があると述べている。

Chion (2002) は、映画における音楽の分類として、感情移入音楽と非感情移入音楽という分類を提案した。感情移入音楽とは、音楽がその場面の、特に、登場人物が感じているとされる感情に関わっていると考えられる音楽である。非感情移入音楽とは、音楽が過酷な場面において何事もないかのように流れ、無関心を守り続けるような効果を持つ音楽のことである。例えば、残虐行為が陽気な手回しのオルガンの音のもとで行われる（見知らぬ乗客、ヒッチコック）場面が挙げられる。このように、音楽と映像が対立した関係にあると、無関心な世界の背景を見せる効果も生みだすことができる場合もある。

## 1.4 音楽と映像の相互作用に関する心理学的研究

Münsterberg (1916/2001) は、映画という新しい現象に最初に注目した心理学者である。彼は、映画における音楽は、緊張を解き放ち、興味を持続させ（注意を喚起し続け）、安らぎを与え、感情を強め、美的経験に貢献すると述べた。

Osgood et al. (1957) は、共感覚に関する研究を背景とし、対象の情緒的意味（印象）を測定するための方法を生みだした。ある言葉の持つ意味は、その辞書的な意味である「外延的意味」と表現的な意味である「内包的意味」に分類される。内包的意味はさらに、過去における経験から連想される名詞からなる「連想的意味」と、「明るい」や「美しい」のように言葉から連想される情緒的なイメージを表わす形容詞からなる「情緒的意味」に分類される。SD (semantic differential) 法では言葉の持つ「情緒的意味」を測定する（神宮，1992）。Osgood et al. は、種々の共感覚に関する研究の結果、対象の持つ印象を形容詞を用いて評定することの有意味性を多くの実験によって確認した。

この SD 法を用いて Tannenbaum (1956) は、映像作品における音楽の影響を心理学的手法により捉える最初の実験を行った。35 分のテレビドラマに音楽を加えた場合と、加えない場合の印象の相違を SD 法によって調べた。SD 法で用いた形容詞は、Osgood et al. (1957) の 3 つの普遍因子に属する形容詞対である。実験の結果、音楽は映像作品の力動性と活動性を強化する効果があることが示された。

その後、注意選択、記憶、推論といった人間の内部における情報処理過程を明らかにしようという認知心理学の分野の発展により、映像作品や映像メディアにおける音と映像の相互作用に関する研究に注目が集まるようになった。

Marshall and Cohen (1988) は、大小の三角形と小さな丸が運動する抽象的な図形を用い

たアニメーションに、ゆったりとした印象と力強い印象の2種類の音楽を組み合わせ、印象評定実験を行い、音楽が映像の印象に及ぼす影響を明らかにした。その結果、活動的な印象を与える音楽を付加すると映像の活動性も高く判断され、迫力のある音楽を付加すると映像の力動性も高く判断されることが示された。活動性と力動性の次元では音楽が映像の印象に対して直接的な影響を与え、音楽が映像の情緒的印象を大きく規定している。また、Marshall and Cohenは調和一連合モデルを提案し、音楽と図形の動きが調和することで、視覚情報と音楽との時間的な調和が見られる部分だけに注意が集中し、その部分と音楽の情緒的印象との間の連合が起こると結論付けた。

Boltz et al. (1991)は、記憶という側面における音楽と映像の相互作用の研究を行った。映画のクライマックスシーンに付加する音楽に対し、映像のムードに一致した音楽、一致しない音楽、音楽なしの条件を設定し、映像に対する記憶がどのように変化するかを調べた。その結果、音楽が映像のムードに一致した条件では一致していない条件と比べ、映像の内容を記述させる再生記憶テストの成績は良い結果が得られた。

Yamada (2008)は、ホラーゲームの映像と音楽を用いた実験で、ゲームの印象は映像よりも音楽で大きく規定されることを示した。金森ら(2012)は、ゲーム音楽とゲーム映像の印象に関する研究を行い、映像と調和した音楽を、同様の印象を与える音楽と入れ替えても、全体の印象と調和度は大きく変化しないことを明らかにした。

岩宮(1992, 1993)、Iwamiya(1994)は、視聴覚情報の統合に関わっていると考えられる諸要因を模式的に表わした。彼は、市販の映像作品の一部を用い、映像あるいは音楽のみを呈示した場合と、両者を一緒に呈示した場合のそれぞれの評定値から、音と映像がもたらす視覚と聴覚の統合過程における相互作用を検討した。

最も下位のレベルで生ずる相互作用は、一方の感覚に対する情報がもう一方の感覚の感度を変化させる現象である。映像の有無が再生音の音質に対する影響を検討することで、このような現象が観測された。フィルタを使って再生帯域を制限すると、再生音の音質は劣化し、貧弱な音になる。しかし、映像を同時に呈示したときには、聴覚系の音質の劣化に対する感度が鈍る。その結果、音質の劣化が分かりにくくなる。映像による視覚情報は、聴覚系の感度を低下させることが明らかになった。

もう少し上位の処理レベルでは、視覚と聴覚に共通して存在する通様相性を通して生ずる共鳴現象が生じていた。ただし、通様相性はその心理的性質に応じて処理レベルが異なり、処理レベルに応じて共鳴現象の生じ方にも差がみられた。「明るさ」という性質は、視覚においても聴覚においても、比較的低次の処理レベルで捕らえられる。この

場合には、視覚と聴覚の情報を統合する機構の活動とは関わりなく、共鳴現象が生ずる。

「明るい」音楽が映像の印象をより「明るく」する。もう少し上位の処理レベルによってもたらされる印象の場合には、共鳴現象に視聴覚情報の統合機構が介在してくる。適切な音楽と映像が組合わされたときのみ、共鳴現象が生ずる。「引き締まった」印象の音楽が、映像を「引き締まった」印象にする。

最上位の処理レベルでは、音楽と映像が一体のものとなって、音楽と映像の評価を高める協合現象も観測された。当然、この協合現象は、音楽と映像を組合せさえすれば生じるといった性質のものではない。制作者の意図のもとに組み合わされた音楽と映像の場合には、ほとんどの場合、協合現象が観測される。しかし、別々の作品の音楽と映像を組み合わせた場合には、このような現象は生じない。

## 1.5 音（あるいは音楽）と映像の調和

Bolivar et al. (1994) は、実験条件を統制することで、音楽と映像の構造的調和 (formal congruency) および意味的調和 (semantic congruency) の効果を初めて実証した。彼らは、友好的 (friendly)、または攻撃的 (aggressive) な映像と音楽を組み合わせ、友好的同士、攻撃的同士の映像と音楽の組み合わせにおいて被験者が調和感を感じることから、意味的調和の存在を示した。さらに彼らは、映像と音楽を組み合わせるときに時間的な関係をずらすことによって調和感が変化することから、時間的 (構造的) 調和の存在を示している。彼らは、音楽が特定の視覚対象物に注意を向けさせる力は、音楽と視覚情報の構造的調和だけでなく、意味的調和によっても生じると述べた。

### 1.5.1 音と映像の構造的調和

Lipscomb (1996) は、人の視覚と聴覚の知覚的統合における視覚領域と聴覚領域のアクセント構造の同期に関する一連の研究を行った。彼は、映画とアニメーションの映像と音楽のアクセントが完全に一致して起こる協和 (consonant)、アクセントは同じ周期で起こるがそのタイミングがずれている位相ずれ (out of phase)、完全に違う周期でアクセントが起こる不協和 (dissonant) の3つの条件を設定し、「同期した-同期していない」という尺度で評定実験を行った。その結果、協和条件で最も評定値が高く「同期した」という印象が強かった。評定値は、位相ずれ、不協和条件の順に低くなっていった。この実験により、音と映像のアクセント構造の同期が、視聴覚情報の同期感に大きく影響を与えることが示された。視聴覚情報の同期感とは、Marshall and Cohen (1988) や Bolivar

et al. (1994) が述べている音と映像の時間的な一致である構造的調和に基づく調和感であるといえる。Lipscomb (1996) により、音と映像のアクセント構造の同期が、構造的調和を成り立たせる要因となりうることが示された。

菅野, 岩宮 (2000) は、音と映像の同期要因および速度対応要因が視聴覚刺激の調和感に及ぼす影響について検討した。映像素材には、空間上に浮かぶ球体を眺める視点を移動させたコンピュータグラフィックスを用いている。音素材は、ベース・パートとドラムス・パートからなるシンプルなものであった。実験要因として、映像素材の動きの速さ (速い, 遅い), 映像素材のカットチェンジの生起頻度 (低頻度, 高頻度), 音素材のテンポ (108, 216 bpm), 映像素材のカットチェンジと音素材の強拍の組み合わせ方 (同期, 位相ずれ, 非同期) の 4 要因を設定した。実験の結果, 映像素材のカットチェンジと音素材の強拍の同期 (同期要因) と, 映像素材の動きの速さと音素材のテンポの対応 (速度対応要因) が調和感を形成する要因となることが示された。この 2 つの要因は, 独立に機能し, 相互作用も見られなかった。このことから, 同期要因と速度対応要因の効果が質的に異なっており, 両者は異なった心的過程を通して生じていることが示唆された。同期要因は構造的調和, 速度対応要因は意味的調和に対応すると解釈できる (岩宮, 2011)。

### 1.5.2 音楽と映像の意味的調和

岩宮, 林 (1999) は, 色彩と音楽の印象に関する印象評定実験を行った。映像刺激は, 「赤」「黄」「黄緑」「緑」「シアン」「青」「青紫」「紫」の照明条件下でピアノ映像風景を模擬した 8 種類の CG である。音刺激には, 8 種類のピアノ曲を用いている。各刺激の印象と, 音刺激と映像刺激の調和感の評定実験の結果, 長調でテンポの速い「明るい」印象の音楽は, 「緑」「黄」「黄緑」「シアン」と相性が良く, 「青」「青紫」「紫」「赤」とは合わなかった。また, テンポの遅い音楽は, 「青」「青紫」「紫」「シアン」と相性が良く, 「緑」「黄」「黄緑」「赤」とは合わなかった。彼らは, 一般に音楽が本来持つ特徴を助長する機能を持つ色彩が, その音楽と調和していると判断されたのだと考え, 得られた調和感は音と映像の意味的調和に基づくものであると解釈した。

岩宮ら (2002) は, 映像の速度と密度, 音楽の調性とテンポが映像作品の調和感に及ぼす影響について検討した。実験に用いた映像素材は, 格子状の仮想平面上に配置した人形が右から左へと一定の速度で移動するアニメーションである。映像の条件として「速い」「遅い」という速度条件と, 「密」「疎」という密度条件を設定している。音素材は, シンセブラス, ベース, ドラムスの 3 つのパートからなる, テンポや調性が変化

しても不自然にならない楽曲であった。音楽の条件としては、「長調」「短調」の調性の条件と、6段階のテンポを設定している。実験の結果、「速い映像と速いテンポ」、「遅い映像と遅いテンポ」、「速い映像と長調」、「遅い映像と短調」、「高密度の映像と速いテンポ」、「低密度の映像と遅いテンポ」の組み合わせが調和する傾向が示された。彼らは、音と映像の調和感に影響する視聴覚構成要素間の諸要因が、映像作品の印象に及ぼす影響を解明した。また彼らは、音楽と映像で類似した印象を生じさせる要因の組み合わせで調和感が高まることから、この実験で用いた視聴覚刺激において生じた調和感は、意味的調和に基づくものであると解釈した。

### 1.5.3 映像メディア制作者の意図

一般に、映像メディアの制作者は、音楽と映像を組み合わせる際に両者の間の主観的調和を図っていると考えられる。

Lipscomb and Kendall (1994)は、映画「Star Trek IV: The Voyage Home, 1986」の映像と音楽をオリジナルのものも含めてランダムに組み合わせた実験刺激を用いて、音楽と映像の調和感の評定実験を行った。その結果、制作者が意図した映像と音楽の組み合わせが最も調和することが示された。

Iwamiya (1994)は、市販の映像作品よりまとまりの良い部分の映像と音楽を用い、元の組み合わせであるオリジナル条件と別の映像と音楽を任意に組み合わせた条件を設け、音楽と映像の調和に関する心理実験を行った。その結果、Lipscomb and Kendallの研究と同様に、制作者が意図した音楽との組み合わせであるオリジナル条件が最も調和することが示された。

## 1.6 連続的な測定方法を用いた実験心理学的研究

本章で紹介した先行研究で使われた刺激は、主として数秒から数十秒の程度の短いものであり、呈示した視聴覚刺激全体に対して、評定を行うものであった。しかし、音楽など、日常に存在する音を対象とする研究では、より現実的な実験手法が必要とされる。連続的な測定方法はそのような実験手法の一つで、生態学的妥当性を高める上でも重要な研究方法であると言える（難波，桑野，2008）。本研究においても、音楽や映像の印象を検討するために連続的な測定方法を用いるので、本節では連続反応測定に関する試みについて紹介する。

Hevner (1936)は、音楽に対する連続的な反応を初めて検討した。実験では、音楽の印



象を表わす複数の形容詞をあらかじめ用意しておく。また、1つの音楽作品をいくつかの section にわけておく。聴取者は、それぞれの section の終わりに、section の印象を最もよく表わす形容詞を示すことが求められた。その結果、1つの音楽作品でも、section によって感じられる印象が異なることが明らかにされた。

1980年代ごろになると、デジタル技術の発達によってリアルタイムで反応を記憶媒体へ記録することが比較的容易にできるようになり、音楽などに対する実験参加者の連続的な反応測定実験が多く行われるようになった (Schubert, 2004)。

難波ら (1977) は、音の大きさ、うるささなど、1つの次元の印象の変化を連続的に捉える方法として、カテゴリ連続判断法を開発した。この方法では、例えば音の「大きさ」を7段階で判断させる場合、「非常に大きい」ならば「7」、「非常に小さい」ならば「1」というように、カテゴリに対応するコンピュータのキーを定めておく。実験参加者には、時々刻々の音の大きさを判断し、そのカテゴリに対応するキーを押すように求める (難波, 桑野, 2008)。この方法を用いて Kuwano and Namba (1985)は、20分間の自動車交通騒音に対する時々刻々の大きさの判断を行った。また彼らは、物理量である騒音レベルと被験者の反応との間のずれを、物理量と反応の相関係数を求めることにより検討した。カテゴリ連続判断法は、鉄道騒音や航空機騒音を対象とした実験でも用いられている (Weber, 1991)。

Gregory (1989) は、ある一次元の評定尺度を用いた連続的な反応を測定する、連続的な反応デジタルインターフェイス (Continuous Response Digital Interface, CRDI) を開発した。また、Schubert (1996) は、二次元 (2つの尺度) を同時に測定するインターフェイスを開発した。このインターフェイスは、独立した2次元の感情を表わす軸 (尺度) で構成された二次元空間を動き回ることにより、反応を示すものである。さらに、Negel et al. (2007) は、映像メディアによって喚起される感情を測定するためのソフト EMuJoy を開発した。EMuJoy は、映像メディアの呈示と連続的な反応の記録を同時に行うことができるソフトウェアである。

難波, 桑野 (1988) は、音楽の印象などのように、1つの次元だけではとらえられない、多岐にわたる印象を連続的に捉える連続記述選択法を提案した。この方法では、予備実験などであらかじめ選択した形容詞のリストを用意し、例えば、「明るい」ならば「a」のキーというように、各形容詞とコンピュータのキーとを対応付けておく。実験参加者は、時々刻々の音の音楽の印象を判断し、その印象に対応するキーを押すように求められる。

この方法を用いて、Namba et al. (1991) は、ムソグルスキー作曲の組曲「展覧会の絵」の「プロムナード」部分について数人の指揮者のオーケストラ演奏を用い、それらの演奏についての印象を捉える実験を行った。彼らは予備実験において、これらの演奏を聴かせ、その印象を表わすのにふさわしい形容詞を自由記述で記入することを求めた。彼らはさらに、予備実験の結果から選んだ 15 の形容詞を用いて、聴取者に連続記述選択法による各演奏の時々刻々の印象の判断を求める本実験を行った。また、聴取後の全体の印象も評定させた。アシュケナージ指揮のオーケストラ演奏に対する連続記述選択実験の結果、曲の始めと最後は「意気揚々とした」「堂々とした」「迫力のある」という印象が優勢であるのに対し、中央部では「落ち着いた」「穏やかな」という印象が優勢となっており、時々刻々と曲の印象が変化していることが示された。また彼らは、時々刻々の判断で各形容詞が選択された比率と、全体判断の結果を比較した。両者の結果はかなり良く一致しているが、一致しないものもあった。例えば、「華やかな」が全体判断では選択頻度が 1 位になっているにもかかわらず、時々刻々の判断では出てこないことが示された。このように、Namba et al. の実験の結果、部分判断では選択されなかった印象が、全体判断を特徴づけることがあり、全体は部分の単純な総和とは言えないことが示唆された。

## 1.7 本研究の目的と位置付け

本研究の目的は、意味的調和に基づく音（あるいは音楽）と映像の調和感が形成される過程を解明し、市販の作品における意味的調和の状況および、その効果を明らかにすることである。

本研究では、音（あるいは音楽）と映像の調和感の連続測定実験、音（あるいは音楽）と映像の印象の連続記述選択実験を実施した。このようなアプローチは、これまでの音（あるいは音楽）と映像の相互作用に関する研究ではなされていなかった。しかし、音楽や映像のように時間経過の中で作品が展開され、それに応じて心理的反応が喚起されるような素材には、これらの連続的に反応を測定する手法は有効であると考えられる。

また、映像作品においては、わざと意味的調和を崩し違和感を与えることによって独特の作品を制作しようとの意図をもった制作者もいる。これまでの音楽と映像の相互作用に関する研究では、このような作風の作品は対象とされなかったが、本研究ではこのような作品における意味的調和の状況に関しても研究対象とし、そこでの意味的調和の状況とその効果を探る。

本章で紹介したように、音（あるいは音楽）と映像の相互作用に関する研究は多く行われているが、映像作品の中での音楽と映像の印象や調和感の変化を対象とした研究、調和感を崩すことの効果を論じた研究はない。本研究により、これまでの研究よりもより市販の映像作品に近い状況での視聴覚融合の様子を捉えることが可能となり、人間が音楽と映像をどのように処理して映像メディア作品が成立するのかについてのより本質的な考察ができる。

本研究では、まず単純な音と映像の素材を用いて、音と映像の印象が類似し意味的調和が形成された視聴覚刺激、音と映像のアクセントが一致し構造的調和が形成された視聴覚刺激に対して、音と映像の調和感の連続測定実験を行い、意味的調和に基づく調和感の形成過程の特徴を明らかにする。

次に、Disney のアニメーションの一部を用いて、音楽と映像の調和感の連続測定実験、音楽と映像それぞれの印象の連続記述選択実験を実施し、市販の映像作品の中で意味的調和が形成されその後どのような展開をしているのかを検討する。同時に、オリジナルの作品とは異なる音楽を組み合わせた視聴覚刺激を用いて、同様の実験を行い、意味的調和に関しての制作者の意図を探る。

さらに、音楽と映像の印象を一致させない「音と画の対位法」と呼ばれる手法を用いた黒澤明監督の映画の一部を用いて、音楽と映像の調和感の連続測定実験、音楽と映像の印象の連続記述選択実験を実施し、作品の中での意味的調和の状況を明らかにする。同時に、刺激全体の印象評定実験を行い、意味的調和を崩すことが作品の評価にどのような効果を及ぼすのかを検討する。「音と画の対位法」では、その効果を自然に導入するために音源を視聴者に提示することが有効であると考えられている。本研究では、音源が画面中に存在することの効果についても、印象評定実験に基づいて検討する。

これらの一連の実験を通した結果を踏まえ、音（あるいは音楽）と映像の意味的調和の形成過程および作品中での意味的調和の状況とその効果についてモデル化し、総括する。本研究により、これまでの研究にない視点から、映像作品における音楽の効果を論じるための知見、作品を制作するための手法を示したいと考える。

## 1.8 本論文の構成

1 章（本章）では、映像作品における音楽と映像の相互作用に関する研究、音（あるいは音楽）と映像の調和感を形成する要因に関するこれまでの知見を示し、本研究の目的とその位置づけ、本論文の構成について述べた。

2章では、構造的調和によって生じる調和感と、意味的調和によって生じる調和感の形成過程の違いを検討するために行った連続測定実験について述べ、調和感の形成過程についての考察を行った。

3章では、Disneyのアニメーションを用いて行った、音楽と映像の調和感の推移状況を探る連続測定実験について述べ、市販の映像作品中での意味的調和の状況と、制作者の意図についての考察を行った。

4章では「音と画の対位法」が使われた黒澤明監督の映画を用いて行った、音楽と映像の調和感の連続測定実験と視聴覚刺激全体の印象評定実験について述べ、「音と画の対位法」の状況と効果について考察した。

5章では、2章から4章までで明らかにした意味的調和に関する知見を総括し、意味的調和の形成過程を指数関数でモデル化し、様々な条件下での調和感形成過程を比較した。さらに、視聴覚情報の融合過程をモデル化し、モデル上での意味的調和の過程について総合的に考察した。

6章では、結論として本研究を総括し、本研究のなした意義を述べた。

## 第2章

### 音と映像の調和感の形成に及ぼす要因

#### —単純な刺激を用いた場合—

##### 2.1 はじめに

視聴覚メディアにおける音と映像の調和感に関する研究によると、映像の中の動きと音のリズムや拍が一致した状態である「構造的調和」、また、映像と音の印象が一致した状態である「意味的調和」という要因が、音と映像の調和感を形成することが明らかになっている (Bolivar et al., 1994)。調和感が形成される過程は、意味的調和によるものと構造的調和によるものとは異なると考えられるが、それらの形成過程に関しては、十分な検討がなされていない。

本研究では、構造的調和によって生じる調和感と、意味的調和によって生じる調和感の形成過程の違いを検討するため、単純な視聴覚刺激を用いて音と映像の調和感の連続測定実験を行った。本章で用いた視聴覚刺激を構成する音は、リズム・パターン、和音といった単純なものであり、音楽とは言えないものなので「音」と称することとする。本論文では、実験によって音と映像の調和感が評定され数値化された値を「調和度」とする。調和度の値とその時間的な変化から、2つの調和における調和感の形成に必要な時間を求め、両者の形成過程を比較、検討した。

##### 2.2 予備実験：連続測定におけるマウス操作による遅延時間の測定

本論文で用いる連続測定とは、映像刺激または音刺激を視聴すると同時に、ディスプレイ上に呈示された尺度上のアイコンをマウスで動かすことで連続的に評定を行うものである。評定開始時、アイコンは尺度の中央に位置している。したがって、刺激を呈示してから、マウスを操作しアイコンをふさわしい位置に置くまでに、ある程度の遅延時間があると考えられる。そこで、連続測定を行うための予備実験として、遅延時間

を把握するための実験を行った。

## 2.2.1 実験方法

### 2.2.1.1 実験参加者

実験参加者は、九州大学の学生 5 名（男性 1 名、女性 4 名、平均年齢 25.0 歳）であった。

### 2.2.1.2 実験刺激

音刺激としてバスドラム1音、映像刺激として水色画像を用いた。映像刺激は、実験参加者がディスプレイの黒色画面から刺激画面への変化を明確に認識できるものがある必要がある。そこで、本節の実験ではコントラストや明度が黒色画面とは明らかに異なる水色画像を映像刺激として用いた。音刺激の呈示レベル（等価騒音レベル）は、約 60 dBとした。

### 2.2.1.3 実験装置

音刺激の作成には Cakewalk Music Creator 5 を使用した。映像刺激の作成には Adobe Premiere 6.0 を使用した。刺激と連続尺度の呈示には、パーソナルコンピュータ（Dell OPTIPLEX 745）を用いた。映像と連続尺度は 17 inch ディスプレイの画面上に呈示した。音刺激は、ドライバユニット（STAX SRM-313）およびヘッドホン（STAX SR-307）を介して、実験参加者の両耳に呈示した。刺激の呈示時刻、尺度上のアイコンの位置と時刻の記録には Microsoft office Excel 2007 のマクロ機能を使用した。

### 2.2.1.4 実験手続き

実験は九州大学大橋キャンパス心理実験室小ブースにて行った。図 2-1 に実験環境の図を示す。連続測定は、ディスプレイ上の映像刺激の下部に呈示した横長の連続尺度と尺度上のアイコンにより行なわれた。実験参加者は、マウスを握ってディスプレイを注視し、音刺激や映像刺激が呈示されたらマウスでアイコンを尺度の中央から右端に動かすよう指示された。

刺激の呈示および尺度上におけるアイコンの位置の取得には、Excel のマクロ機能を用いた。プログラムにより刺激を呈示し、その時刻を記録する。また、アイコンの位置が変わるたびに、その尺度上での座標と時刻を 10 ms 単位で記録した。刺激が呈示されてからアイコンが尺度の右端に達するまでの時間を「マウス操作による遅延時間」とし、これを求めた。

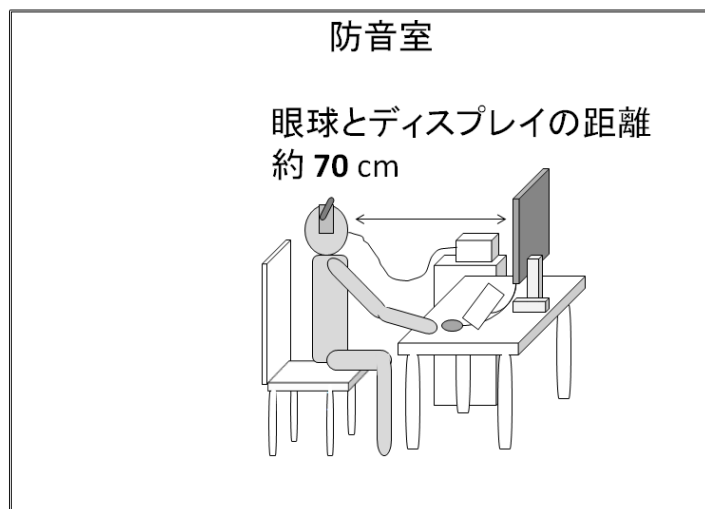


図 2-1 実験環境

## 2.2.2 結果

マウス操作による遅延時間の測定実験の結果、刺激の呈示から尺度の右端までアイコンを動かすのにかかった時間は、0.77～1.16 秒（平均 0.90 秒）であった。対応のある  $t$  検定を行ったところ、音刺激と映像刺激の間で遅延時間に差はなかった ( $t(4) = 1.35$ ,  $p > 0.1$ )。マウス操作による遅延時間が 1 秒程度なので、刺激呈示後 2 秒以降は十分に刺激の評定が行われていると考える。よって、本論文の連続測定によって得られた結果は、刺激開始後 2 秒以降のデータを考察の対象とする。

## 2.3 実験 1：音と映像のタイミングが同期した刺激における音と映像の調和感の連続測定

実験 1 では、音のリズムや映像の切り替わりのタイミングが一致した状態である構造的調和が成立した単純な刺激に対して、音と映像の調和感の連続測定を行った。得られた調和度から、構造的調和に基づく調和感が形成されるのに要する時間を検討した。

### 2.3.1 実験方法

#### 2.3.1.1 実験参加者

実験参加者は、九州大学の学生 15 名（男性 13 名、女性 2 名、平均年齢 23.2 歳）であった。

### 2.3.1.2 実験刺激

実験1で用いる視聴覚刺激は、構造的調和が成立することを意図して作成された。視聴覚刺激を構成する音素材は、ドラムス・パートによるリズムパターン（8ビート、120 bpm）である。映像素材は、4種類の単色の画像（ピンク色、水色、緑色、黄色）が順番に替わるものとした。これらを組み合わせて刺激を作成した。

音素材のドラムスの1拍目と3拍目に合わせて、映像素材の単色画像が切り替わるように組み合わせたものを「同期刺激」とした。また、意図的に単色画像が切り替わるタイミングとドラムスの拍が一致しないように組み合わせたものを、「非同期刺激」とした。図2-2に同期刺激、非同期刺激の模式図を示す。さらに、同期と非同期の部分が30秒毎に入れ替わるものを、「同期—非同期—同期刺激」「非同期—同期—非同期刺激」として、これらの4つを実験1の刺激とした。各刺激の長さはいずれも90秒間であった。呈示音圧レベルは約60 dB（等価騒音レベル）とした。

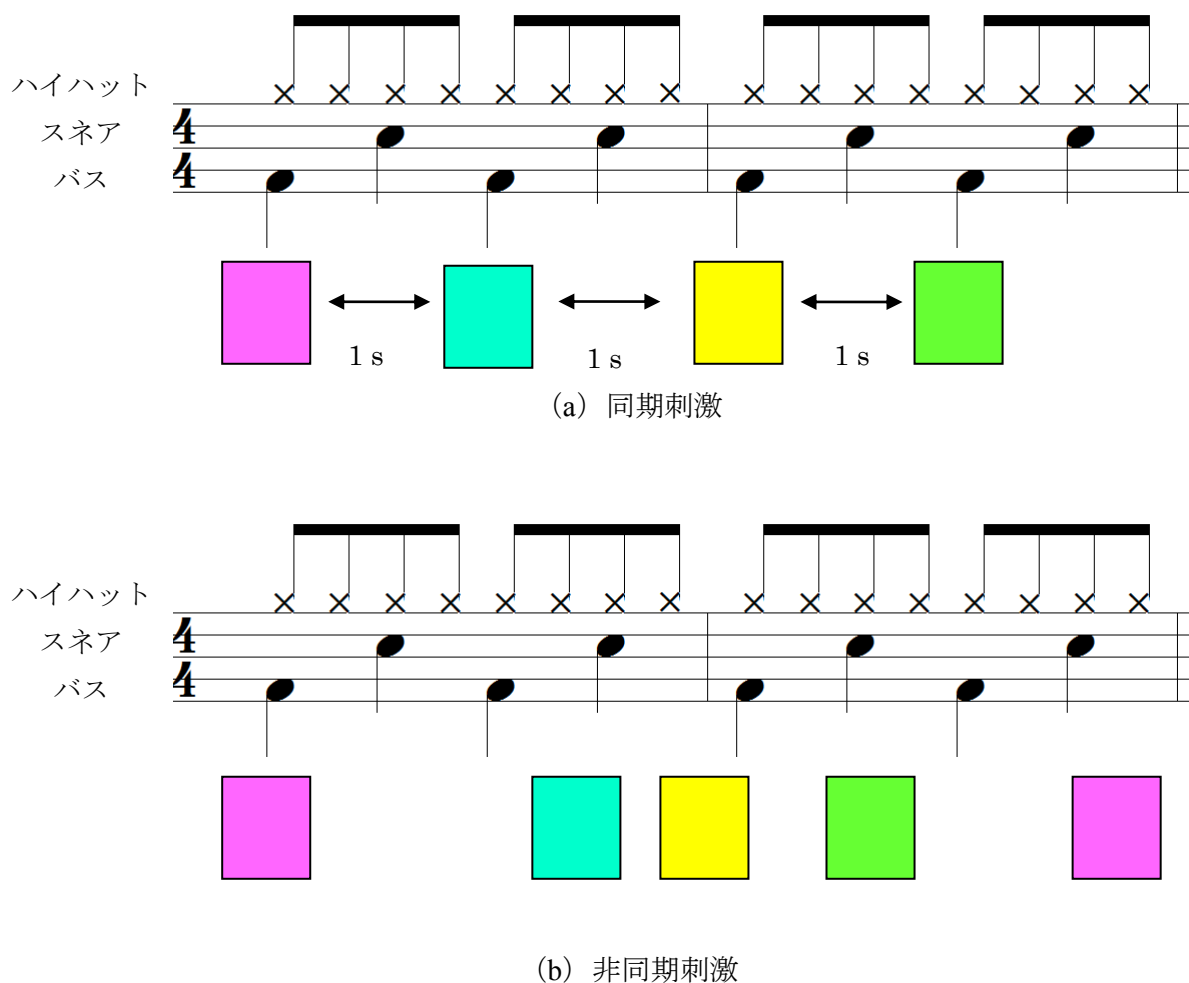


図 2-2 実験刺激



### 2.3.1.3 実験装置

音と映像の調和感の連続測定には主観評定の連続測定にふさわしいソフトウェアであるEMuJoy (Negel, et al., 2007) というソフトウェアを使用した。この実験は、オリジナルソフトウェア開発前に実施したので、このソフトウェアを用いた。このソフトウェアは、ディスプレイに表示された尺度上のアイコンをマウスで動かすと、その時刻と尺度の評定値がコンピュータに記録されるものである。その他の実験装置は、予備実験 (2.2.1) と同様である。

### 2.3.1.4 実験手続き

実験参加者には、視聴覚刺激を視聴しながら、両端が「調和している (+1) - 調和していない (-1)」である横長の連続尺度上で、マウスを左右に操作してアイコンを動かす、その瞬間に感じた音と映像の調和感を評定するように教示した。EMuJoyではアイコンの位置が変わるたびに、その尺度上での値と、その時の時刻が15 ms単位で記録される。

予備実験や2.4節の実験で用いたオリジナルソフトウェアで測定したデータと直接比較できるように、EMuJoyによって得られた15 ms毎の値を10 ms単位の値に変換した。EMuJoyにおいてデータがサンプリングされてから次にサンプリングされるまでの15 msの間は値は一定と考え、10 ms毎の値を抽出した。例えば、0 ms, 15 ms, 30 ms, 45 ms 時点での値が記録されたとすると、0 msから14.999... msの間は一定の値、15 msから29.999... msの間は一定の値、30 msから44.999... msの間は一定の値であるとするので、10 msの値は0 msの値と等しく、20 msの値は15 msの値と等しく、30 msと40 msの値は30 msの値と等しいとする。

## 2.3.2 結果と考察

図 2-3 に連続測定によって得られた、各刺激における音と映像の調和度の結果を示す (全実験参加者の値の平均値)。予備実験の結果より、刺激呈示後 2 秒以降の値を考察する。

実験参加者毎に、同期刺激、非同期刺激における刺激呈示後 2 秒から 90 秒までの調和度の平均を求めた。全実験参加者の同期刺激における平均調和度は 0.77 であり、非同期刺激における平均調和度は-0.46 であった。対応のある t 検定を行った結果、同期刺激の平均調和度と非同期刺激の平均調和度の差は有意であった。 $(t(14) = 8.43, p < .001)$ 。この結果から、映像と音のタイミングが同期した状態である構造的調和が成立している

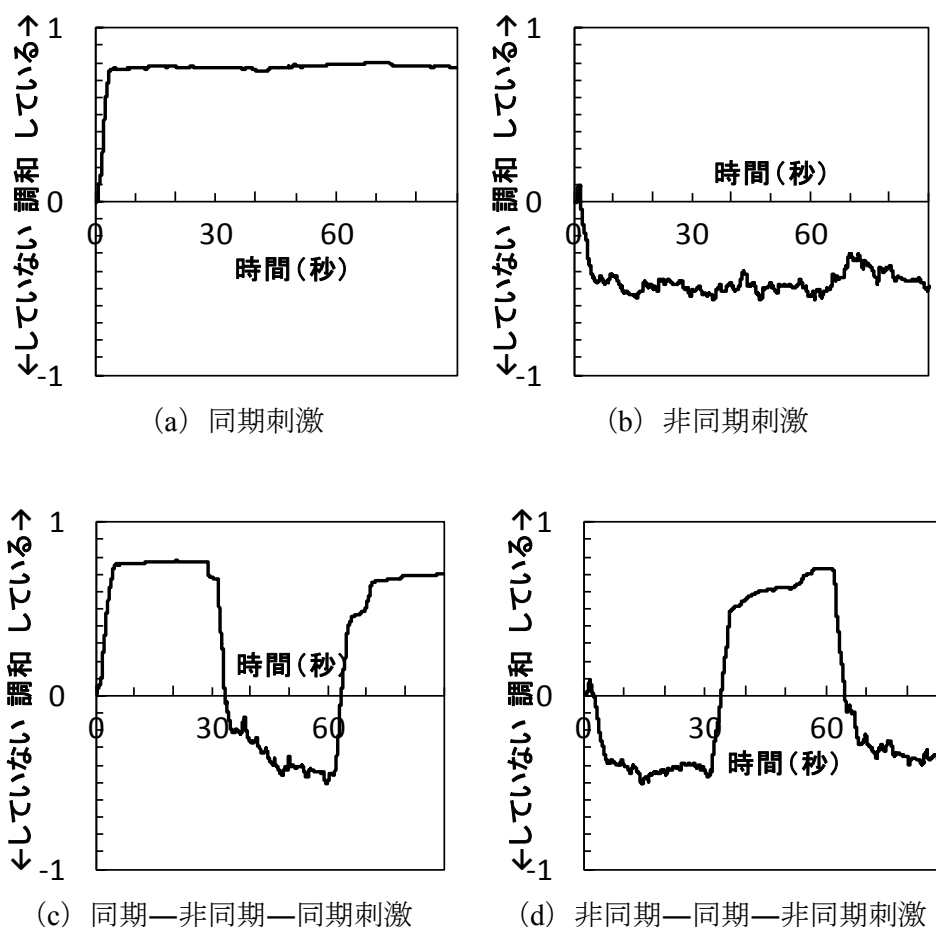


図 2-3 各刺激における音と映像の調和度（全実験参加者の値の平均値）

と、視聴者は音と映像に調和を感じる事が示された。この結果は視聴覚刺激を視聴した後に、調和感の評定を行った過去の研究結果とも一致している（菅野, 岩宮, 2000）。

また、同期-非同期-同期刺激における刺激呈示後 2 秒から 30 秒, 32 秒から 60 秒, 62 秒から 90 秒のそれぞれの調和度の平均を実験参加者毎に求めた。全実験参加者にわたり平均したところ, 平均調和度はそれぞれ, 0.74, -0.31, 0.56 であった。Bonferroni 法による多重比較を行ったところ, 刺激呈示後 32 秒から 60 秒の間の非同期部分と, 2 秒から 30 秒および 62 秒から 90 秒の同期部分との平均調和度に有意な差が見られた ( $p < .001$ ,  $p < .001$ )。

非同期-同期-非同期刺激でも同様の傾向が見られた。刺激呈示後 2 秒から 30 秒, 32 秒から 60 秒, 62 秒から 90 秒の全実験参加者の平均調和度を求めたところ, それぞれ -0.39, 0.53, -0.24 であった。Bonferroni 法による多重比較を行ったところ, 刺激呈示後

32 秒から 60 秒の同期の部分における平均調和度と、2 秒から 30 秒および 62 秒から 90 秒の非同期の部分における平均調和度との間に、有意な差が見られた ( $p<.001$ ,  $p<.001$ )。

これらのことから、途中で同期・非同期の状態が変化すると、それに伴い調和感も変化することがわかる。

また、同期刺激では、調和度が最も高い値に達した後の値は一定であり (0.75 ~ 0.78)、非同期刺激では調和度が -0.3 ~ -0.6 の間で変動するという現象が見られた。この現象は、視聴覚の情報処理過程において音と映像が非同期の場合にも、同期を感じようとしたために見られたものと考えられる。視聴覚の情報処理過程には、視覚と聴覚の時間的な同期性を見出そうとする傾向があるように思われる。

次に、同期刺激と同期-非同期-同期刺激における刺激呈示後 2 秒から 30 秒の間の調和度の推移から、構造的調和に基づく調和感の形成に要する時間を求めた。はじめに、28 秒間の調和度から 0.5 秒毎の調和度の変化率を求めた。また、刺激呈示後 2 秒における調和度と 28 秒における調和度の差を 28 で割ったものを 28 秒間の平均上昇値とし、これを求めた。調和度の 0.5 秒毎の変化率が平均上昇値よりも下回る点をピーク点、ピーク点に達するまでの時間をピーク到達時間と定めた。ピーク点が調和度の上昇が終わった点であり、ピーク到達時間を調和度の形成が完了した時間とする。このような方法で実験参加者毎、刺激毎にピーク到達時間を求めた。図 2-4 にピーク到達時間導出の模式図を示す。

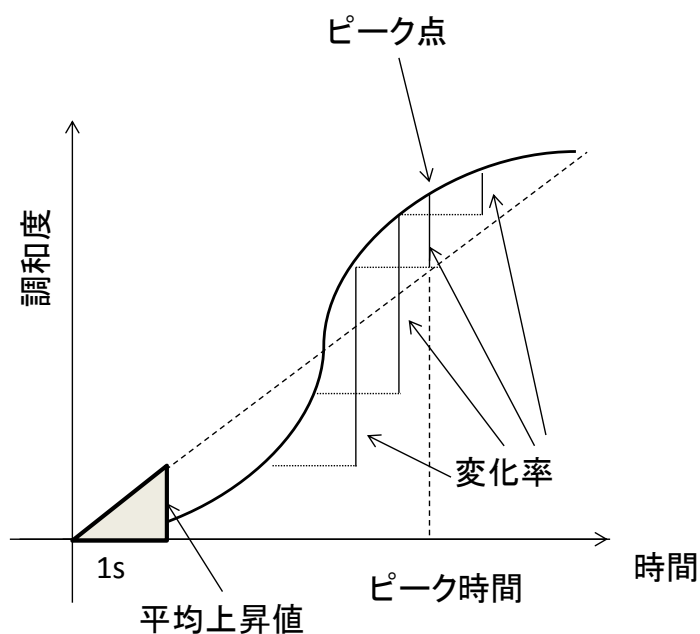


図 2-4 ピーク時間導出の模式図

表 2-1 各群に分類されるデータ

1回目	2回目以降
同期刺激（2秒から30秒）, 同期-非同期-同期刺激（2秒から30秒） のうち, 同期部分の評定が1回目の実験参加者 のデータ	同期刺激（2秒から30秒）, 同期-非同期-同期刺激（2秒から30秒） のうち, 同期部分の評定が2回目以降の実験参 加者のデータ

表 2-2 ピーク到達時間の平均値

		1回目	2回目以降	総和
ピーク到	データ数	9	16	25
達時間	平均(秒)	7.4	8.1	7.8
	標準偏差	3.8	4.5	4.16

ここで、同期部分の評定が2回目以降になると、1度評定を行っているために調和感の形成に要する時間が早まることも考えられる。そこで、同期部分が含まれている同期刺激、同期-非同期-同期刺激、非同期-同期-非同期刺激のいずれの評定も行っておらず同期部分の評定が1回目である実験参加者と、先にあげた3つの刺激のいずれかの評定を既に行い同期部分の評定が2回目以降である実験参加者へと分類する。表 2-1 に各群に分類されるデータを示す。

表 2-2 に、両群の平均ピーク到達時間を示す。平均を算出する際、調和度の値が小さい実験参加者5名（1回目；3名，2回目以降；2名）のデータは除外した。同期部分の評定が「1回目」である条件における平均ピーク到達時間は7.4秒であり、「2回目以降」の条件においては8.1秒であった。対応のないt検定を行ったところ、両者のピーク到達時間に有意な差は見られなかった（ $t(23) = -0.426$ ， $p > 0.1$ ）。全体の平均ピーク到達時間は7.8秒であった。

以上から、構造的調和が成立している視聴覚刺激に対して、音と映像の調和感が形成されるのに要する時間は、7秒から8秒程度であることが示唆された。また、測定が1回目か2回目以降かでピークに達する時間に差が無いことから、構造的調和の形成に要する時間には、刺激を視聴したことがあるかどうかの影響は無いことがわかる。

## 2.4 実験 2：音と映像の印象が類似した刺激における音と映像の調和感の連続測定

実験2では、音と映像の印象が類似した状態である意味的調和が成立した単純な刺激に対して、音と映像の調和感の連続測定を行った。得られた調和度から、意味的調和に基づく調和感が形成される時間を検討した。

### 2.4.1 刺激選定のための予備実験

意味的調和が成立した視聴覚刺激を作成するには、印象が類似した音と映像を組み合わせる必要がある。このような視聴覚刺激を作成するために、類似した印象を持つ音素材、映像素材を選定するための予備実験を行った。音素材、映像素材の「明るさ」の印象の連続測定実験を行い、明るい印象、暗い印象の音素材と映像素材を選定した。

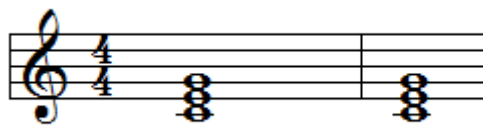
#### 2.4.1.1 実験方法

##### 2.4.1.1.1 実験参加者

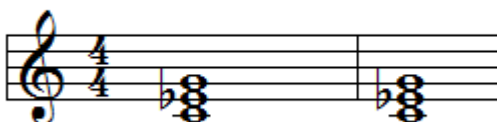
実験参加者は、九州大学の学生5名（男性2名、女性3名、平均年齢22.6歳）であった。

##### 2.4.1.1.2 実験刺激

音素材はピアノ音で約4.4秒毎に鳴らされる長三和音（ハ長調、長三和音、根音261.63 Hz、4/4拍子、全音符、55 bpm）と、短三和音（ハ短調、短三和音、根音261.63 Hz、4/4拍子、全音符、55 bpm）とした。それぞれを音素材A1、A2とする（図2-5）。各素材の長さは90秒であった。



(a) A1



(b) A2

図 2-5 刺激選定のための予備実験で用いた音素材

音楽的な文脈のない状況で和音を聴取し、その後知覚した感情を調べた研究によると、長和音は短和音よりも、「うれしい」という反応をした聴取者の数は、30人中29人であることが示されている (Crowder, 1984)。また、音楽を用いた研究でも、長調は「喜び」、短調は「悲しみ」と知覚されることが明らかにされている (Peretz et al., 1998)。従って、本研究で用いる長三和音・短三和音は、実験参加者に、「明るい」・「暗い」印象を生じさせることが予想される。

映像素材には6種類の映画より、笑顔映像3種類、泣き顔映像3種類を抜粋して使用した。いずれの映像も、正面から人物の顔が写され人物の表情がはっきりと認識でき、その場面が数秒間続くようなものである。顔面の感情表現を正しく識別できるかを検証した研究では、幸福や悲しみを表わす表情の写真を示した時に、その写真の人物とは異なる文化に生きる被験者でも幸福、悲しみと正しく認識できることが明らかになっている (Ekuman and Friesen, 1971)。したがって、本実験で用いる笑顔・泣き顔映像は、それぞれ明るい印象・暗い印象を実験参加者に生じさせると考えられる。各素材の長さは90秒とし、抜粋したままでは90秒の長さに足りない素材は、繰り返し再生した。映像素材のために抜粋した映画を以下に示す。

- ・ 西の魔女が死んだ (2008)より女の子の笑顔
- ・ 犬と私の10の約束 (2008)より女の子の笑顔
- ・ Nuovo Cinema Paradiso (1988)より男の子の笑顔
- ・ The Godfather Part III (1990)より女性の泣き顔
- ・ Simon Birch (1995)より男の子の泣き顔
- ・ Sad Movie (2005)より男性の泣き顔

#### 2.4.1.1.3 実験装置

実験装置は、本章予備実験 (2.2.1.3) で使用したものと同様であった。

#### 2.4.1.1.4 実験手続き

実験参加者には、音素材、映像素材を視聴しながら、両端が「明るい (1) - 暗い (-1)」である横長の連続尺度上でアイコンを左右に動かし、音素材、映像素材の明るさという印象に対して評定を行うように教示した。尺度の形容詞が左右どちらに配置されるかはランダムに入れ替えた。

アイコンを動かした時の座標と時刻の記録方法は、予備実験 (2.2.1.4) と同様である。アイコンの位置が変わるたびに記録された10 msごとの座標を、その時点での各素材の「明るさ」とする。

#### 2.4.1.2 結果

A1 (長三和音) の明るさの値は0.4程度、A2 (短三和音) の明るさの値は-0.6程度で一定値であり、それぞれ明るい印象、暗い印象をもつ音素材であると確認された。笑顔映像3種類の明るさの値はどれも0.6程度であり、明るい印象を持っていた。また、泣き顔映像3種類の明るさの値はどれも-0.6程度であり、暗い印象を持っていた。本実験では音と映像の調和感が形成される過程を検討するため、実験刺激に用いる映像は途中で印象の変化がないものがふさわしい。最も値に変化がなく安定していた映像は、西の魔女が死んだ (2008) から抜粋した女の子の笑顔の場面と、The Godfather Part III (1990) からの女性の泣き顔の場面であった。よって、この2つの映像を本実験の刺激用素材として選んだ。笑顔映像をV1、泣き顔映像をV2とする。図2-6に連続測定で得られた、A1, A2, V1, V2における明るさの評定結果を示す (全実験参加者の平均値)。

A1, A2, V1, V2の評定値のピーク点を前節 (2.3.2) と同様の方法で求めたところ、それぞれ5秒、11秒、8秒、8秒であった。これらの時間が、音・映像の印象が十分に形成される時間であると考えられる。

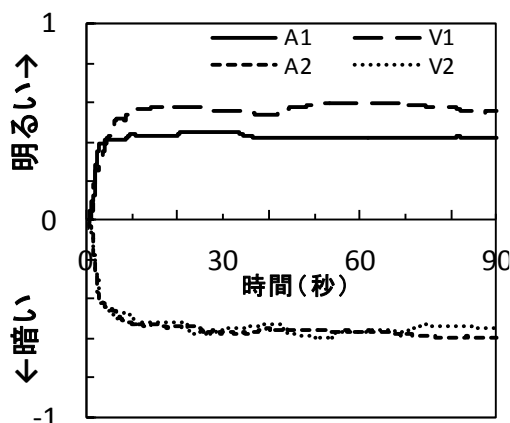


図 2-6 A1, A2, V1, V2 における明るさの評定値 (全実験参加者の平均値)

## 2.4.2 音と映像の印象が類似した刺激における音と映像の調和感の連続測定

### 2.4.2.1 実験方法

#### 2.4.2.1.1 実験刺激

予備実験により選ばれた素材を用いて、類似した印象の音素材と映像素材を組み合わせた「S1 刺激」「S2 刺激」、異なる印象の音素材と映像素材を組み合わせた「D1 刺激」「D2 刺激」、これらの刺激が 30 秒毎に入れ替わる「S1-D1-S1 刺激」「D1-S1-D1 刺激」「S2-D2-S2 刺激」「D2-S2-D2 刺激」の 8 つの視聴覚刺激を作成し、実験 2 の刺激とした。視聴覚刺激の詳細を以下に示す。

#### 長三和音群

**S1刺激** : 長三和音 (A1) + 笑顔映像 (V1)

**D1刺激** : 長三和音 (A1) + 泣き顔映像 (V2)

**S1-D1-S1刺激** : 長三和音 (A1) + 笑顔 (V1) → 長三和音 (A1) + 泣き顔 (V2)  
→ 長三和音 (A1) + 笑顔 (V1)

**D1-S1-D1刺激** : 長三和音 (A1) + 泣き顔 (V2) → 長三和音 (A1) + 笑顔 (V1)  
→ 長三和音 (A1) + 泣き顔 (V2)



### 短三和音群

**S2刺激**：短三和音 (A2) + 泣き顔映像 (V2)

**D2刺激**：短三和音 (A2) + 笑顔映像 (V1)

**S2-D2-S2刺激**：短三和音 (A2) + 泣き顔 (V2) → 短三和音 (A2) + 笑顔 (V1)  
→ 短三和音 (A2) + 泣き顔 (V2)

**D2-S2-D2刺激**：短三和音 (A2) + 笑顔 (V1) → 短三和音 (A2) + 泣き顔 (V2)  
→ 短三和音 (A2) + 笑顔 (V1)

#### 2.4.2.1.2 実験装置

実験装置は、本章予備実験 (2.2.2) で使用したものと同様であった。

#### 2.4.2.1.3 実験手続き

実験参加者には、視聴覚刺激を視聴しながら、両端が「調和している (1) - 調和していない (-1)」である横長の連続尺度上でアイコンを動かし、その瞬間に感じた音と映像の調和感を評定するように教示した。評定値の記録方法は予備実験 (2.2.1.4) と同様であった。

#### 2.4.2.1.4 実験参加者

実験参加者は、九州大学の学生 17 名 (男性 10 名, 女性 7 名, 平均年齢 22.8 歳) であった。

#### 2.4.2.2 結果と考察

図 2-7 に連続測定によって得られた、各刺激における音と映像の調和度の推移を示す (全実験参加者の値の平均値)。

実験参加者毎に、S1 刺激 (長三和音+笑顔映像), D1 刺激 (長三和音+泣き顔映像), S2 刺激 (短三和音+泣き顔映像), D2 刺激 (短三和音+笑顔映像) における、刺激呈示後 2 秒から 90 秒までの調和度の平均を求めた。全実験参加者にわたり平均した S1 刺激における平均調和度は 0.30 であり、D1 刺激における平均調和度は-0.35 であった。対応のある t 検定を行った結果、S1 刺激と D1 刺激の平均調和度の間に有意な差が見られた ( $t(16) = 6.05, p < .001$ )。また、S2 刺激における全実験参加者の平均調和度は 0.53 であり、D2 刺激における平均調和度は-0.29 であった。対応のある t 検定を行った結果、S2 刺激と D2 刺激の平均調和度の間に有意な差が見られた ( $t(16) = 6.70, p < .001$ )。

以上のように、明るい印象同士の音と映像の組み合わせである S1 刺激と、暗い印象同士の音と映像の組み合わせである S2 刺激の平均調和度は、音と映像の印象が異なるもの同士の組み合わせである D1 刺激, D2 刺激の平均調和度と比べ高い値であった。

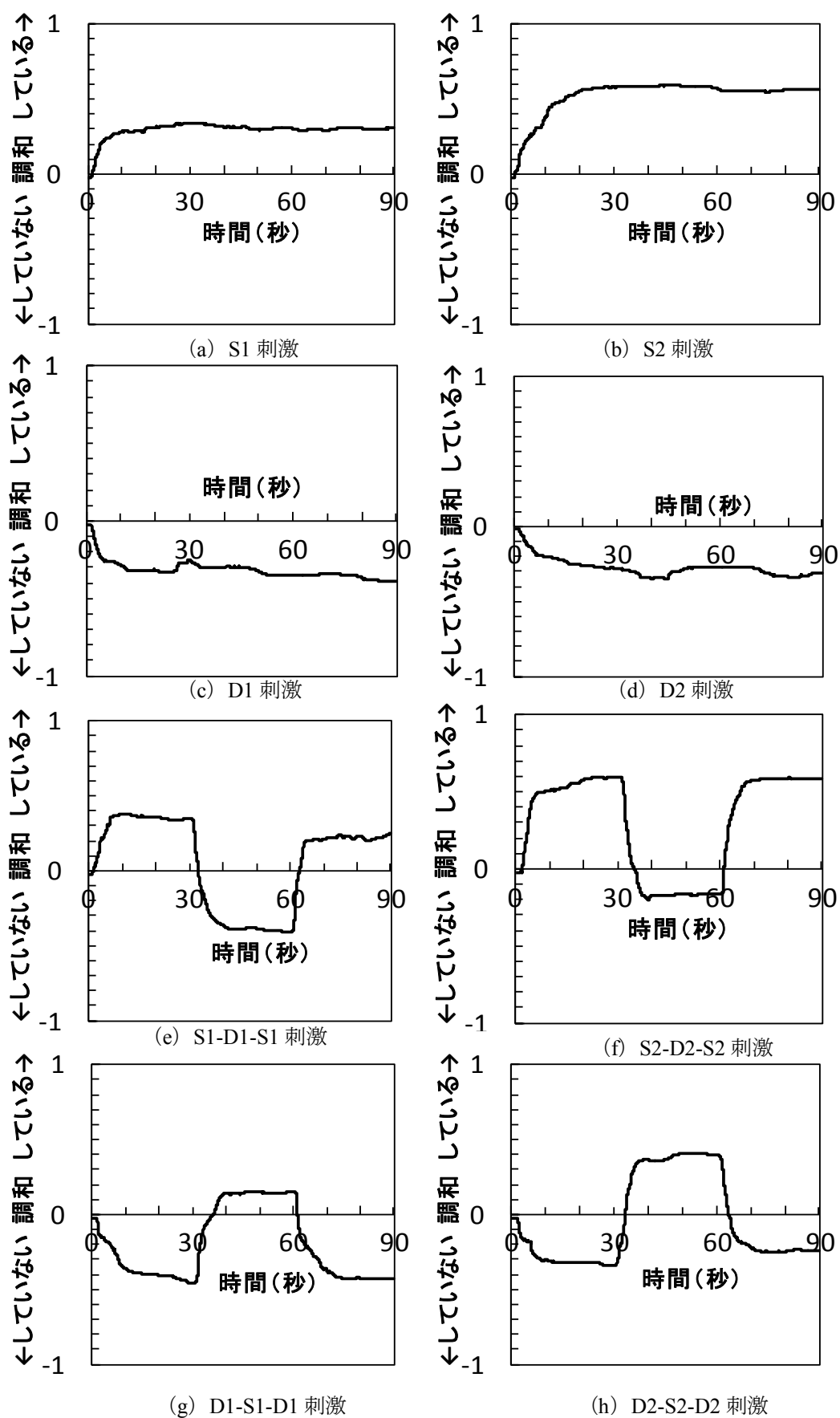


図 2-7 各視聴覚刺激における音と映像の調和度 (全実験参加者の値の平均値)

これらの結果から、映像と音の印象が類似した状態である意味的調和が成立していると、視聴者は調和感を感じることを示される。この結果は、視聴覚刺激を視聴した後に調和感の評定を行った過去の研究結果とも一致する（岩宮ら，2002）。

また、実験参加者毎に、S1-D1-S1 刺激における刺激呈示後 2 秒から 30 秒、32 秒から 60 秒、62 秒から 90 秒の間のそれぞれの調和度の平均を求めた。全実験参加者にわたり平均したところ、平均調和度はそれぞれ 0.33, -0.35, 0.20 であった。Bonferroni 法による多重比較を行ったところ、音と映像の印象が類似していない D1（刺激呈示後 32 秒から 60 秒）と、音と映像の印象が類似した部分である S1（刺激呈示後 2 秒から 30 秒、および 62 秒から 90 秒）の平均調和度の間には、有意な差が見られた ( $p<.001$ ,  $p<.01$ )。

D1-S1-D1 刺激でも同様の傾向が見られた。刺激呈示後 2 秒から 30 秒、32 秒から 60 秒、62 秒から 90 秒の間のそれぞれの調和度の平均を求め、全実験参加者にわたり平均したところ、それぞれ -0.36, 0.10, -0.38 であった。Bonferroni 法によると、印象が類似した S1（刺激呈示後 30 秒から 60 秒）と、印象が類似していない D1（刺激呈示後 2 秒から 30 秒、および 62 秒から 90 秒）における平均調和度の中に有意な差が見られた ( $p<.01$ ,  $p<.01$ )。

同様に、S2-D2-S2 刺激における刺激呈示後 2 秒から 30 秒、32 秒から 60 秒、62 秒から 90 秒の間のそれぞれの調和度の平均を求めた。全実験参加者にわたり平均したところ、平均調和度はそれぞれ 0.51, -0.12, 0.55 であった。Bonferroni 法による多重比較を行ったところ、印象が類似していない D2（刺激呈示後 32 秒から 60 秒）と、印象が類似した S1（刺激呈示後 2 秒から 30 秒、および 62 秒から 90 秒）の平均調和度の間の差は有意であった ( $p<.001$ ,  $p<.001$ )。

D2-S2-D2 刺激でも同様の傾向が見られた。刺激呈示後 2 秒から 30 秒、32 秒から 60 秒、62 秒から 90 秒の間のそれぞれの調和度の平均を求め、全実験参加者にわたり平均したところ、それぞれ -0.26, 0.31, -0.25 であった。Bonferroni 法によると、印象が類似した状態である S2（刺激呈示後 30 秒から 60 秒）と、印象が類似していない状態である D2（2 秒から 30 秒、および 62 秒から 90 秒）における平均調和度の中に、有意な差が見られた ( $p<.001$ ,  $p<.01$ )。

これらのことから、途中で音と映像の印象の類似性が変化すると、それに伴い調和感も変化することが明らかになった。

また、音と映像の印象が類似している S1 刺激、S2 刺激と比べ、類似していない D1 刺激、D2 刺激の調和度の値は一定ではなく不安定であるという現象が見られた。音と

映像の印象が類似しておらず調和度が低い状態は、調和度が高い状態よりも不自然であり、視聴者に否定的に捉えられていると考えられる。人間は、否定的と評価をしたものには「不快」な感情を持ちその対象を回避するという傾向がある (Ortony et al., 1988)。D1, D2 刺激ではこのような不快な状態を回避しようとしたために調和度の値が不安定となったものではないかと考えられる。

次に、S1 刺激, S1-D1-S1 刺激における刺激呈示後 2 秒から 30 秒の間の調和度の推移から、音と映像が共に明るい印象で類似し意味的調和が成立した視聴覚刺激の調和感が形成されるのに必要な時間を求める。また、S2 刺激, S2-D2-S2 刺激における刺激呈示後 2 秒から 30 秒の間の調和度から、音と映像が共に暗い印象で類似し意味的調和が成立した視聴覚刺激の調和感が形成されるのに必要な時間を求める。実験 1 (2.3.2) と同様の方法で、実験参加者毎、刺激毎にピーク到達時間を求める。

表2-3 各群に分類されるデータ  
(上段) S1 (長三和音+笑顔映像)

1回目	2回目以降
S1刺激 (2秒から30秒) , S1-D1-S1刺激 (2秒から30秒) のうち, S1部分の評定が1回目の実験 参加者のデータ	S1刺激 (2秒から30秒) , S1-D1-S1刺激 (2秒から30秒) のうち S1部分の評定が2回目以降の 実験参加者のデータ

(下段) S2 (短三和音+泣き顔映像)

1回目	2回目以降
S2刺激 (2秒から30秒) , S2-D2-S2刺激 (2秒から30秒) のうち, S2部分の評定が1回目の実験 参加者のデータ	S2刺激 (2秒から30秒) , S2-D2-S2刺激 (2秒から30秒) のうち S2部分の評定が2回目以降の 実験参加者のデータ

ここで、S1 部分の評定が1度目の場合と、2回目以降の場合では、調和感の形成に要する時間が異なることも考えられる。そこで、S1 部分が含まれた S1 刺激、S1-D1-S1 刺激、D1-S1-D1 刺激のいずれの評定も行っておらず、S1 部分に対する評定が1回目である実験参加者と、先にあげた3つの刺激のいずれかの評定を既に行い、S1 部分に対する評定が2回目以降である実験参加者とに分類する。同様に、S2 部分が含まれた S2 刺激、S2-D2-S2 刺激、D2-S2-D2 刺激のいずれの評定も行っておらず、S2 部分に対する評定が1回目である実験参加者と、3つの刺激のいずれかの評定を既に行い、S2 部分に対する評定が2回目以降である実験参加者とに分類する。表 2-3 に各群に分類されるデータを示す。

表2-4 ピーク到達時間の平均，中央値  
(上段) S1 (長三和音+笑顔映像)

		1回目	2回目以降	総和
ピーク到	データ数	7	17	24
達時間	平均 (秒)	17.0	12.4	13.75
	標準偏差	3.45	2.48	9.96
	中央値 (秒)	18.25	7	9.5

(下段) S2 (短三和音+泣き顔映像)

		1回目	2回目以降	総和
ピーク到	データ数	13	16	29
達時間	平均 (秒)	20.0	13.3	16.34
	標準偏差	8.18	9.83	9.59
	中央値 (秒)	21.25	10.5	16.75

表 2-4 に、S1, S2 における両群の平均ピーク到達時間を示す。平均を求める際、調和度の値が小さい実験参加者のデータ (S1・1 回目条件の 2 名分) は除外した。S1 部分 (長三和音+笑顔映像) では、「1 回目」における平均時間は 17.0 秒であり、「2 回目以降」における平均時間は 12.4 秒であった。また、S2 部分 (短三和音+泣き顔映像) では、「1 回目」における平均時間は 20.0 秒であり、「2 回目以降」における平均時間は 13.3 秒であった。これらは、音素材・映像素材単独の印象の評定値のピーク点への到達時間 (A1 : 5 秒, A2 : 11 秒, V1 : 8 秒, V2 : 8 秒) より長い。意味的調和の成立には、音と映像それぞれの印象を捉えたうえで、両者の印象を比較する必要があると考えられる。よって視聴覚刺激の意味的調和が形成され調和度がピークに達するには、音素材・映像素材それぞれの印象がピークに達するよりも時間がかかるものと考えられる。

「1 回目」と「2 回目以降」の条件間で、平均ピーク到達時間に統計的に有意な差があるかどうかを調べるため、統計的検定を行った。データに正規性が確認できなかったので、Wilcoxon の順位和検定を行った。S2 部分での「2 回目以降」と「1 回目」の平均ピーク到達時間の差に有意な傾向が得られた ( $p < .1$ )。S1 部分では、有意な差は見られなかった。

以上から、意味的調和が成立している視聴覚刺激に対して、音と映像の調和感が形成されるのに要する時間は、12 秒から 20 秒程度であることが示された。また、S1 部分 (長三和音+笑顔映像)、S2 部分 (短三和音+泣き顔映像) とともに評定が 2 回目以降になると、1 回目の場合よりも、ピーク到達時間は短くなる傾向が見られた。ただし、S1 部分の場合、時間差は統計的に有意なものではない。一度記憶されると、調和感が形成されるのにあまり時間を要さない場合もあると考えられる。

### 2.4.3 構造的調和と意味的調和の形成過程の比較

調和感の形成に要する時間が、構造的調和によるものと意味的調和によるものとの統計的に有意な差があるのかを検討する。

図 2-8 に、「同期・1 回目」「長三和音+笑顔・1 回目」「短三和音+泣き顔・1 回目」の各条件における平均調和度の推移を示す。平均調和度のグラフからも、意味的調和が成立した条件である「長三和音+笑顔・1 回目」「短三和音+泣き顔・1 回目」の方が、構造的調和が成立した「同期・1 回目」の場合よりも調和度が上昇するのに時間を要している様子が見られる。

表 2-5 に実験 1, 実験 2 で得られた平均ピーク到達時間を示す。各群の等分散性が確認できなかったため、Non parametric 検定である Kruskal-Wallis 検定を行ったところ、各

条件間に有意な差が見られた ( $H(5) = 16.175, p < .01$ )。そこで、多重比較として、Mann-Whitney の U 検定を Bonferroni の不等式による修正を行ったうえで実施した。

「短三和音+泣き顔・1回目」のピーク到達時間は、「同期・1回目」「同期・2回目以降」におけるピーク到達時間よりも、有意に長い ( $p < .05, p < .05$ )。また、「長三和音+笑顔・1回目」のピーク到達時間においても、「同期・1回目」「同期・2回目以降」におけるピーク到達時間に比べ、有意に長い傾向が見られた ( $p < .1, p < .1$ )。このことから、意味的調和が成立している視聴覚刺激に対する調和感が形成される時間は、構造的調和が成立している場合よりも、長い時間がかかることが示された。この時間の違いは、2つの調和を形成する情報処理過程の違いを反映しているものと考えられる。構造的調和の場合は音と映像のアクセントの同期を捉え、両者が一致したときに調和していると判断される。一方、意味的調和の場合は、音と映像それぞれの印象を捉えそれらを記憶し、両者の印象を比較して一致したときに調和していると判断される。時間的に一致しているかどうかを照合するという比較的単純な過程による構造的調和の形成よりも、音と映像それぞれの印象を捉えたうえで、両者の印象を比較する意味的調和の形成の方が、時間を要するのであろう。

ただし、「短三和音+泣き顔・2回目以降」のピーク到達時間と、「同期・1回目」「同期・2回目以降」のピーク到達時間との間には有意な差は見られなかった。同様に、「長三和音+笑顔・2回目以降」のピーク到達時間と、「同期・1回目」「同期・2回目以降」のピーク到達時間の間にも有意な差は見られなかった。このことから、意味的調和の場合には、一度その刺激を見たことがある場合は、形成に要する時間は構造的調和におけるものと差がないほど短くなる場合もあると考えられる。

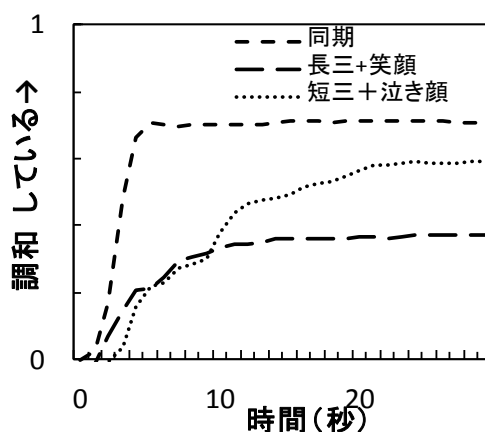


図 2-8 「同期・1回目」「長三和音+笑顔・1回目」「短三和音+泣き顔・1回目」の各条件における平均調和度の推移

表 2-5 平均ピーク到達時間

	同期・ 1回目	同期・ 2回目以 降	長三和音 +笑顔・ 1回目	長三和音 +笑顔・ 2回目以 降	短三和音 +泣き顔・ 1回目	短三和音 +泣き顔・ 2回目以 降
平均ピーク到達 時間(秒)	7.4	8.1	17.0	12.4	20	13.3

## 2.5 本章のまとめ

構造的調和、意味的調和が成立した単純な視聴覚刺激を用いて、連続測定により音と映像の調和感を測定した。

測定によって得られた調和度の値は、同期した状態において高く、非同期の状態においては低かった。この傾向により、構造的調和が成立することにより調和感が高まることが確認された。また、音と映像の印象が類似した状態において調和度は高く、類似していない状態において低かった。この傾向から、意味的調和が成立することにより調和感が高まることが確認された。



調和感が形成されるのに要する時間を求めたところ、構造的調和の形成に要する時間は、7秒から8秒程度であった。一方、意味的調和の形成に要する時間は12秒から20秒程度であり、構造的調和における場合よりも長い時間であった。初めて視聴する視聴覚刺激の場合には、意味的調和と構造的調和との間で、調和感の形成に要する時間に有意な差が見られた。このことから、意味的調和の方が構造的調和の形成よりも時間を要することが明らかになった。ただし、意味的調和の場合、一度視聴した刺激の場合は、調和感の形成時間が短くなる傾向が得られた。

## 第3章

# 市販の映像作品を用いた音楽と映像の調和感の連続測定

### 3.1 はじめに

2章では、音と映像の構造的調和、意味的調和の状態をコントロールし、音と映像のアクセントが同期した条件および非同期である条件、印象が類似した条件および類似していない条件を意図的に作り出し、それらにおける音と映像の調和感が形成される過程を検討した。意味的調和による調和感は、構造的調和による調和感よりも、形成に時間を要することを明らかにした。

3章では、音楽と映像の調和感に関する映像コンテンツ制作者の意図の効果を探るため、オリジナルの作品をそのままに使用した刺激と、別の作品の音楽と映像を組み合わせた刺激を用いて、音楽と映像の間に成立する調和感の状況について考察した。また、映像作品の調和感を形成する要因として考えられる音楽と映像の印象についても、連続的な測定方法で測定し、音楽と映像の調和感との関係を考察した。

実験で用いた方法は、音楽と映像の調和感の測定には2章で用いた連続測定法を、音楽や映像の印象の測定には連続記述選択法を用いた。

### 3.2 実験3：市販の作品を用いた視聴覚刺激の音楽と映像の調和感の連続測定

実験3では、音楽と映像の調和感に関する制作者の意図を探るために、市販の映像作品と、制作者の意図を無視して音楽と映像を組み合わせたものに対して、音楽と映像の調和感の連続測定を行った。実験刺激として、Disneyのアニメーション映画の一部を抜粋したものと、別々の作品同士の音楽と映像を組み合わせたものを用いた。

## 3.2.1 実験方法

### 3.2.1.1 実験参加者

実験参加者は、九州大学の学生 15 名（男性 8 名，女性 7 名，平均年齢 23.1 歳）であった。

### 3.2.1.2 実験刺激

実験刺激には、DVD で市販されている「FANTASIA の禿山の一夜（Disney, 2000, 開始後 1 時間 43 分 16 秒から 1 時間 45 分 16 秒まで）」、「FANTASIA 2000 の威風堂々（Disney and Ernst, 2000, 開始後 56 分 23 秒から 59 分 3 秒まで）」よりそれぞれ 160 秒間を抜粋したものを使用した。160 秒間という抜粋時間は、作品としてのまとまりと変化がありつつストーリーが完結し、実験刺激として長すぎないという基準で、「威風堂々」の切り取りの長さを定めたものである。「禿山の一夜」に関しては、明確なストーリーが無くより自由な抜粋時間が選択できるが、「威風堂々」と同じ 160 秒で音楽の終了と合わせて無理なく区切れることから、同一の長さにした。この 2 作品は、クラシックの楽曲にアニメーション作品を組み合わせたもので、台詞や効果音は含まれず、音楽と映像の調和感を検討するのに好都合な素材である。図 3-1 に「禿山の一夜」と「威風堂々」から抜粋した場面のストーリーのブロック図を示す。

抜粋した 2 種類の視聴覚刺激を A1V1, A2V2 とし、これらを「オリジナル刺激」と呼ぶ。また、2 つのオリジナル刺激の音楽を交換し、音楽と映像の組み合わせが元の作品とは異なる刺激 A1V2, A2V1 を作成した。これらを「組み合わせ刺激」と呼ぶ。ここで、A1 は「禿山の一夜」の音楽、V1 は「禿山の一夜」の映像、A2 は「威風堂々」の音楽、V2 は「威風堂々」の映像を表す。オリジナル刺激 2 種類、組み合わせ刺激 2 種類に対して音楽と映像の調和感の連続測定実験を行った。

### 3.2.1.3 実験装置

実験装置は、2 章実験 1 (2.2.2) で使用したものと同様であった。

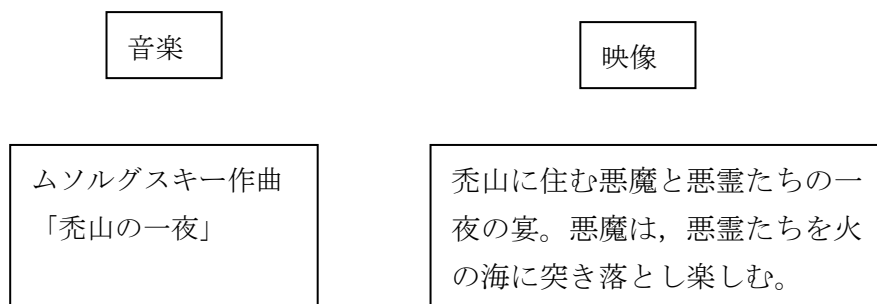
### 3.2.1.4 実験手続き

実験手続きは、2 章実験 1 (2.2.3) と同様であった。

## 3.2.2 結果と考察

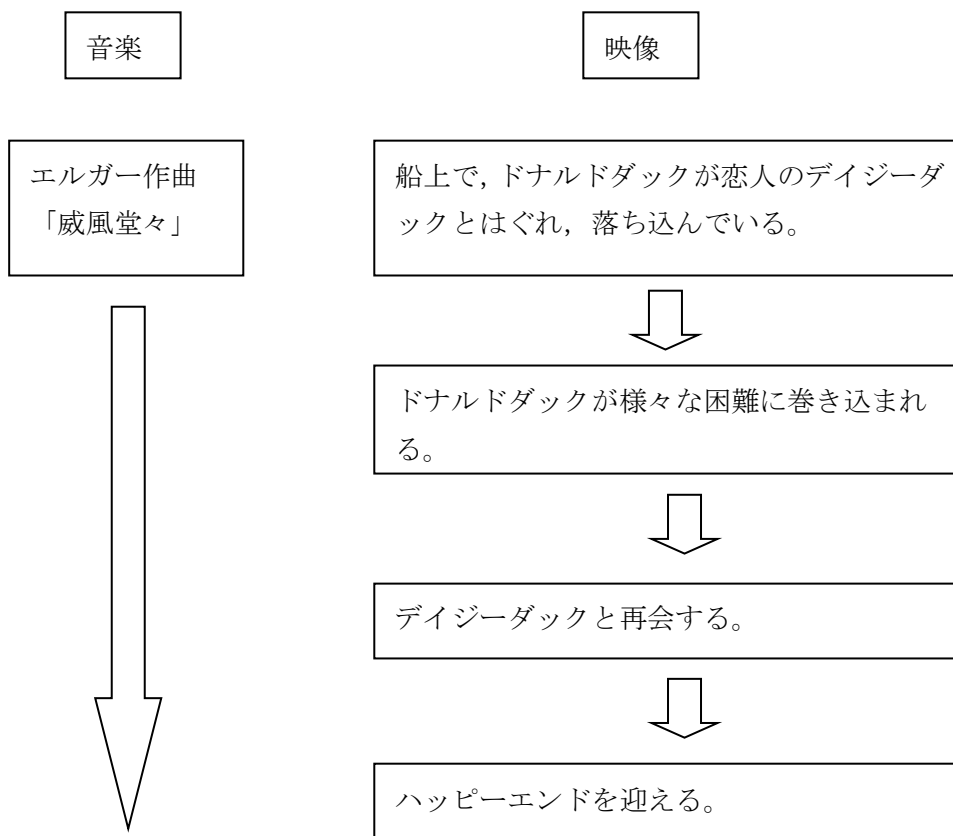
2 章の予備実験によると、マウス操作に要する時間は 2 秒であった。これを考慮して、本章でも調和感の連続測定で得られたデータは、2 秒以降のものを考察の対象とする。

図 3-2 に、A1V1, A2V2, A1V2, A2V1 に対する各時間における音楽と映像の調和度



## (a) 「禿山の一夜」

ムソルグスキー作曲「禿山の一夜」にのせた、  
明確な物語のないアニメーション



## (b) 「威風堂々」

エルガー作曲「威風堂々」にのせて主人公のドナルドダックを中心とした  
起承転結のある物語

図 3-1 「禿山の一夜」「威風堂々」より抜粋した場面のストーリー

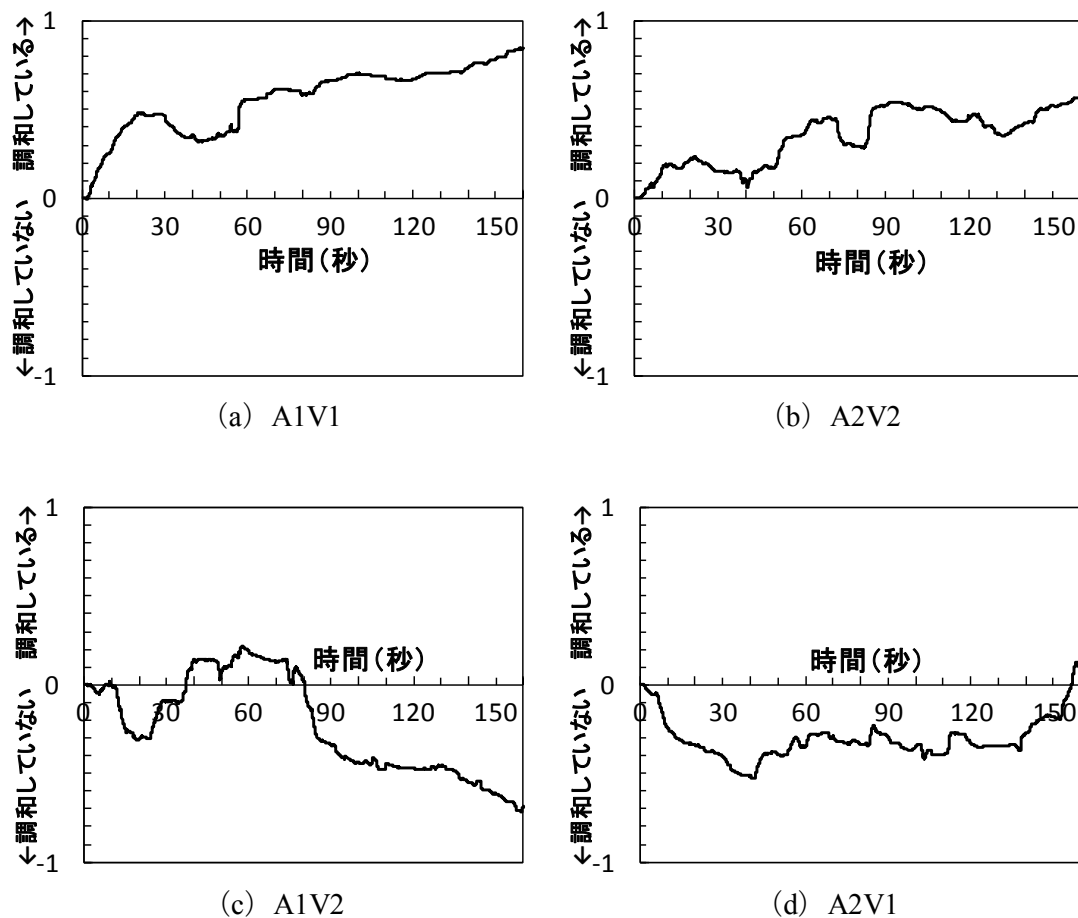


図 3-2 市販の作品を用いた各視聴覚刺激における各時点での音楽と映像の調和度

を示す（全実験参加者の平均値）。音楽と映像が元々の作品通りの組み合わせである A1V1, A2V2 の刺激開始後 2 秒以降における調和度の平均値は、正の値である（A1V1 ; 平均  $M = 0.56$ , 標準偏差  $SD = 0.11$ , A2V2 ; 平均  $M = 0.34$ , 標準偏差  $SD = 0.22$ ）。オリジナル作品とは異なる音楽と映像を組み合わせさせた A1V2, A2V1 の刺激開始後 2 秒以降における調和度の平均値は、負の値である（A1V2 ;  $M = -0.23$ ,  $SD = 0.26$ , A2V1 ;  $M = -0.30$ ,  $SD = 0.29$ ）。オリジナル刺激（A1V1, A2V2）と組み合わせ刺激（A1V2, A2V1）の調和度の平均値を比較するため、対応のある  $t$  検定を行い、両群間の平均値の差を検定した。その結果、オリジナル刺激の調和度は組み合わせ刺激の調和度よりも、有意に高い値であることが明らかになった ( $t(14) = 10.97$ ,  $p < .001$ )。このことから、制作者は、音楽と映像が調和するよう意図して制作していると考えられる。このような傾向は、岩宮（1992）や Lipscomb and Kendall（1994）の研究でも認められている。

さらに、連続測定の結果からよりきめ細かに調和度の推移を観察する。図 3-2 による

と、オリジナル刺激である A1V1, A2V2 の両刺激において、調和度は、刺激開始から終了の時点まで、途中で上下することはないもののしだいに高まる傾向が見られた。t 検定によると、A1V1, A2V2 とともに、刺激呈示後 160 秒経過後の平均調和度は 2 秒経過後の平均調和度に比べ、有意に高い値であった (A1V1 ;  $t(14) = -22.11, p < 0.001$ , A2V2 ;  $t(14) = -8.27, p < 0.001$ )。組み合わせ刺激では、このように最後に向かって高まるような傾向は見られなかった。

2 章と同様の方法で、調和度のピーク到達時間を求めたところ、A1V1 は 156.5 秒、A2V2 は 158 秒であった。ピーク到達時間はいずれも、刺激が終了する間際である。2 章で示されたように、単純な刺激を用いた場合、意味的調和が成立し音楽と映像の調和感がピークに達するのに必要な時間は 17 秒から 20 秒程度である。本章で対象とした、A1V1, A2V2 においては、意味的調和が成立し調和感がピークに到達するのに必要な時間を過ぎても、調和感が上昇し続けている。

### 3.3 実験 4 : 市販の作品を用いた視聴覚刺激の各時点での音楽と映像の印象の連続記述選択実験

実験 3 では、音楽と映像の調和感の連続測定により、オリジナル刺激 A1V1, A2V2 の調和度が組み合わせ刺激 A1V2, A2V1 の調和度よりも有意に高いことが示された。さらに、オリジナル刺激 A1V1, A2V2 の両刺激とも、調和度は刺激の終盤の時間帯に向かって徐々に上昇する傾向が見られた。

実験 4 では、調和感の要因であると考えられる、音楽と映像の意味的調和について、刺激を視聴しながら連続的に評価を行う方法で検討した。それぞれの刺激の音楽 A1, A2 および映像 V1, V2 に対する印象を連続記述選択法と呼ばれる手法で評価実験を行い、実験 3 で用いた視聴覚刺激の音楽と映像の印象がどれだけ類似しているかを検討した。さらに、音楽と映像の調和度と音楽と映像の類似性の関係について解析し、意味的調和の形成過程について考察した。

#### 3.3.1 実験方法

##### 3.3.1.1 実験参加者

実験参加者は、九州大学の学生 19 名 (男性 12 名, 女性 7 名, 平均年齢 22.2 歳) であった。

### 3.3.1.2 実験刺激

実験刺激は、実験1で用いた視聴覚刺激の音楽および映像を単独で呈示したA1(A1V1の音楽)、A2(A2V2の音楽)、V1(A1V1の映像)、V2(A2V2の映像)であった。

### 3.3.1.3 実験装置

実験装置は、2章実験2で使用したものと同様であった。

### 3.3.1.4 実験手続き

実験4では、音楽と映像の調和感が感じられる過程を検討するために、音楽のみ(A1, A2)あるいは映像のみ(V1, V2)を実験刺激として、それぞれの印象評定実験を行った。印象評定法には、連続記述選択法を用いた(難波, 桑野, 2008)。連続記述選択法とは、音楽や映像などの時間経過とともに変化する刺激に対し、各時点で用意された印象を表わす形容詞の中から、あてはまる印象を選択する方法である。連続記述選択法を用いたのは、映像や音楽の多岐にわたる印象を、時間経過とともに効率的に捉えるのに最適であると考えたためである。印象評定に用いた形容詞としては、過去の研究(岩宮, 1992; 岩宮ら, 2002; 金ら, 2005)を参考にして、映像および音楽の印象を網羅すると考えられているものを選択し、「陽気な」「陰気な」「緊張感のある」「のんびりした」「迫力のある」「物足りない」「スピード感のある」「スピード感のない」(8種類)を用いた。

本実験のように、印象(情緒的意味)を評定する実験では、個人によって評定値に差が生じ、実験参加者群によって得られる結果が異なる可能性もある。この点に関し、谷口(1995)は、音楽の印象を測定する尺度を構成し、その信頼性を調べている。彼は、音楽の印象を測定する尺度として「高揚尺度」「強さ尺度」「軽さ尺度」「荘重尺度」を用いて、異なる実験参加者群による音楽の印象評定実験を行い、作品、尺度ごとに評定値の相関係数を求めたところ、80%以上で0.50以上の値が得られた。この結果より、音楽の印象を測定する実験において、異なる実験参加者群でも同様の結果が得られることが明らかになった。本実験で用いる形容詞である「明るさ」は「高揚尺度」に、「迫力のある」は「強さ尺度」に、「スピード感のある」は「軽さ尺度」に、「緊張感のある」は「荘重尺度」に含まれる。本実験で用いる形容詞についても、実験参加者によらずある程度一貫した結果が得られると考えられる。

連続記述選択法では、刺激の印象を表す形容詞をあらかじめ用意し、各形容詞とコンピュータのキーとを対応付けておく。実験参加者には、音楽を聴きながらあるいは映像を見ながら、その時点における印象を判断し、対応するキーを押すように求めた。複数の印象を同時に感じた場合には、複数のキーを交互に押すように指示した。例えば、A,

B2 つのキーを押す場合は、ABABAB...のように押すように指示した。各実験参加者が選択した形容詞は、1秒ごとに1秒間の時間間隔で計測する。1秒間に、複数の形容詞のキーが押された場合はあればそれぞれをカウントする。1秒間に同じ形容詞のキーが複数回押されても1個とカウントする。

### 3.3.2 結果と考察

図 3-3 に、連続記述選択法において、映像 V1 および音楽 A1 に対して各 1 秒間に各形容詞を選択した人数を 1 秒ごとに示す。映像 V1 に対しては、全般的に「緊張感のある」「陰気な」の形容詞が選択されていた。V1 は、起承転結のない単調な映像である。印象も全体を通してあまり変化が見られない。音楽 A1 の印象も、全般的に「緊張感のある」「陰気な」が選択され、「のんびりした」「陽気な」はほとんど選択されていない。さらに最後の部分は、音楽、映像の両方とも「スピード感のある」が選択されている。A1V1 の調和度が高いのは、類似した印象の音楽と映像が組み合わせられ、意味的調和（岩宮，2011；Bolivar et al., 1994）が形成されているためと考えられる。

図 3-4 に、連続記述選択法において、映像 V2 および音楽 A2 に対して各 1 秒間に各形容詞を選択した人数を 1 秒ごとに示す。映像 V2 は、主人公のドナルドダックを中心に起承転結のある物語が展開している。図 3-4 に示すように、始まりから 50 秒あたりまでは、映像 V2、音楽 A2 とともに「のんびりした」が選択されている。50 秒から 90 秒あたりは、両方とも、「緊張感のある」「スピード感のある」が選択されている。最後の部分（140 秒以降）は、両方とも「陽気な」が選択されている。これらの時間帯では、A2 と V2 は同じ印象を持つために、意味的調和が形成され、A2V2 の調和度は高く評定されたと考えられる。110-140 秒にかけては A2 と V2 で特に印象が一致しているわけではないが、両方の印象が対立していることもなく、調和度は高い水準を保っている。一部ではあるが、動物のグループが行進するシーン（100 - 105 秒）で、行進する動物の足の動きが音楽のリズムと同期しているシーンがあり構造的調和の効果がみられるが、ごく短時間である。なお、本研究で用いた A1V1、A2V2、A1V2、A2V1 の 4 つの視聴覚刺激において、音楽と映像の同期の効果が生ずる可能性のある箇所はこの箇所のみである。よって、残りの部分から感じられる調和感は、意味的調和に起因するものと考えられる。



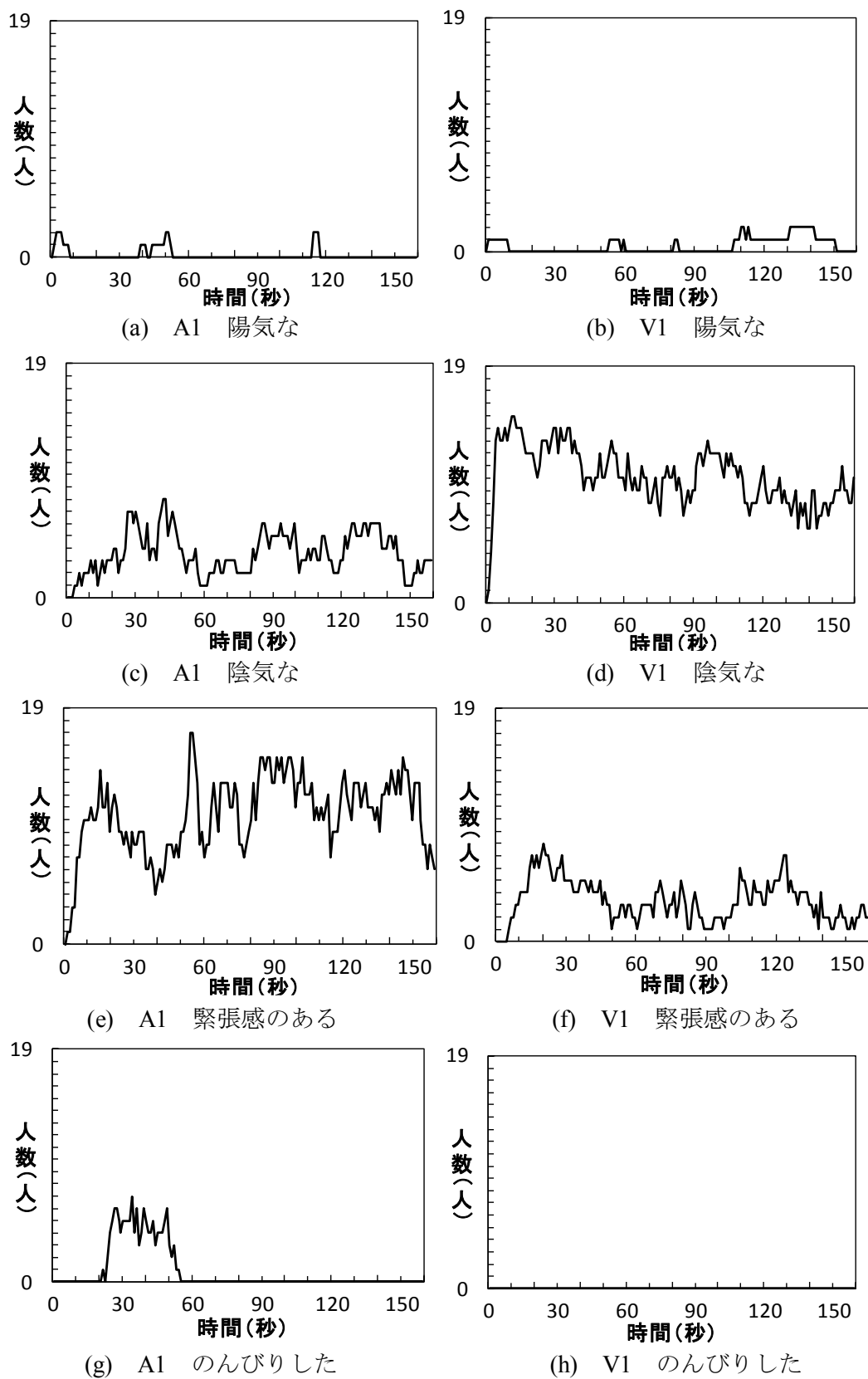


図 3-3 連続記述選択法において V1 および A1 に対して各形容詞を選択した人数 (1)

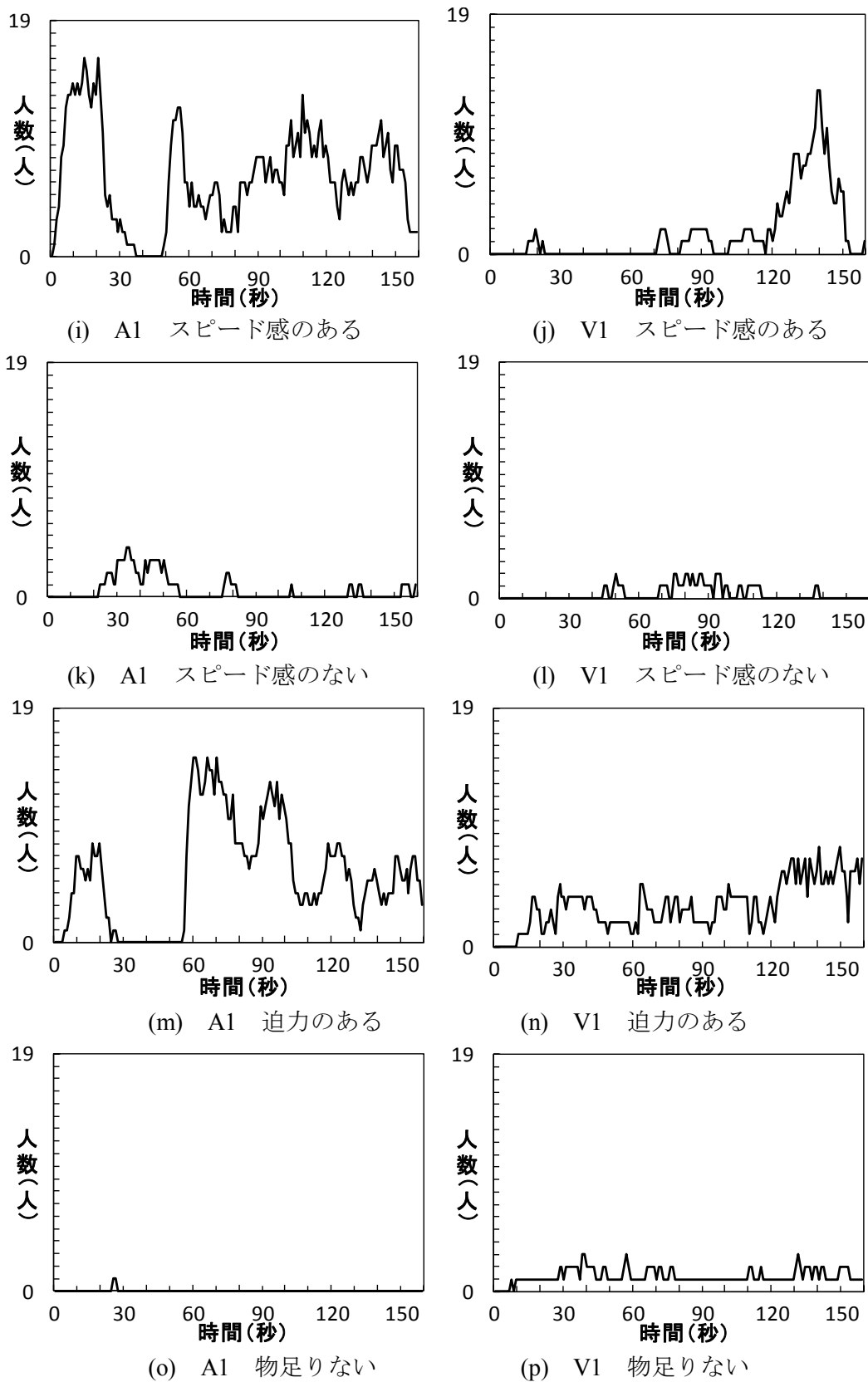


図 3-3 連続記述選択法において V1 および A1 に対して各形容詞を選択した人数 (2)

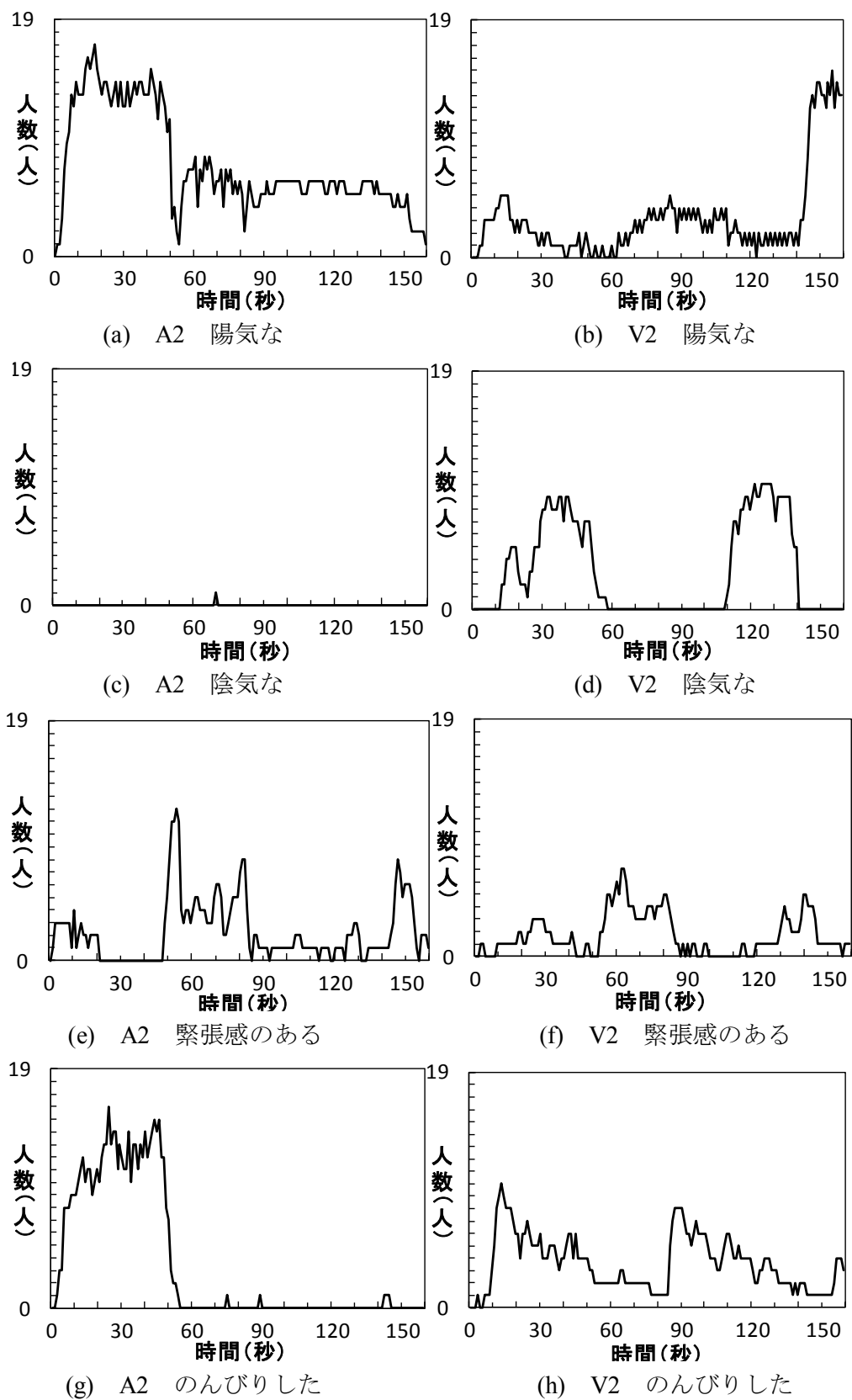


図 3-4 連続記述選択法において V2 および A2 に対して各形容詞を選択した人数 (1)

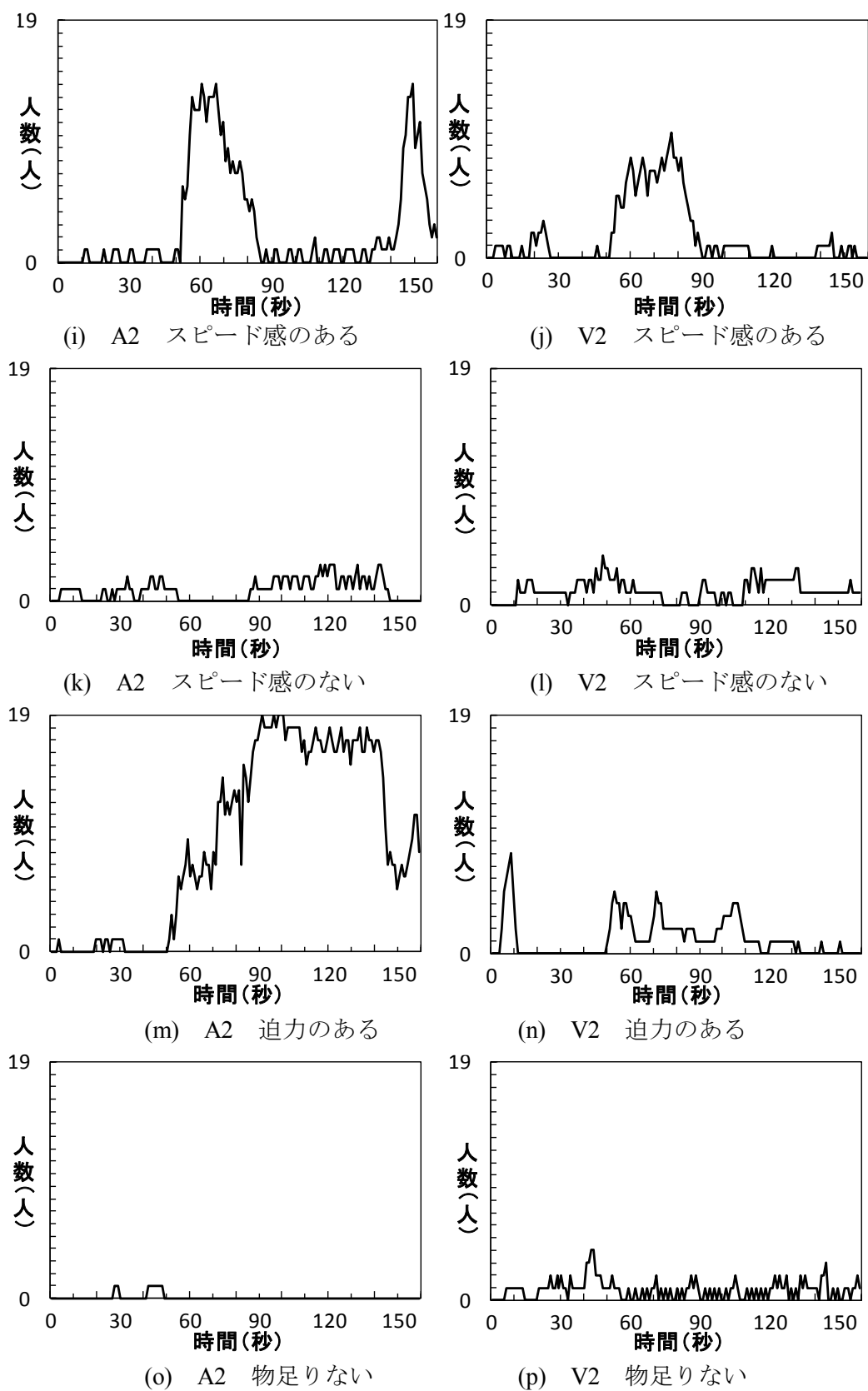


図 3-4 連続記述選択法において V2 および A2 に対して各形容詞を選択した人数 (2)

さらに、音楽と映像の印象の一致の程度を定量的に見積もるために、1秒ごとに選択された形容詞をもとに、「音楽と映像の印象の一致度」という指標を用いる。音楽と映像の印象の一致度は、1秒間における「〈音楽と映像で共通して選択された形容詞ののべ人数〉－〈同時に選択された反対の意味を持つ形容詞ののべ人数〉」と定義する。この定義では、同じ形容詞を複数の実験参加者が選択した場合は、その人数をカウントする。音楽と映像で反対の意味の形容詞を選択した場合は、その数を減ずるものとする。

図 3-5 に、各刺激呈示時間における各刺激の印象の一致度を示す。オリジナル刺激を構成する A1 と V1 および A2 と V2 の一致度は、多くの部分がプラスになり、同じ形容詞が選択されている様子が分かる。組み合わせ刺激を構成する A1 と V2 および A2 と V1 の一致度は、全般的にオリジナル刺激を構成する A1 と V1, A2 と V2 の一致度よりも低く、多くの部分が 0 付近である。刺激呈示後 0 秒から 160 秒までの 161 個の一致度の値を平均すると、オリジナル刺激を構成する A1 と V1 および A2 と V2 は 8.30, 4.43 となり、組み合わせ刺激を構成する A1 と V2 および A2 と V1 は -0.39, -5.08 となる。オリジナル刺激の音楽と映像 (A1 と V1, A2 と V2) と組み合わせ刺激の音楽と映像 (A1 と V2, A2 と V1) の一致度の 2 群間における差を比較するため、平均値の差の検定を行った。2 群間の等分散性が仮定できなかったため、Welch の検定を行った。その結果、2 群間に有意な差が見られた ( $t'(612.165)=18.09, p<.001$ )。

一致度を用いた定量的な検討によっても、オリジナル刺激 A1V1, A2V2 では組み合わせ刺激 A1V2, A2V1 に比べ、音楽と映像の印象に対して同一の形容詞が選択されていることが明らかになった。このことから、オリジナル刺激において作者の意図により意味的調和が形成されている様子がわかる。組み合わせ刺激を構成する A1 と V2 および A2 と V1 では、全般的に一致度が低く音楽と映像の印象が一致していない。このため組み合わせ刺激では意味的調和が形成されず、全般的には調和感が低くなったと考えられる。ただし、A1V2 の 40-70 秒付近のように、若干ではあるが調和度が高まる部分もある。この部分では、音楽と映像の両方ともに「緊張した」が選択されており、意味的調和が偶然に形成されていたと思われる。

### 3.3.3 音楽と映像の調和度と音楽と映像の印象の一致度の関係

オリジナル刺激 A1V1, A2V2 の調和度の平均値は両刺激とも正の値であった。また、オリジナル刺激を構成する音楽と映像の印象の一致度も、A1 と V1 および A2 と V2 ともに正の値であった。このことから、オリジナル刺激の調和感を形成している要因は、音楽と映像の印象の一致である意味的調和であることが明らかになった。

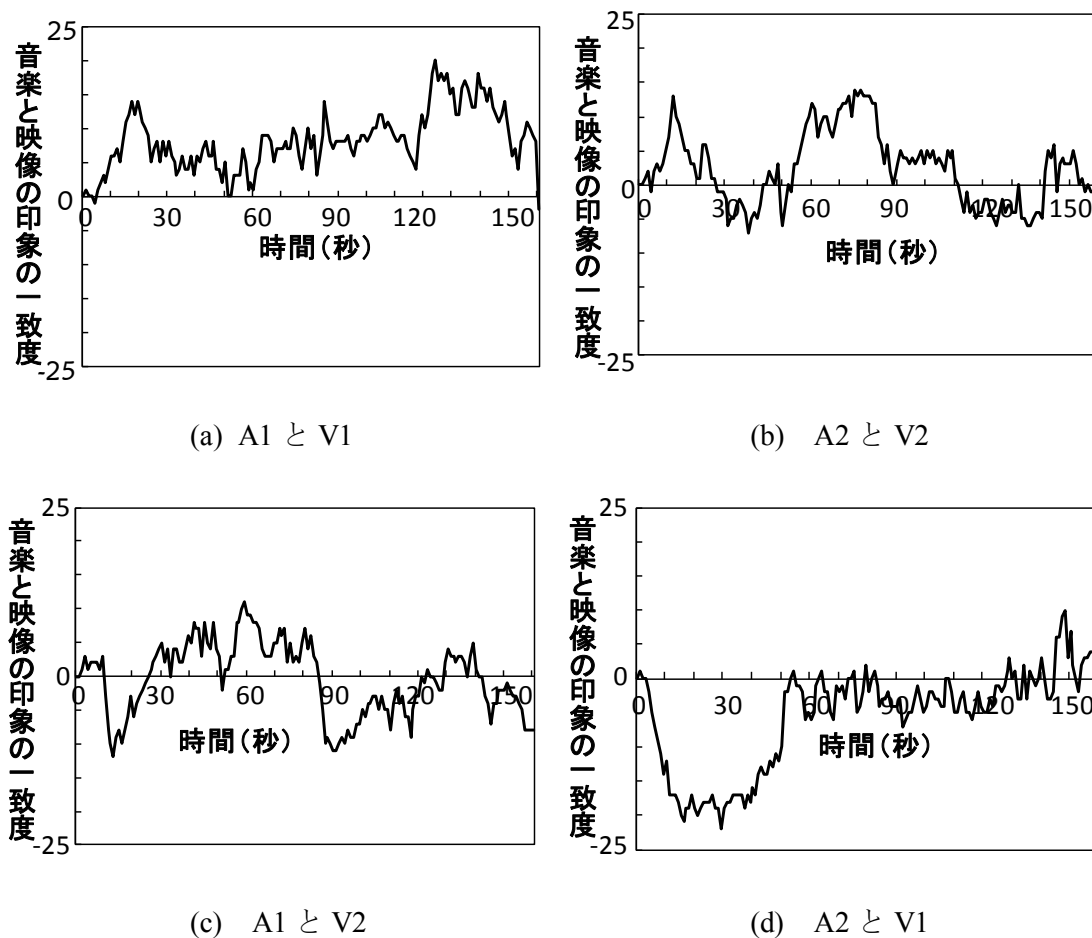


図 3-5 各刺激の印象の一致度

オリジナル刺激を構成する音楽と映像の印象の一致度は、A1 と V1 および A2 と V2 ともに刺激呈示後から正の値であり、初期の段階で意味的調和が形成されている状況にある。オリジナル刺激の調和度は徐々に上昇し、20 秒付近で一度ピークに達する。これは、2 章で示した意味的調和が形成されピークに到る時間とほぼ一致する。

しかし、図 3-5 (b) A2 と V2 の 110-140 秒のように、途中で音楽と映像の印象の一致度が低く (0 点付近)、音楽と映像の印象が類似していない箇所もある。このような箇所でも調和度は高いレベルを維持していることから、一旦意味的調和が成立して調和感が上昇すると、途中でその要因が弱まっても調和感は維持されるものと考えられる。

3.2.2 で指摘したように、オリジナル刺激の A1V1, A2V2 においては、2 章で得られた意味的調和が成立し調和感がピークに到達するのに必要な時間を過ぎても、調和度がさらに上昇し続けている。A1V1 の音楽である A1 においては、後半が「迫力のある」(図 3-3 (m)) 印象になっている。また、A2V2 では映像 V2 の後半が「迫力のある」(図

3-4 (n)) 印象となっている。作品の制作者は、本研究で用いた刺激の後半に作品の盛り上がりを見せていると考えられる。こういった盛り上げる展開をした視聴覚刺激の場合、意味的調和が形成された音楽と映像の調和感はさらに上昇することが示唆される。

### 3.4 実験 5：意味的調和の形成過程に関する実験

実験 3、実験 4 により、オリジナル刺激 (A1V1, A2V2) では、音楽と映像の印象が類似した状態である意味的調和が図られており、組み合わせ刺激 (A1V2, A2V1) の調和度よりも有意に高いことが示された。さらに、刺激呈示直後より調和度は上昇傾向にあることから、2 章でも示されたように調和感が形成されるには時間がかかることが示された。

実験 5 では、実験 1、実験 2 と同じ映像作品を用いて、A1V1 や A2V2 の始まりの場面よりも前から抜粋し A1V1 や A2V2 の前半部であった場面が後半部になるように抜粋した視聴覚刺激や、A1V1 や A2V2 の後半部から抜粋した視聴覚刺激、A1V1 の前半と後半を入れ替えた視聴覚刺激を用いて、意味的調和が形成される過程について検討した。

もし、どのような映像作品においても意味的調和が形成されるのに時間がかかり、調和度は低い値から上昇する傾向を持つならば、十分に意味的調和が形成されておらず調和度が低い水準であった A1V1 や A2V2 の始めの場面でも、その場面の前から呈示すると調和度は高い水準となるはずである。

#### 3.4.1 実験方法

##### 3.4.1.1 実験参加者

実験参加者は、九州大学の学生 20 名 (男性 13 名, 女性 7 名, 平均年齢 22 歳) であった。これらは、実験 3 の実験参加者とは異なる。

##### 3.4.1.2 実験刺激

視聴覚刺激には、実験 3 で用いた A1V1, A2V2 と同じ映像作品を用いた。同じ DVD の A1V1 の始まりよりも 81 秒前から 160 秒間を抜粋したものを A1V1\_b, A1V1 の始まりの 79 秒後から 81 秒間だけを抜粋したものを A1V1\_c と呼ぶ。A1V1 においては明確なストーリーがないので、元の刺激 A1V1 の中間点 (79 秒後) を抜粋の区切りにした。同じ DVD の A2V2 の始まりよりも 111 秒前から 160 秒間を抜粋したものを A2V2\_b, A2V2 の始まりの 49 秒後から 111 秒間を抜粋したものを A2V2\_c と呼ぶ。A2V2 はストーリーがあるので、A2V2 のストーリーの転換点 (49 秒後) を抜粋の区切りにした。また、

A1V1 の前半 79 秒間と後半 81 秒間を入れ替えたものを A1V1\_i と呼ぶ。これらの 5 つの視聴覚刺激に対して、音楽と映像の調和感の連続測定実験を行った。なお、A2V2 はストーリーを伴うものであり、A1V1\_i のように前半と後半を入れ替えると違和感が生じるため、そのような刺激は使わなかった。

#### 3.4.1.3 実験装置

実験装置は、2 章実験 1 (2.2.2) で使用したものと同様であった。

#### 3.4.1.4 実験手続き

実験手続きは、2 章実験 1 (2.2.3) と同様であった。

### 3.4.2 結果と考察

実験結果を図 3-6 に示す。A1V1\_b は、A1V1 が始まる 81 秒前から抜粋した刺激である。A1V1 に対する評定の始まりから 2 秒後の全実験参加者の調和度の平均値は 0.00 であった。一方、A1V1\_b における同じ箇所の全実験参加者の調和度の平均値は、0.28 であった。Welch の検定によると、A1V1\_b におけるこの箇所の調和度は、A1V1 に比べ有意に高い ( $t'(19.75) = 2.85, p < 0.01$ )。

A1V1\_c は A1V1 が始まって 79 秒後から始まる刺激である。A1V1\_c の始まりから 2 秒後の全実験参加者の調和度の平均値は 0.02 である。一方、A1V1 における同じ箇所の全実験参加者の調和度の平均値は 0.59 であった。Welch の検定によると、A1V1 におけるこの箇所の調和度は、A1V1\_c に比べ有意に高い ( $t'(15.81) = 11.63, p < 0.001$ )。

A2V2, A2V2\_b, A2V2\_c の調和度の比較からも、同様の傾向が得られた。A2V2 の始まりから 2 秒後の全実験参加者の調和度の平均値は 0.00 であった。一方、A2V2 が始まる 111 秒前から刺激が呈示される A2V2\_b における同じ箇所の全実験参加者の調和度の平均値は、0.35 であった。両調和度の差は統計的に有意であった ( $t'(19.07) = -4.93, p < 0.001$ )。また、A2V2 が始まって 49 秒後から始まる A2V2\_c の始まりから 2 秒後の全実験参加者の調和度の平均値は 0.00、A2V2 における同じ箇所の全実験参加者の調和度の平均値は 0.19 であった。統計的に有意傾向ありのレベルではあるが、A2V2 の場合の方が A2V2\_c の場合よりも調和度の平均値は高い ( $t'(14.06) = 1.93, p < 0.1$ )。同じ箇所でも刺激呈示後ある程度時間が経過してからのの方が、調和度は高くなっている。

図 3-6 (c) は A1V1 の前半部分 (79 秒間) と後半部分 (81 秒間) を入れ替えた A1V1\_i に対する調和度を、A1V1 と同じ順序になるように入れ替えて表示したものである。A1V1\_i の刺激呈示後 2 秒後の全実験参加者の調和度の平均値は 0.01 であり、A1V1 における同じ箇所 (A1V1 の刺激呈示後 81 秒) の全実験参加者の調和度の平均値は 0.59



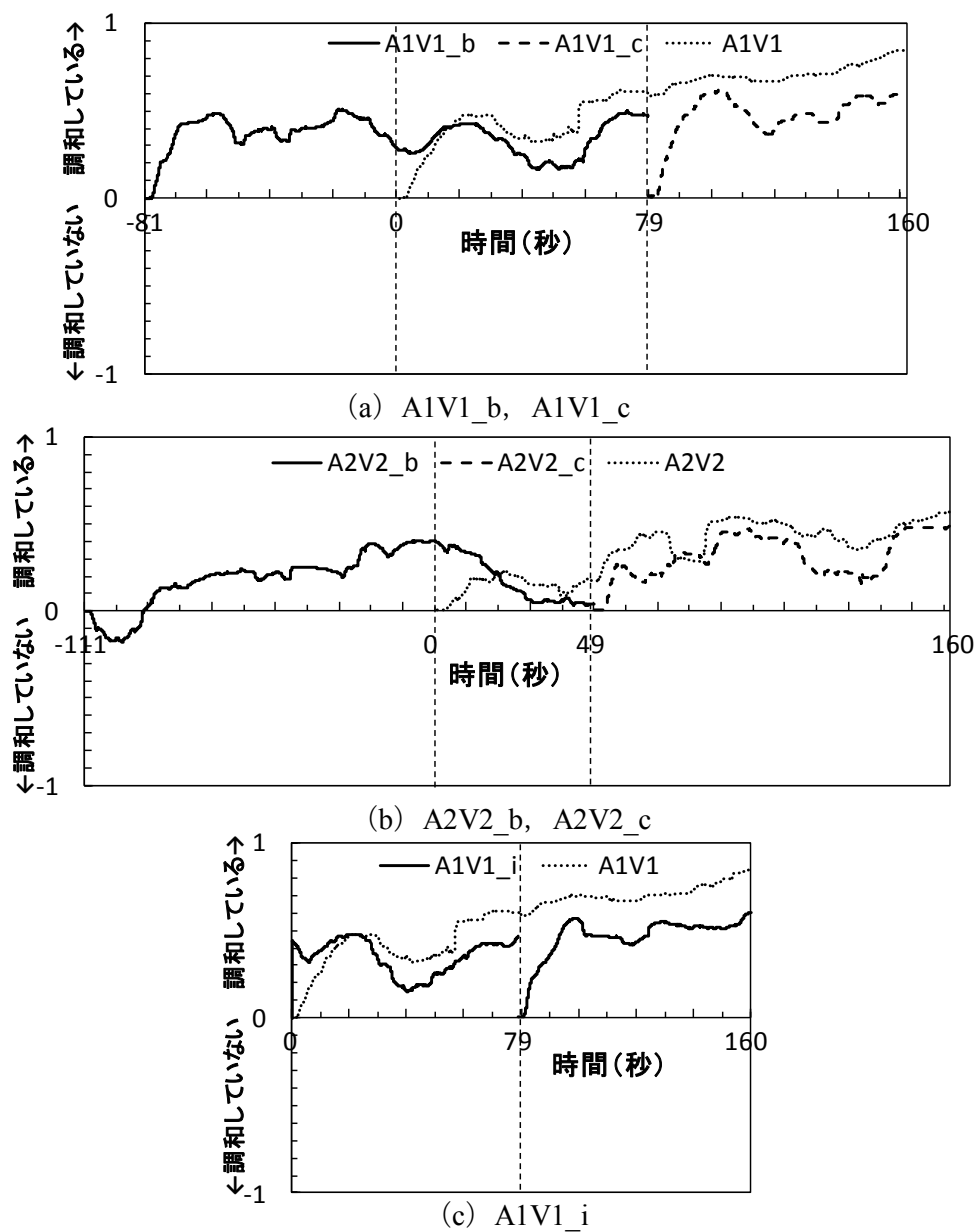


図 3-6 市販の作品を用いた各視聴覚刺激における各時点での音楽と映像の調和度

であった。両調和度の差は統計的に有意であった ( $t'(14.42) = 11.97, p < 0.001$ )。

以上のように、同じ場面でも刺激呈示直後には調和度は低く、ある程度時間が経過した後では調和度は高い傾向が見られた。これは、刺激呈示直後には意味的調和がまだ十分に形成されておらず、刺激呈示後ある程度時間が経過すると十分に意味的調和が形成されるためであると考えられる。このことから、視聴覚コンテンツにおける意味的調和の形成には時間を要することが明らかとなった。

### 3.5 本章のまとめ

市販の映像作品の一部を用いた音楽と映像の調和感に関する連続測定実験により、制作者は意図的に音楽と映像の印象を類似させ、意味的調和の効果を演出していることが示された。任意の音楽と映像では、音楽と映像の印象は類似せず、調和感は得られない。

音楽と映像の調和感は、刺激呈示直後ではそれほど高くないが、ある程度の時間をかけて、しだいに調和感が高まっている。前半と後半を入れ替えて呈示したり、中間部から呈示したりしても、調和度は時間をかけて上昇する傾向が見られた。市販の作品でも、意味的調和の形成に時間がかかることが示された。これは、2章で示した結果を裏付けるものである。任意の音楽と映像では、このような傾向は見られなかった。

また、視聴覚刺激 A2V2 で示されたように、一度調和度が上昇すると、その後に音楽と映像の印象が類似していない部分があっても、調和度は高い水準を保っている。このことから、意味的調和が形成されると、その要因が弱まっても効果は持続することが示唆された。

## 第4章

# 黒澤明の作品における音と画を対比させる手法の効果

### 4.1 はじめに

2章において、条件を統制した単純な視聴覚刺激を用いて、音と映像の印象が類似することにより音と映像の調和感が高まることを示し、意味的調和に基づく調和感の形成過程を検討した。3章においては、Disneyのアニメーション作品を用いて、市販の作品における時々刻々の音楽と映像の印象、および音楽と映像の調和感を検討した。3章で用いた作品においては、音楽と映像の印象を類似させ意味的調和に基づく調和感を形成しようとする制作者の意図が示唆された。

3章で示したように、通常、映像作品における音楽はその場の状況や登場人物の心情に合わせた印象の音楽が利用され、意味的調和が図られている。ところが、映像作品の中には、わざと音楽と映像の印象を一致させないことで、独特の効果をねらったものがある。

映画監督の黒澤明は、音と映像の関係に独特のこだわりをもっており、映像で展開されているシーンの印象とは全く正反対の音楽を用いることで、観客に非常に強烈な印象を与えることを狙ったと証言されている（西村，1998）。このような手法は黒澤作品のいくつかでみられ、黒澤明は「劇と音楽の対位的な処理」と呼んでいた。後に映画評論家の西村は、この手法を「音と画の対位法」と称してその効果を論じている（西村，1998）。西村の著書（西村，1998）によると、「黒澤監督は、音と画の対位法によって、映像または音だけではできない、また、映像と音楽が調和しているだけでもできない、新たな心理空間を映像作品に創造した」と述べられている。ただし、その効果について、実証した研究はない。

そこで4章では、音と画の対位法が用いられた映画の一部を抜粋した刺激、およびその部分の音楽を入れ替えた刺激を用いて、刺激の時々刻々の印象および刺激全体の印象の評定実験を行い、音と画の対位法がその瞬間の音楽と映像の調和感や印象、および視

聴後の印象に及ぼす効果を明らかにした。本論文では、音と画の対位法を単に「対位法」と略すこともある。なお、音と画の対位法は、音楽用語の対位法の用語を借りてはいるが、あくまでも比喩的な表現で、音楽用語での対位法とは異なる。本論文では音楽用語としての対位法は使用していない。

## 4.2 実験 6：視聴覚刺激の音楽と映像の調和感の連続測定

実験 6 では、黒澤明監督作品から音と画の対位法が用いられている場面を抜粋したものを実験刺激として、音楽と映像の調和感の連続測定を行った。対位法が使われた場面と使われていない場面の音楽と映像の調和度を比較した。

### 4.2.1 実験方法

#### 4.2.1.1 実験刺激

実験刺激として、映画「野良犬」(黒澤明, 1949) から 400 秒間を抜粋して使用した (TDV2683, 開始より 1 時間 53 分 18 秒経過時より 1 時間 59 分 58 秒経過後まで)。図 4-1 に抜粋した場面のストーリーのブロック図を示す。これを視聴覚刺激 AV とする。AV には音楽が付加された場面が 4 箇所あり、これらの箇所をそれぞれ AV-①, AV-②, AV-③, AV-④とする。そのうちの AV-②, AV-③, AV-④は音と画の対位法が用いられたとされる箇所である (西村, 1998)。AV-②, AV-③は、刑事と犯人の緊張した格闘劇の合間に、のんびりとしたピアノの曲 (Sonatinen Op.20-1, Kuhlau) が流れる場面である。AV-④は逃走した上に捕えられた犯人と刑事が 2 人で倒れこむ場面に、子どもたちの歌声 (蝶々, 日本唱歌) が聞こえてくるという場面である。AV-①には、対位法が用いられておらず、通常の映像作品のように音楽が付けられたと思われる。

また、制作者の意図が反映されない視聴覚刺激として、視聴覚刺激 AV の音楽が付加されている場面と同じ箇所に、著者が音楽を付加して視聴覚刺激を作った。ただし、元の音楽とは逆の印象が生ずると予想される音楽を付加した。この視聴覚刺激を A'V とする。A'V の音楽が付加された箇所を、それぞれ A'V-①, A'V-②, A'V-③, A'V-④とする。A'V-①は黒澤監督の対位法をまねて、A'V-②, A'V-③, A'V-④は映像の印象に合うような音楽をつけた。A'V-①の部分には「天国と地獄」(黒澤明, 1963) からの抜粋 (Die Forelle, Op.114, Schubert), A'V-②, A'V-③の部分には「野良犬」(黒澤明, 1949) の①の部分 (AV-①) の音楽, A'V-④の部分には「とらわれた時間」(秋山裕和) の一部を組み

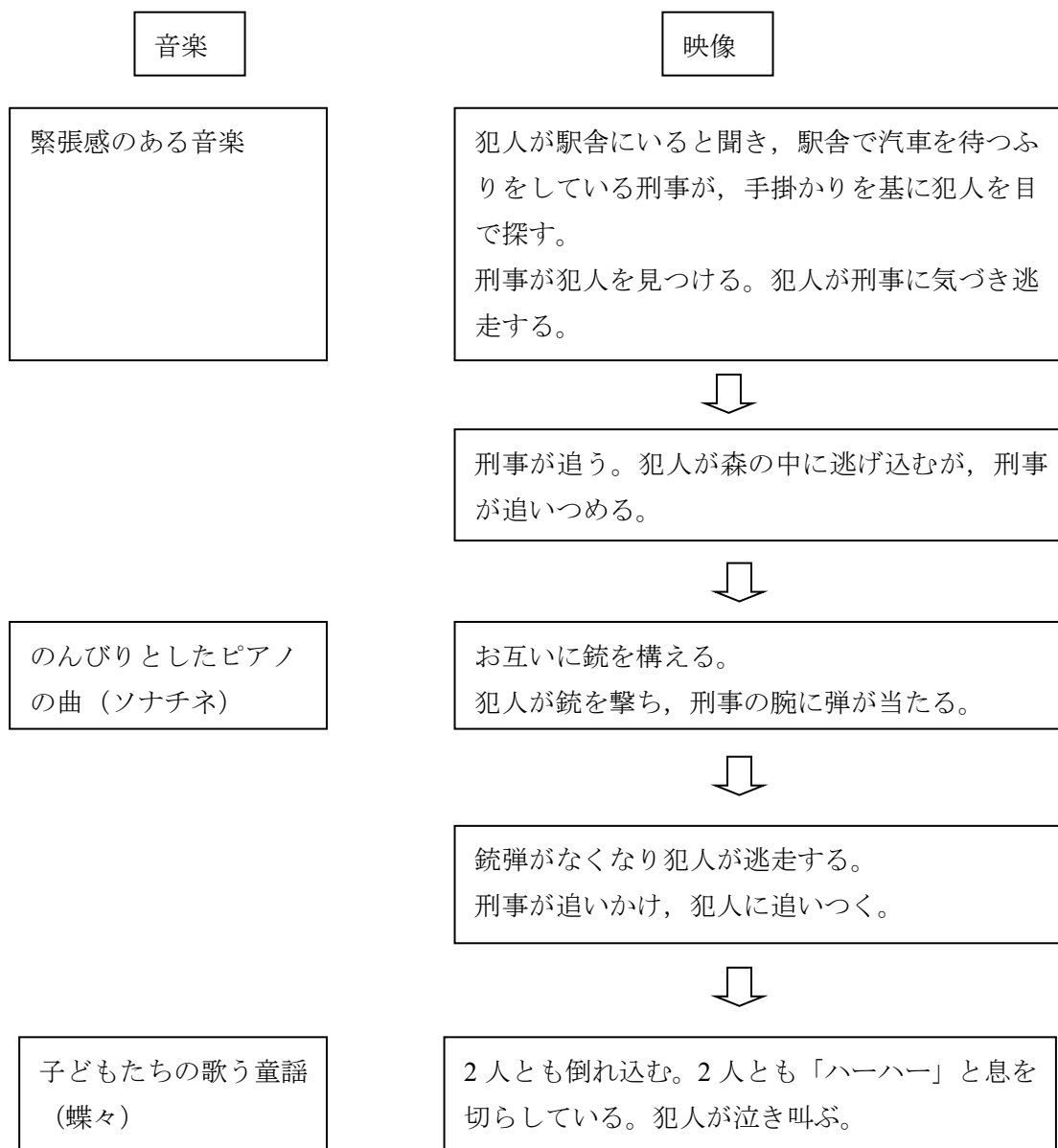


図 4-1 映画「野良犬」より抜粋した場面のストーリー

合わせている。ただし，もとの作品の音楽以外の音は残すように編集した。AVには，冒頭部（図 4-2 (a) の①の部分）に 1 秒程度の台詞が 2 ヶ所，終盤付近（図 4-2 (b) の④の部分）には「ハーハー」という 15 秒程度の息づかい含まれるが，それらはできるだけそのまま活かされるように A'V を作成した。台詞の 1 ヶ所は音楽のとぎれ目にあるので，その部分を抜粋して用いた。息づかいは，④の直前の音楽がない場面で同様の息づかいがあり，その部分を抜粋して活用した。実験者により，AV と A'V でその部分

に関しては、聴感上ほとんど差がないことが確認された。A'Vに含まれない台詞は、音楽中にかすかに聞こえる程度で意味も明瞭でなかったため、その有無がAVとA'Vの印象の差に影響を与えないと判断した。

これらAVとA'Vに対して、音楽と映像の調和感の連続評定実験を行った。刺激中には、音楽のない部分も含まれるが、その部分の評定はしないでもいいものとした。

#### 4.2.1.2 実験参加者

AVに対する調和感の連続測定実験の実験参加者は、21歳から27歳の九州大学の学生20名（男性14名、女性6名、平均年齢23.1歳）であった。A'Vに対する調和感の連続測定実験の実験参加者は、21歳から24歳の九州大学の学生20名（男性13名、女性7名、平均年齢22.5歳）であった。各実験参加者は、AVを用いた実験かA'Vを用いた実験のいずれかにしか参加していない。実験参加者にはこの作品を視聴したことがあるかを確認したが、この作品を視聴したものはいなかった。

#### 4.2.1.3 実験装置

実験装置は、2章実験1(2.2.2)で使用したものと同様であった。

#### 4.2.1.4 実験手続き

実験手続きは、2章実験1(2.2.3)と同様であった。刺激の呈示レベル（等価騒音レベル）は、約64dBとした。

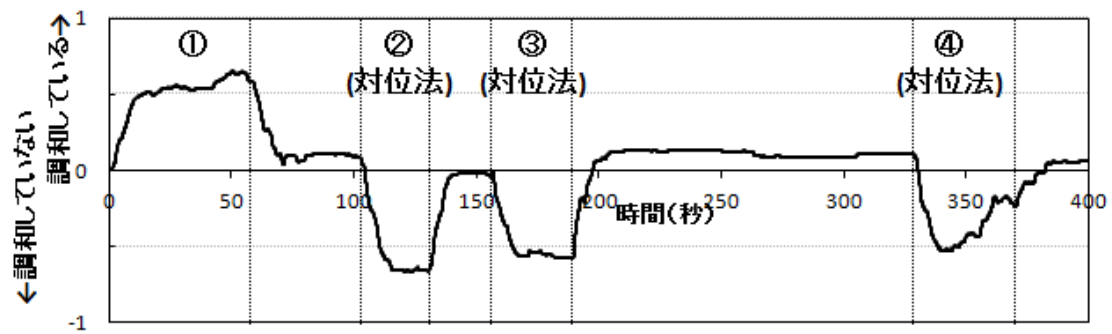
### 4.2.2 結果と考察

図4-2に、視聴覚刺激AV、A'Vにおける、時々刻々の音楽と映像の調和度を示す（全実験参加者の平均）。視聴覚刺激中、音楽がつけられた4つの部分を、図中に①から④で示す（本研究では、これらの部分を検討対象とする）。

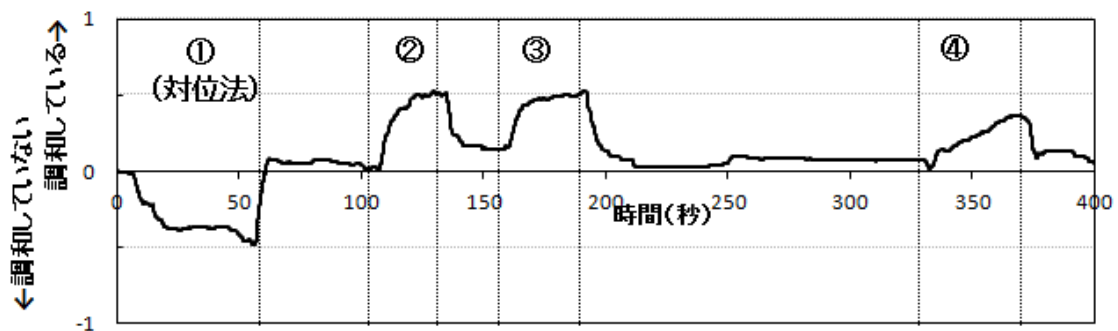
AV-①、A'V-②、A'V-③、A'V-④では、全般的に音楽と映像の調和度は高い。AV-①の部分の音楽と映像の調和度の平均値は0.50で、正の値であった。また、A'V-②、A'V-③、A'V-④での調和度の平均値は0.35、0.41、0.21で、正の値であった。AV-①はもともと音楽と映像の調和が図られたと思われる部分で、A'V-②、A'V-③、A'V-④は著者が映像の印象に合わせて音楽を組み合わせた部分である。AV-②、AV-③、AV-④、A'V-①では、全般的に音楽と映像の調和度は低い。AV-②、AV-③、AV-④での調和度の平均値は-0.52、-0.48、-0.33で、負の値であった。また、A'V-①の部分の音楽と映像の調和度の平均値は-0.31で、負の値であった。AV-②、AV-③、AV-④は、対位法が使われている部分で、A'V-①は、著者が対位法をまねて音楽を組み合わせた部分である。これらの結果より、黒澤監督が対位法を用いた部分では音楽と映像の調和感が低いことが示され、

著者が制作した視聴覚刺激は意図通りの刺激になっていたことが確かめられた。

対位法を用いた群 (AV-②, AV-③, AV-④, A'V-①) と対位法を用いていない群 (AV-①, A'V-②, A'V-③, A'V-④) の2群間の調和度の差を比較するため平均値の差の検定を行った。2群間の等分散性が仮定できなかつたため, Welch の検定を行った。その結果, 2群間に有意な差が見られた ( $t'(147.87) = 17.83, p < .001$ )。



(a) AV



(b) A'V

図 4-2 各視聴覚刺激における各時点での音楽と映像の調和度

## 4.3 実験 7：視聴覚刺激の音楽と映像の印象の連続記述選択実験

実験 7 では、実験 6 で用いた刺激の印象を検討するため、音刺激・映像刺激・視聴覚刺激に対して連続記述選択実験を行った。

### 4.3.1 実験方法

#### 4.3.1.1 実験刺激

視聴覚刺激 AV の「音」を A, 「映像」を V, A'V の「音」を A' とする。視聴覚刺激 AV, A'V, 音刺激 A, A', 映像刺激 V に対して、時々刻々の印象の評定実験を行った。

#### 4.3.1.2 実験参加者

AV, A, V に対する印象評定実験の実験参加者は、21 歳から 27 歳の九州大学の学生 20 名（男性 14 名，女性 6 名，平均年齢 23.1 歳）であった。A'V, A' に対する印象評定実験の実験参加者は、21 歳から 24 歳の九州大学の学生 20 名（男性 13 名，女性 7 名，平均年齢 22.5 歳）であった。いずれも実験 6 に参加した実験参加者（4.2.1.2）と同じである。

#### 4.3.1.3 実験装置

実験装置は、2 章実験 1（2.2.2）で使用したものと同様であった。

#### 4.3.1.4 実験手続き

印象評定に用いた形容詞は、音刺激、映像刺激の評定では「陽気な」「陰気な」「緊張感のある」「のんびりした」「迫力のある」「物足りない」「スピード感のある」「スピード感のない」（8 種類）であった。視聴覚刺激（映像＋音）の評定ではこれらに「ユニークな」「ありふれた」（2 種類）を加えた。視聴覚刺激（映像＋音）では、8 種類の形容詞での評定に加え、「ユニークな」「ありふれた」（2 種類）だけでの評定も行った。その他は 3 章 3 節の実験（3.3.1.4）と同様である。

### 4.3.2 結果と考察

図 4-3 から図 4-22 に結果を示す。連続記述選択法の実験結果において、比較的多く選択された形容詞のデータを用いて、調和感の高低の差が生ずる要因を検討する。①の部分では「緊張感のある」「のんびりした」、②③④では「陽気な」「陰気な」の形容詞を検討の対象とする。



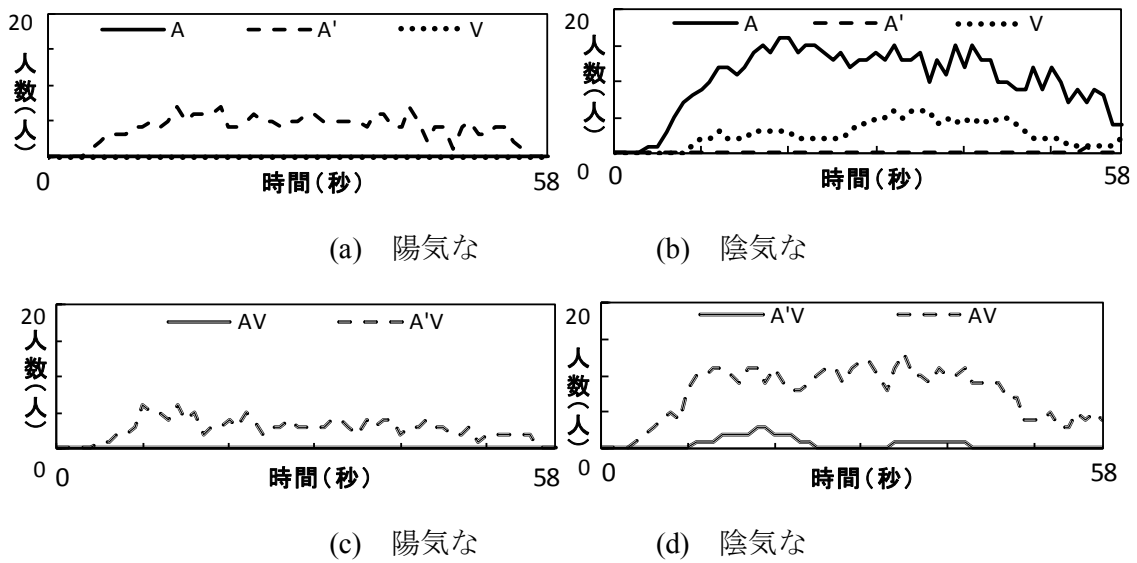


図 4-3 各刺激において①の時間帯の各時点で

「陽気な」「陰気な」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

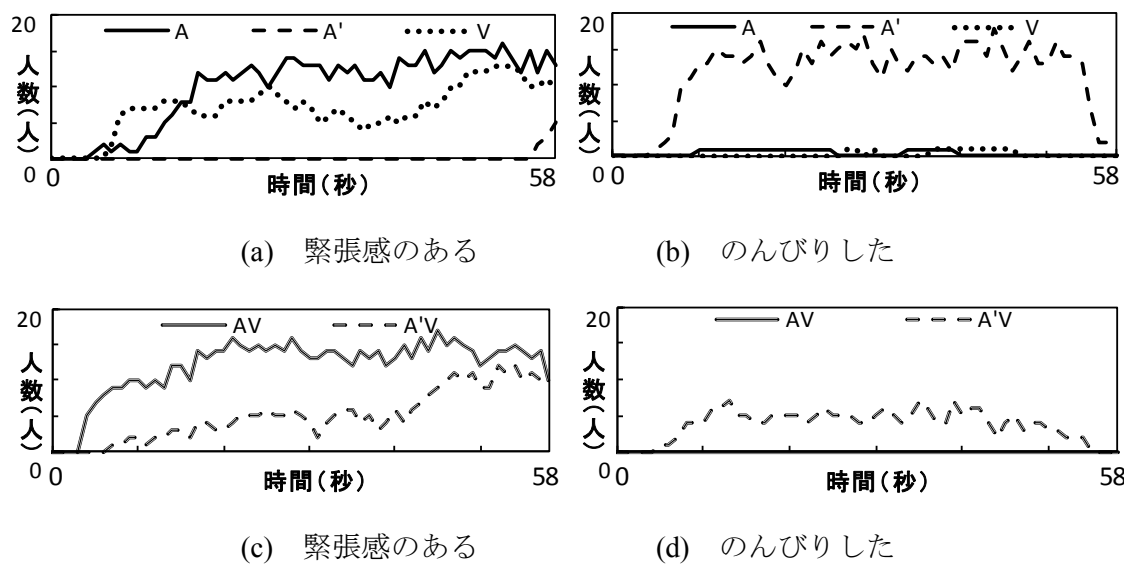


図 4-4 各刺激において①の時間帯の各時点で

「緊張感のある」「のんびりした」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

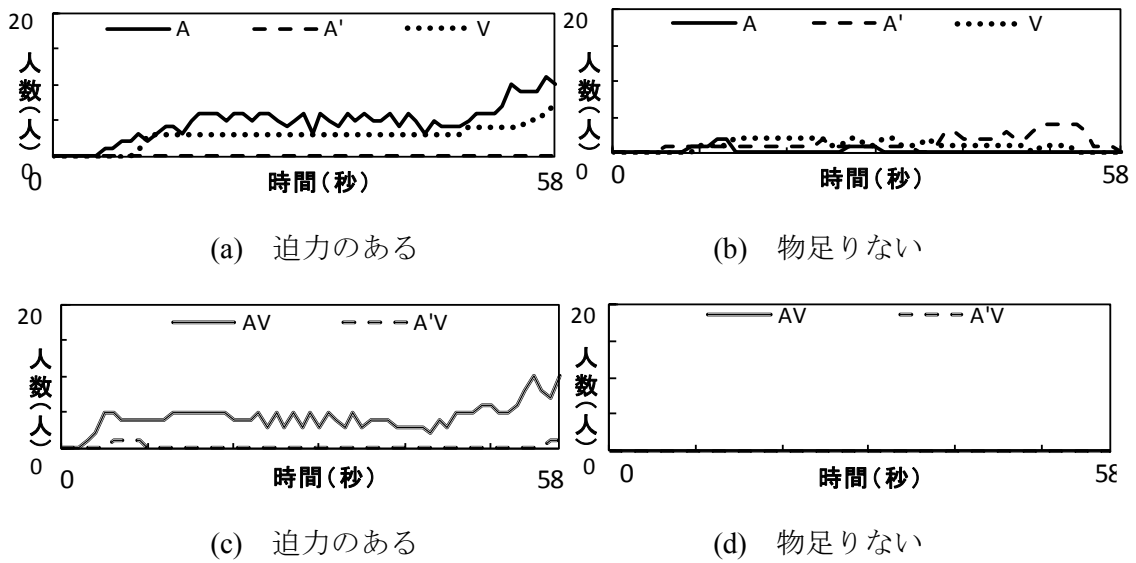


図 4-5 各刺激において①の時間帯の各時点で  
「迫力のある」「物足りない」を選択した人数  
(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

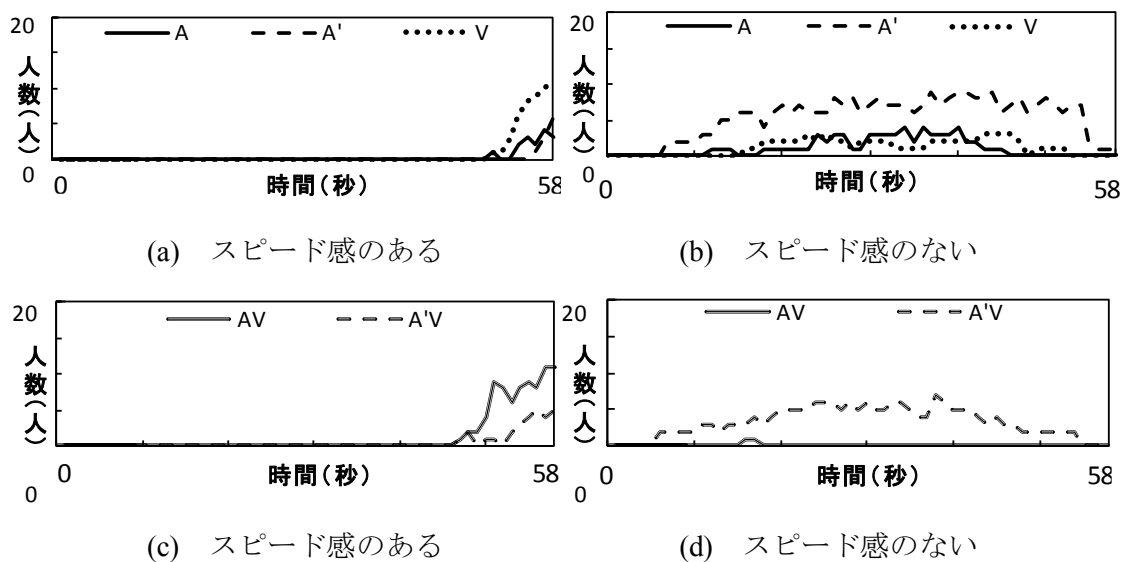
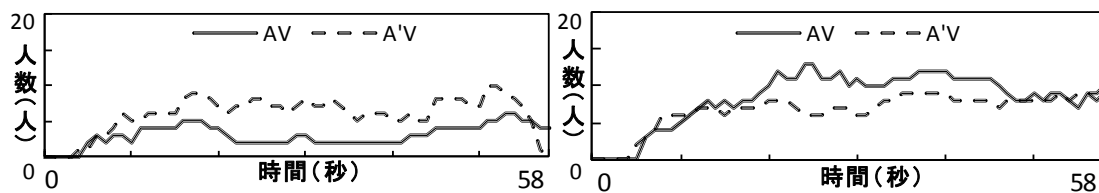


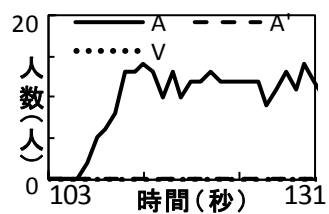
図 4-6 各刺激において①の時間帯の間の各時点で  
「スピード感のある」「スピード感のない」を選択した人数  
(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V



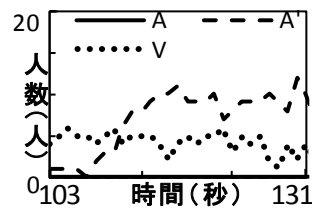
(a) ユニークな

(b) ありふれた

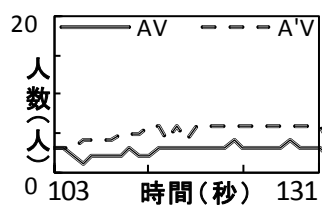
図 4-7 視聴覚刺激 AV, A'V において①の時間帯の間の各時点で  
「ユニークな」「ありふれた」を選択した人数



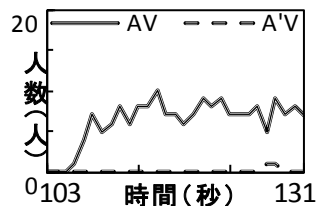
(a) 陽気な



(b) 陰気な



(c) 陽気な

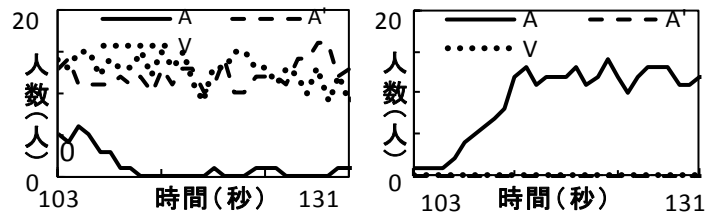


(d) 陰気な

図 4-8 各刺激において②の時間帯の各時点で

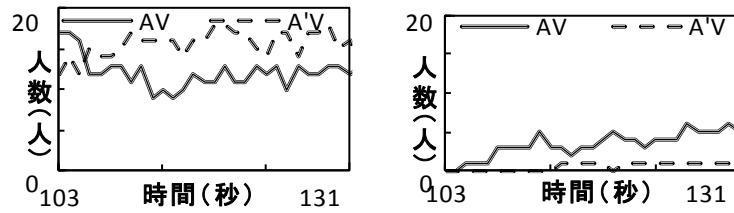
「陽気な」「陰気な」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V



(a) 緊張感のある

(b) のんびりした



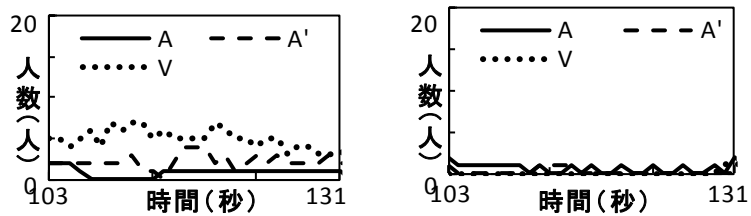
(c) 緊張感のある

(d) のんびりした

図 4-9 各刺激において②の時間帯の各時点で

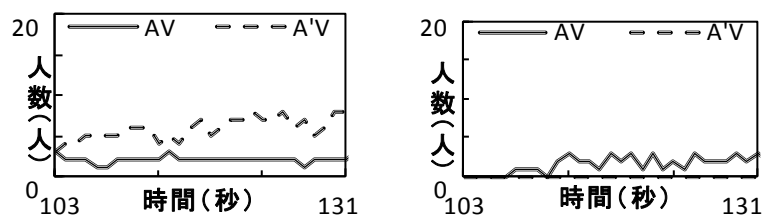
「陽気な」「陰気な」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V



(a) 迫力のある

(b) 物足りない



(c) 迫力のある

(d) 物足りない

図 4-10 各刺激において②の時間帯の各時点で

「迫力のある」「物足りない」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

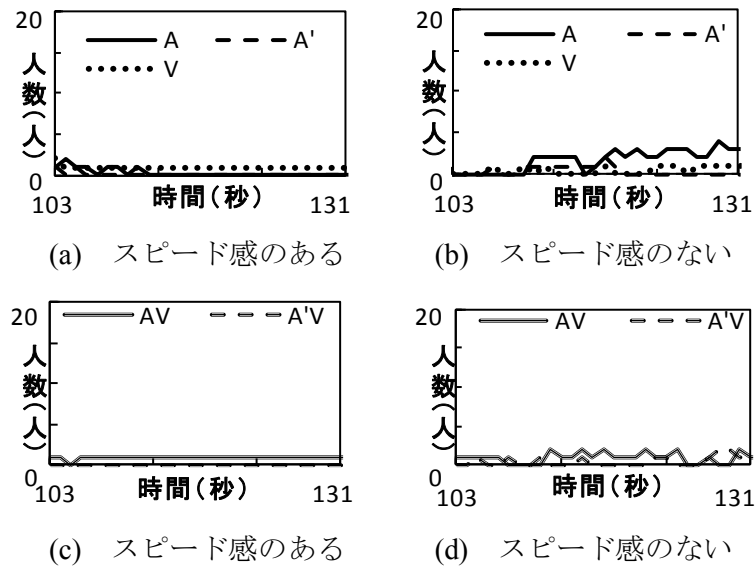


図 4-11 各刺激において②の時間帯の各時点で  
「スピード感のある」「スピード感のない」を選択した人数  
(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

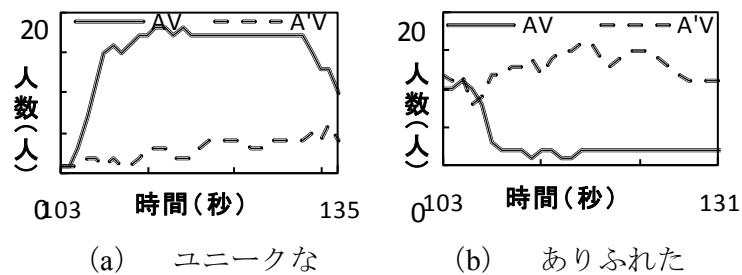


図 4-12 視聴覚刺激 AV, A'V において②の時間帯の間の各時点で  
「ユニークな」「ありふれた」を選択した人数

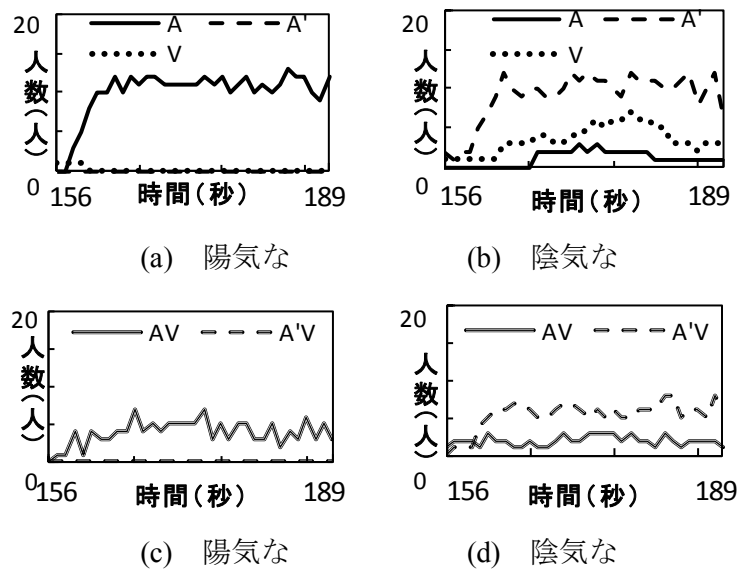


図 4-13 各刺激において③の時間帯の各時点で

「陽気な」「陰気な」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

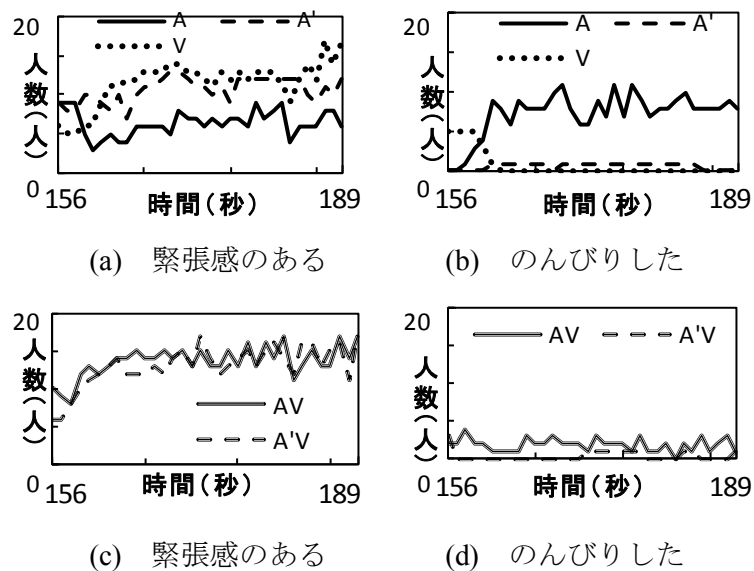


図 4-14 各刺激において③の時間帯の各時点で

「緊張感のある」「のんびりした」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

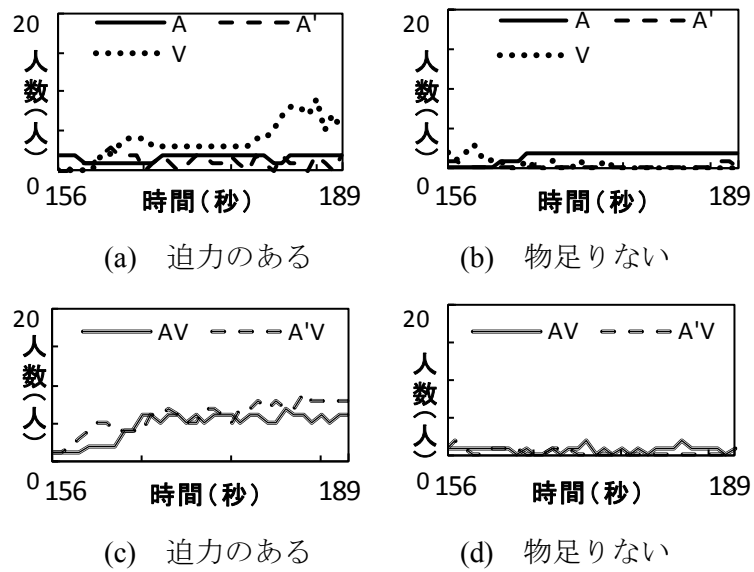


図 4-15 各刺激において③の時間帯の各時点で  
「迫力のある」「物足りない」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V    (下段) 視聴覚刺激 AV, A'V

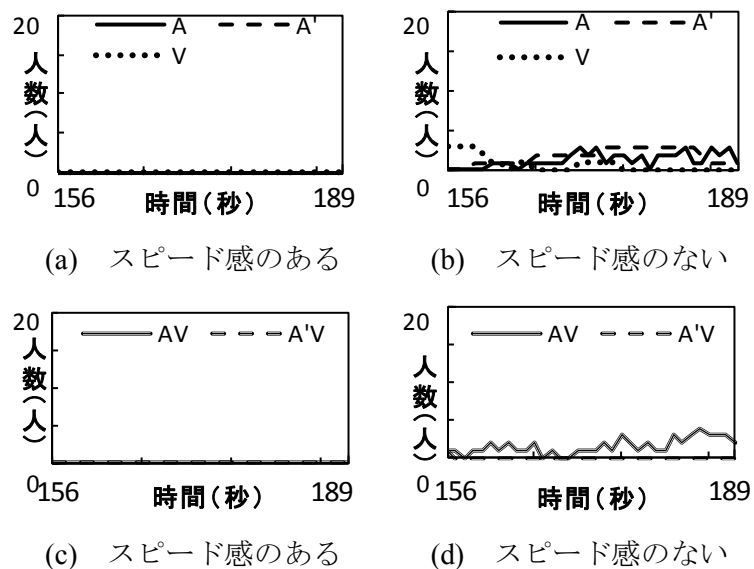


図 4-16 各刺激において③の時間帯の各時点で

「スピード感のある」「スピード感のない」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V    (下段) 視聴覚刺激 AV, A'V

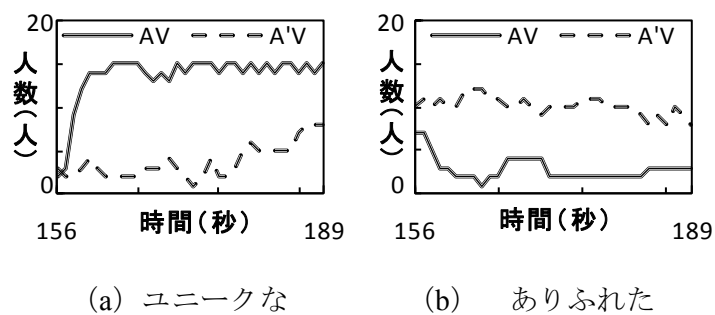


図 4-17 視聴覚刺激 AV, A'V において③の時間帯の各時点で  
「ユニークな」「ありふれた」を選択した人数

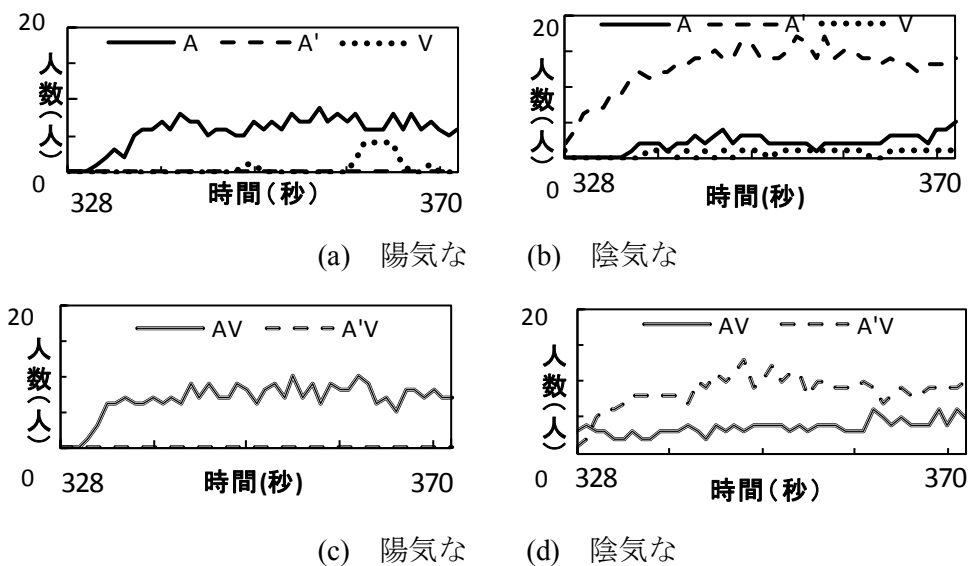


図 4-18 各刺激において④の時間帯の各時点で  
「陽気な」「陰気な」を選択した人数  
(上段) 音刺激 A, A', 映像刺激 V    (下段) 視聴覚刺激 AV, A'V



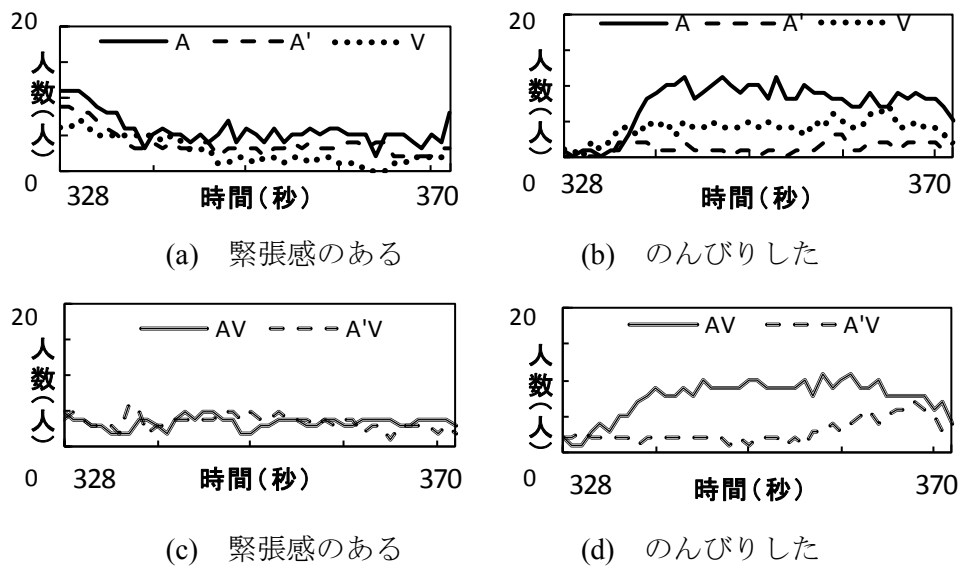


図 4-19 各刺激において④の時間帯の各時点で  
「緊張感のある」「のんびりした」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

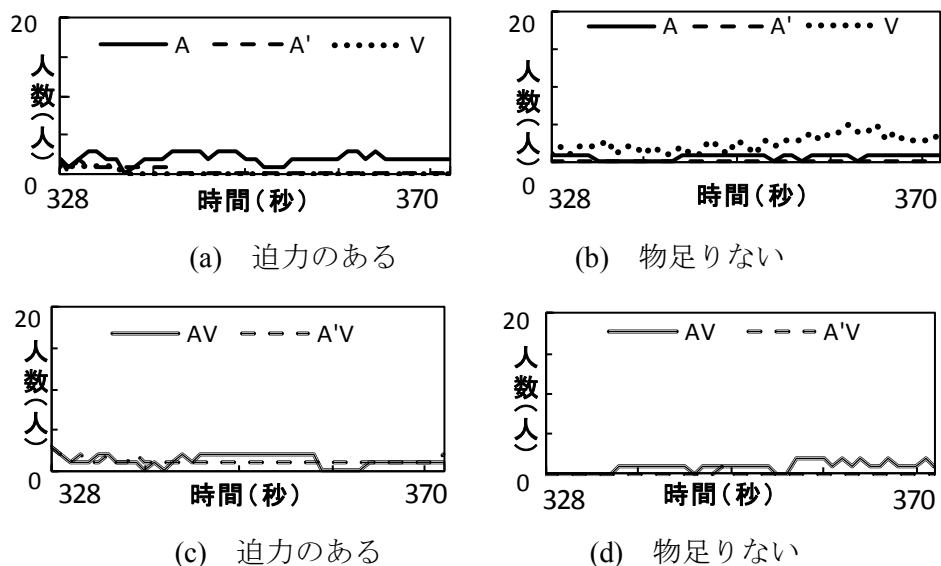


図 4-20 各刺激において④の時間帯の各時点で  
「迫力のある」「物足りない」を選択した人数

(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

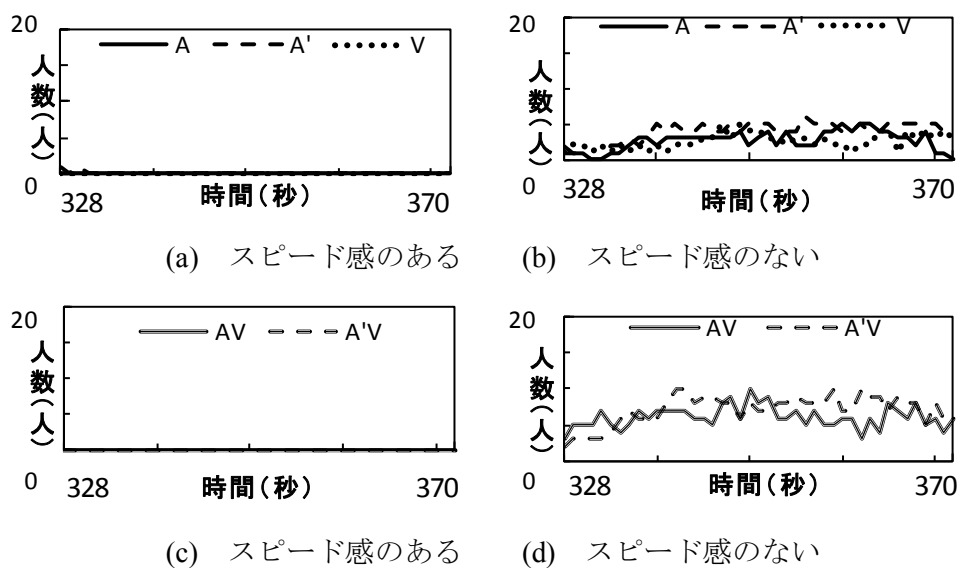


図 4-21 各刺激において④の時間帯の各時点で  
「スピード感のある」「スピード感のない」を選択した人数  
(上段) 音刺激 A, A', 映像刺激 V (下段) 視聴覚刺激 AV, A'V

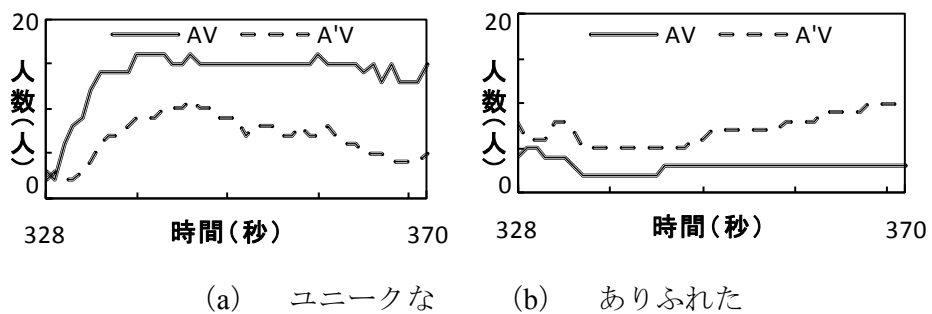


図 4-22 視聴覚刺激 AV, A'V において④の時間帯の間の各時点で  
「ユニークな」「ありふれた」を選択した人数

図 4-4 は、①の部分の音刺激、映像刺激、視聴覚刺激（映像刺激＋音刺激）に対して、連続記述選択法において「緊張感のある」「のんびりした」の形容詞が選択された人数を示す。刺激全般にわたって、A-①では「緊張感のある」の選択者が多く、A'-①ではその反対の意味を持つ「のんびりした」の選択者が多い。A'-①の結果は、著者の意図通りの刺激になっていることを示すものである。V-①においては、刺激全般にわたって「緊張感のある」の選択者が多い。

AV-①においては、A-①と V-①がいずれも「緊張感のある」との印象を持たれており、同じような印象の音楽と映像が組み合わせられ意味的調和が形成されていると考えられる。実際に、図 4-2 に示されているように、この部分では、音楽と映像の調和度も高い水準を保っている。図 4-4 によると、この視聴覚刺激においては、「緊張感のある」の選択者が多く、「のんびりした」の選択者はいない。A'V-①の視聴覚刺激においては、映像は「緊張感のある」印象であるにもかかわらず、音楽は逆の意味を持つ「のんびりした」印象で、意味的調和が形成されておらず、図 4-2 (b) に示されているように、調和度も低い状態である。この視聴覚刺激においては、「緊張感のある」の選択者と「のんびりした」の選択者が混在している。

図 4-8 は、②の部分の音刺激、映像刺激、視聴覚刺激に対して、「陰気な」「陽気な」の形容詞が選択された人数を示す。V-②においては、「陰気な」の選択者が多いのに対し、A-②ではその反対の意味を持つ「陽気な」の選択者が多い。A'-②では「陰気な」の選択が多く、A-②とは反対で V-②と同様の印象の音楽という著者の意図通りの刺激になっていることが示された。AV-②は、音と画の対位法が用いられた部分であり、制作者の意図通り音楽と映像の印象が異なるものとなっており意味的調和が形成されていない。また、図 4-2 (a) に示されているように、この部分では音楽と映像の調和度は低い水準である。図 4-8 によると、AV-②の視聴覚刺激においては、「陰気な」の選択者と「陽気な」の選択者が混在している。著者が音楽と映像の意味的調和を形成することを意図した A'V-②では、図 4-2 (b) に示すように、音楽と映像の調和度は高い水準を保っている。図 4-8 に示すように、A'V-②の視聴覚刺激の印象としては、「陰気な」の選択者が多く、「陽気な」の選択者はまれである。

図 4-13 は、③の部分の音刺激、映像刺激、視聴覚刺激に対して、「陰気な」「陽気な」の形容詞が選択された人数を示す。AV-③は、AV-②の続きの場面で音楽も AV-③と同様に音と画の対位法が用いられた部分である。図 4-8 で得られた傾向と同様に、V-③においては「陰気な」の選択者が多いのに対し、A-③ではその反対の「陽気な」の選択者が

多く、制作者の意図通り音楽と映像の印象が異なるものとなっている。A'-③では「陰気な」の選択者が多い。図 4-2 (a) に示されているように、AV-③における音楽と映像の調和度は低い。図 4-13 によると、AV-③の視聴覚刺激においても、「陰気な」と「陽気な」の選択者が混在している。著者が音楽と映像の意味的調和を形成することを意図した A'V-③では、図 4-2 (b) に示すように音楽と映像の調和度は高く、視聴覚刺激に対しては、図 4-13 に示すように「陰気な」の選択者が多く「陽気な」の選択者はまれである。

図 4-18 は、④の部分の音刺激、映像刺激、視聴覚刺激に対して、「陰気な」「陽気な」の形容詞が選択された人数を示す。この部分の A-④は子どもたちの「蝶々」の合唱で、「陽気な」の選択者が多く、A'-④では「陰気な」の選択が多い。V-④では、選択数は少ないが全般を通して「陰気な」の選択があり、一部で「陽気な」の選択もある。AV-④の場面は、犯人が逮捕された場面であり動きがないため V-④の印象が明確ではないが、やはり音と画の対位法が用いられた部分では、制作者の意図通りに、音楽と映像の印象が異なるものとなっている。図 4-2 (a) に示されているように、この部分の音楽と映像の調和度は低い。図 4-18 によると、AV-④の視聴覚刺激においても、「陰気な」と「陽気な」の選択者が混在している。著者が音楽と映像の意味的調和を形成することを意図した A'V-④では、図 4-2 (b) に示すように音楽と映像の調和度は高く、視聴覚刺激に対しては、図 4-15 に示すように「陰気な」の選択者が多く「陽気な」の選択者はいない。

3.3.2 と同様の方法で、連続記述選択法の結果から AV, A'V の 1 秒毎の印象の一致度を算出した。図 4-23 に AV, A'V の各時間における一致度を示す。各秒の一致度を平均すると、AV-①の音楽と映像の平均一致度は 11.22 で正の値であった。AV-②, AV-③の平均一致度はそれぞれ-10.79, -4.00 で負の値であり、AV-④は 2.07 と正の値であるものの小さい値であった。A'V-①の部分の音楽と映像の平均一致度は-8.08 で負の値であった。A'V-②, A'V-③, A'V-④の平均一致度はそれぞれ 16.38, 13.12, 2.26 で正の値であった。

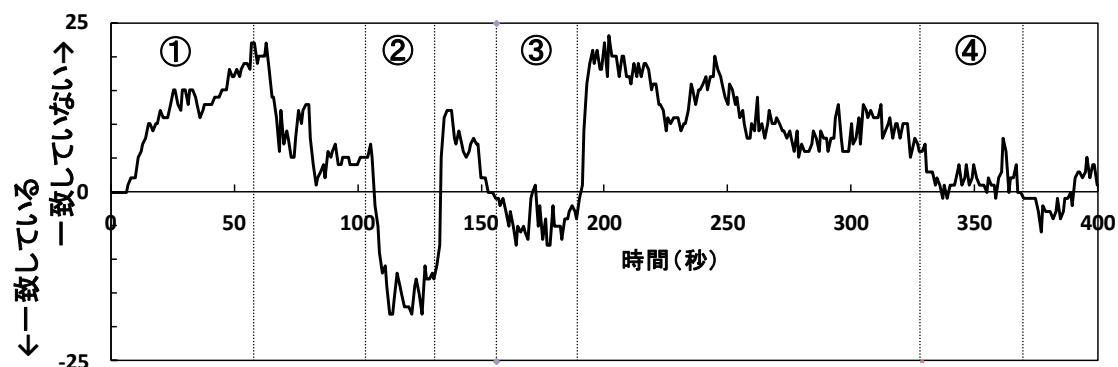
対位法を用いた群 (A'V-①, AV-②, AV-③, AV-④) と対位法を用いていない群 (AV-①, A'V-②, A'V-③, A'V-④) の 2 群間の一致度の差を比較するため t 検定を行った。その結果、2 群間に有意な差が見られた ( $t(328) = 20.439, p < .001$ )。

以上のように、通常のように音楽が使われる場面では、類似した印象の音楽と映像が組み合わせられていることが、印象の一致度という定量的な数値からも明らかになった。

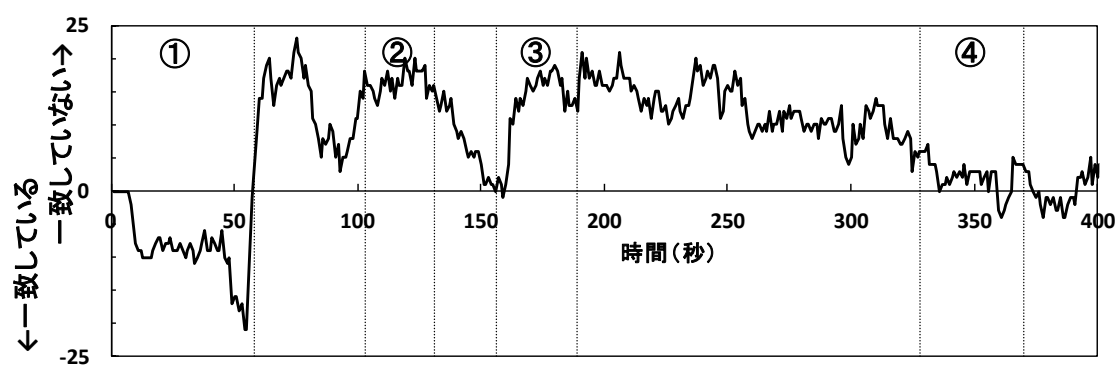
このような場面では、音楽と映像の「意味的調和」が形成されており、音楽と映像の調和感が生じている。これは、3章の Disney の作品を用いた実験の結果と一致するものである。

意味的調和が形成された状況では、視聴覚刺激の印象は統一した印象になる傾向がみられた。一方、対位法が用いられた場面では、異なる印象の音楽と映像が組み合わせられており、意味的調和が形成されず、音楽と映像の調和感は低かった。さらに、視聴覚刺激の印象は、ある印象に統一されず、反対の印象が混在していた。視聴覚刺激の印象は、音の印象の影響を受けることも映像の印象を受けることもありえるのである。この傾向は、西村が指摘する（西村，1998）、音と画の対位法により音と映像を独立した要素として機能させるという黒澤監督の意図が実現されていることを示すものである。制作者の意図を反映させず、著者が対位法をまねて作った視聴覚刺激でも、同様の傾向が得られている。

また、対位法が使われた視聴覚刺激では、「ユニークな」の形容詞も多く選択されていた。AV-①では「ありふれた」が多く（図 4-7）、AV-②では「ユニークな」が多く選択されている（図 4-12）。通常とは異なった独特な音楽と映像の関係のものに「ユニークな」印象が持たれているようである。



(a) AV



(b) A'V

図 4-23 音楽と映像の印象の一致度

## 4.4 実験 8：視聴覚刺激全体の印象評定

実験 8 では、音と画の対位法が視聴覚刺激全体の印象にもたらす効果を検討した。実験 6 で使用した刺激と新たに作成した 2 つの刺激を用いて、刺激視聴後の印象を評定させる実験を行った。

### 4.4.1 実験方法

#### 4.4.1.1 実験刺激

4.2 の実験で用いた AV, A'V の「音楽」と「映像」を用いて、映像との調和度が高い音楽だけを組み合わせた AV-high と、調和度が低い音楽だけを組み合わせた AV-low を作成した。この AV-high, AV-low と、実験 6 で用いた AV, A'V の 4 つを実験 8 の刺激とする。

#### 4.4.1.2 実験参加者

実験参加者は九州大学の学生 20 名（男性 10 名，女性 10 名）であった。年齢は 21 歳から 27 歳（平均年齢 23.3 歳）であった。このうち，2 章の刺激 AV の実験に参加した者が 12 名，A'V の実験に参加した者が 5 名で，残りの 3 名は以前の実験には参加していない。ただし，評定結果に過去の実験への参加の影響は見られなかった。

#### 4.4.1.3 実験装置

実験装置は，2 章実験 1（2.2.2）で使用したものと同様であった。

#### 4.4.1.4 実験手続き

実験参加者には，1 つの刺激の視聴が終わった後，視聴覚刺激の印象に関わる「まとものある—まとものない」「違和感のない—違和感のある」「面白い—面白くない」「記憶に残る—記憶に残らない」「ユニークな—ありふれた」「変化のある—単調な」「葛藤を感じる—葛藤を感じない」「含みのある—含みのない」「良い—悪い」「緊張感のある—のんびりした」「迫力のある—ものたりない」「スピード感のある—スピード感のない」「臨場感のある—臨場感のない」「陽気な—陰気な」の形容詞対で構成した 14 個の 7 段階評定尺度（0 から 6）を用いて，印象を評定させた。さらに，すべての刺激の視聴が終わった後に，4 つの刺激を「良い」と思う順に順位付けをさせた。

### 4.4.2 結果と考察

形容詞対ごとに，刺激を要因として一元配置の分散分析を行った結果，「まとものある—まとものない（ $F(3, 57) = 4.82, p < .05$ ）」「違和感のない—違和感のある（ $F(3, 57) = 4.06, p < .05$ ）」「面白い—面白くない（ $F(3, 57) = 4.61, p < .05$ ）」「記憶に残る—記憶に残らない（ $F(3, 57) = 9.44, p < .05$ ）」「ユニークな—ありふれた（ $F(3, 57) = 18.4, p < .05$ ）」「変化のある—単調な（ $F(3, 57) = 10.2, p < .05$ ）」「葛藤を感じる—葛藤を感じない（ $F(3, 57) = 5.43, p < .05$ ）」「含みのある—含みのない（ $F(3, 57) = 9.40, p < .05$ ）」「良い—悪い（ $F(3, 57) = 3.93, p < .05$ ）」の 9 尺度で刺激の主効果が有意であることが確認された。これらの形容詞対の評定値に対し，多重比較（Bonferroni 法）を行った。図 4-24 に，各刺激の平均評定値を示す。

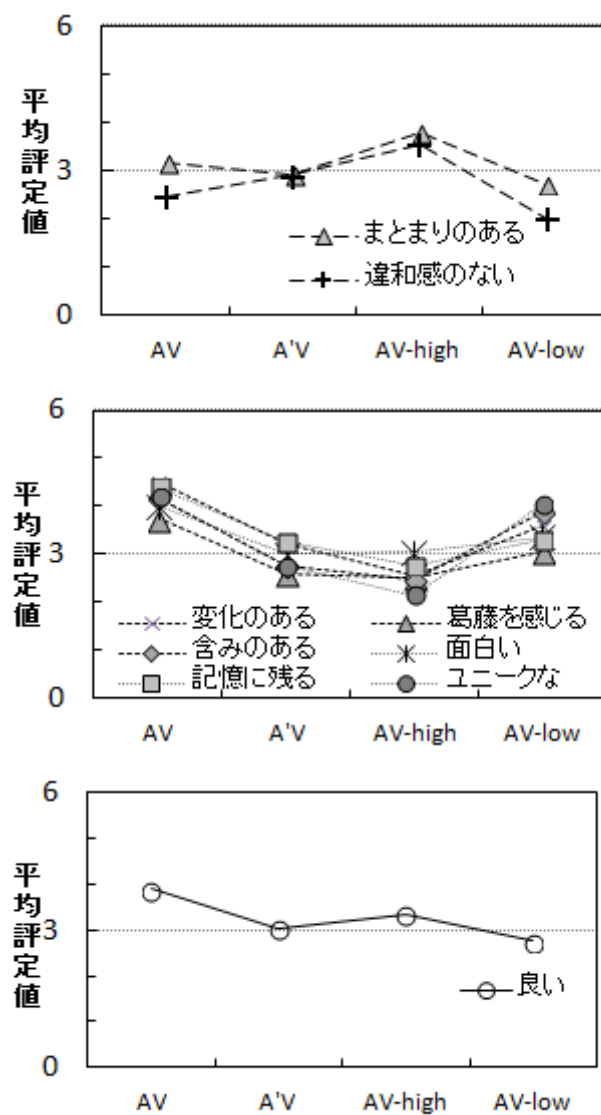


図 4-24 実験 8 における各視聴覚刺激における平均評定値



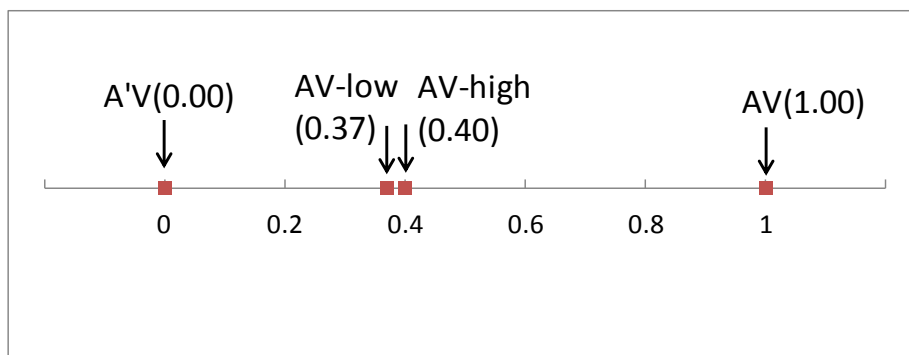


図 4-25 実験 8 における各視聴覚刺激における「良さ」の尺度値

視聴覚刺激の総合的な評価を表す「良い—悪い」以外の 8 つの尺度をクラスター分析（ウォード法）によって分類したところ、「まとまりのある」「違和感のない」のグループと「面白い」「記憶に残る」「ユニークな」「変化のある」「葛藤を感じる」「含みのある」のグループに 2 分できることが示された。

「まとまりのある」「違和感のない」といった「まとまりのよさ」を表わす尺度においては、AV-high の評定値が、対位法を用いた部分を含む AV, A'V, AV-low を上回り、最も高い。AV-high と AV-low の間には、これらの形容詞の評定値に有意差が認められた（「まとまりのある」では  $p < .1$ , 「違和感のない」では  $p < .05$ ）。

「面白い」「記憶に残る」「ユニークな」「変化のある」「葛藤を感じる」「含みのある」といった「ユニークさ」を表す尺度では、AV, A'V, AV-low の評定値が高く、オリジナルの作品の一部である AV の評定値が最も高い。一方、AV-high に対するこれらの形容詞の評定値は、最も低い。これらの尺度においては、AV と AV-high の間で有意差が認められた（いずれも  $p < .05$ ）。

以上の傾向から、AV-high のように、対位法を用いず音楽と映像をの印象を類似させた視聴覚刺激は、まとまりのよい印象となることがわかる。一方、AV のように対位法が用いられた場面を含む視聴覚刺激は、複雑でユニークな印象となる。AV-high は、音楽が使われたすべての場面において音楽と映像の意味的調和が成立しており、調和度もすべての場面において高い値である。そのため、AV-high は、違和感がなく、まとまりがあるように受け取られるのだろう。AV, A'V, AV-low は、音楽と映像の意味的調和が成立していない場面が含まれている。意味的調和が成立していない場面があると、まとまりはよくないが、面白さや、複雑さの印象が強くなると考えられる。

総合的な評価を表す「良さ」の評定値が最も高いのは、音楽と映像の組み合わせがオリジナル作品と同じである AV であった。次いで、良さの評価は、AV-high, A'V, AV-low と続く。AV と AV-low の間でのみ、評定値に有意差が認められた ( $p<.05$ )。また、4つの刺激を「良い」と思う順に順位付けたデータより、4つの刺激の「良さ」の順位を距離尺度へ変換したところ(武藤, 1982)、刺激の「良さ」は、AV (1.00), AV-high (0.40), AV-low (0.37), A'V (0.00) の順で高かった(図 4-25)。いずれの場合も、対位法が用いられた AV の評価が最も高く、音楽と映像の調和度が全ての場面において高い AV-high は、2 番目に良い評価であった。ただし、図 4-24 の AV と AV-high の間には有意差はない。AV-high のように視聴覚刺激の音楽と映像の意味的調和を成立させると、総合的な評価は良くなるが、AV のように対位法を用い視聴覚刺激の音楽と映像の意味的調和を成立させなくても、意味的調和が成立した刺激と同程度の評価を保つかより評価が高まる場合があると考えられる。

A'V, AV-low でも、対位法が用いられていたが、総合的な評価は低い。AV では、対位法の場面の画面中にピアノ演奏風景や合唱する子供たち(音源)が描写されているが、A'V, AV-low では、対位法の場面の画面中に音源の映像的描写がない。黒澤監督も、音と画の対位法について「(映像が描く世界において) 現実に聞こえてくる音でないと、効果は出ない」と主張している(西村, 1998)。これを受けて、西村は「音と画の対位法は、ストーリーの中で必然性を持たせることによって初めて成立する。心理的な不自然さを感じさせないためには、状況をきちんと説明しておかねばならない」と述べている(西村, 1998)。音楽の音源が画面上に現れることで、音楽を(映像の世界の)現実に聞こえてくる音にすることができる。そのため、音楽の音源が画面上に特定できるかどうか、作品の良さに関わっていると推測される。

## 4.5 実験 9：音楽の音源が画面に映っていない視聴覚刺激

### 全体の印象評定

実験 8 では、対位法が用いられた刺激の中でも、音楽の音源が映像の画面に映っている刺激の評価が高いという結果が得られた。そこで、実験 9 では、音楽の音源が画面に映っていることの効果を検証するため、対位法の場面の映像に音楽の音源が映っていない条件で印象評定実験を行った。

## 4.5.1 実験方法

### 4.5.1.1 実験刺激

前節の実験で用いた AV, A'V, AV-high, AV-low の映像を加工し、音と画の対位法が用いられている場面②③④で流れる音楽の音源を認識できないようにした。②③の場面では、画面中のピアノの音の音源である家全体にモザイク処理を施し、家があることを認識できないようにした。④の場面では、合唱の声の音源である歩いている子供たちの列全体にモザイク処理を施して、何が動いているのかわからないようにした。無音でピアノだけが映っている場面は削除した。映像に付加された音楽は前節の実験で用いたものと同様である。これらを AV-B, A'V-B, AV-high-B, AV-low-B と呼び、本実験の刺激とした。

### 4.5.1.2 実験装置

実験装置は、2章実験1(2.2.2)で使用したものと同様であった。

### 4.5.1.3 実験手続き

実験手続きは、実験8(4.4.1.3)と同様であった。

### 4.5.1.4 実験参加者

実験参加者は、九州大学の学生8名(男性5名,女性3名,平均年齢25歳)であった。本実験の参加者には、本章の他の実験に参加した者はいなかった。また、この作品を視聴したことがあるかを確認したが、この作品を視聴したものはなかった。

## 4.5.2 結果と考察

視聴覚刺激の総合的な評価を表す「良い—悪い」以外の8つの尺度をクラスター分析(ウォード法)によって分類したところ、3章の実験と同様に、「まとまりのある」「違和感のない」のグループと「面白い」「記憶に残る」「ユニークな」「変化のある」「葛藤を感じる」「含みのある」のグループに2分できることが示された。図4-26に、各尺度に対する各刺激の平均評定値を示す。

「良さ」の評定値は、AV-high-Bが最も高く、AV-B, A'V-B, AV-low-Bと続く。4つの刺激を「良い」と思う順に順位付けたデータより、4つの刺激の「良さ」の順位を距離尺度へ変換したデータでは(武藤, 1982), AV-high-B(1.00), A'V-B(0.71), AV-B(0.27), AV-low-B(0.00)の順番で「良さ」が高かった(図4-27)。いずれにおいてもAV-highの評価が最も高かった。

「まとまりのある」「違和感のない」といった形容詞の評定値は、AV-high-B, A'V-B,

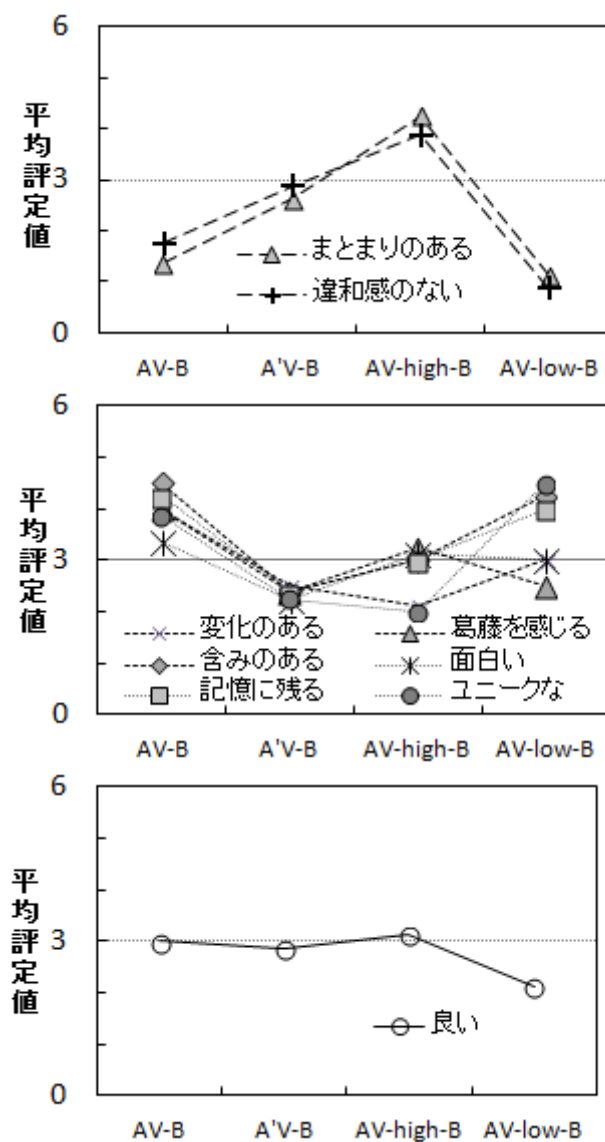


図 4-26 実験 9 における各視聴覚刺激における平均評定値

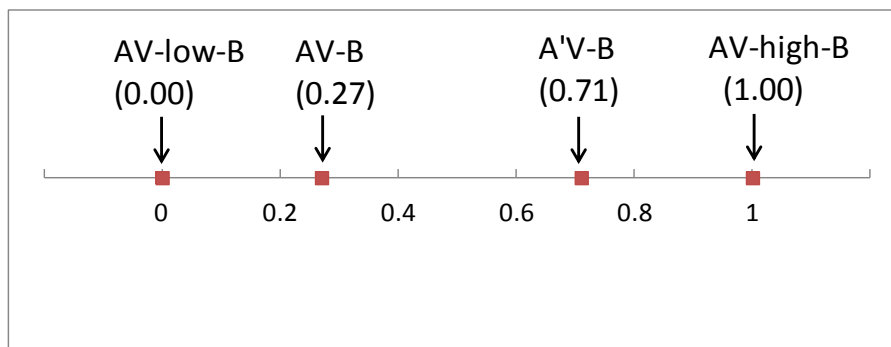


図 4-27 実験 9 における各視聴覚刺激における「良さ」の尺度値

AV-B, AV-low-B の順で高かった。多重比較 (Bonferroni 法) を行ったところ, 「まとまりのある」「違和感のない」の評定値は, AV-high-B と AV-B の間で, 有意差が認められ (それぞれ  $p<.1$ ,  $p<.05$ ), AV-high-B と AV-low-B の間にも有意差が認められた (いずれも  $p<.05$ )。

このように, 「まとまりのよさ」「良さ」の評価は, 刺激全般にわたって意味的調和がとれた AV-high-B で最も高く, 刺激全般にわたって意味的調和がとれていない AV-low-B で最も低い。意味的調和がとれた場面ととれていない場面が混在する AV-B, AV-B はその中間に位置する。

次に, 実験 8 と実験 9 で得られたデータを比較する。「面白い」「記憶に残る」「ユニークな」「変化のある」「葛藤を感じる」「含みのある」といった「ユニークさ」を表す尺度では, 各刺激とも実験 8 と実験 9 で得られた評定値との間に有意差はみられなかった。音源が見えるか見えないかという条件の違いは, 「ユニークさ」には影響を及ぼさないようである。

実験 9 で得られた対位法の場面で音源が画面上にない AV-B の「良さ」の評定値は, 実験 8 で得られた音源が画面中にある AV の評定値に比べると低い。t 検定を行ったところ, AV-B と AV の評定値に有意差があることが確認された ( $p<.1$ )。また, 「違和感のない」「まとまりのある」といった形容詞の評定値も, AV-B は, AV に比べて低かった。「まとまりのある」については, t 検定を行った結果, AV-B と AV の評定値に有意差が認められた ( $p<.05$ )。対位法が使われていない AV-high と AV-high-B の「良さ」「まとまりのある」「違和感のない」の評定値には, 有意差は見られなかった。このことから, AV-B の「良さ」や「まとまりのある」の評定値が低いのは, 映像を加工したことによる影響ではなく, 対位法の場面で画面中に音源が見えなかったためと考えられる。

AV-high, AV-high-B で用いられた音楽は, Cion の主張する音と映像の関係を表わす「3つの輪」の中の「オフの音」であり, 「non-diegetic sound」と位置づけられる。また, AV で用いられた音と画の対位法の場面の音楽は, 音源が画面中に存在している, または隣接しているため, Cion の主張する「インの音」「フレーム外の音」であり, 「diegetic sound」とみなせる。一方, AV-high-B で音楽の音源を認識できないように加工した場面においては, 音楽の音源は画面中に存在しておらず, 音楽は「オフの音」であり

「non-diegetic sound」となる。この「音と映像の関係」と本実験の結果から, 音楽が「オフの音」, 「non-diegetic sound」である場合には, 音楽と映像が調和していると視聴覚刺激の総合的な評価は高いが, 調和していないと視聴覚刺激の総合的な評価は低くなる。

音楽が「インの音」「フレーム外の音」であり、「diegetic sound」である場合には、音と映像が調和していなくても、視聴覚刺激の総合的な評価は高くなりうる。

以上の実験結果から、対位法の場面で画面中に音源がない場合、作品全体の違和感が増し、まとまりのなさがより際だつ印象となり、作品の総合的な評価が低くなることが示された。またこの場合、音楽と映像の意味的調和が成立している作品の評価は高く、意味的調和が成立していない場面があると作品の評価は低くなる。

## 4.6 本章のまとめ

本研究では、黒澤明監督の映画の一部を用いて、音と画の対位法の効果を心理実験により検証した。対位法が用いられていない、通常のように音楽が付けられたとされる場面では、音楽と映像の意味的調和が図られており、調和感は高い。一方、対位法が使われた場面では、反対の意味をもつ音楽と映像が組み合わせられており、音楽と映像の調和感は低かった。しかし、その場面はユニークな印象を与えていた。

音と画の対位法が含まれた視聴覚刺激全体は、複雑さや面白さの印象が強い。また、視聴覚刺激の総合的な良さの評価は、全般的に音楽と映像の意味的調和がとれた視聴覚刺激と同程度かそれ以上に高く、黒澤作品において対位法の手法が効果的に利用されていることがわかる。

しかし、対位法の場面で音楽の音源が画面上に存在しないと、作品の総合評価は下がることが示された。音楽と映像を調和させない音と画の対位法の効果を十分に出すためには、音楽の音源が画面上に存在することが必要であると考えられる。黒澤監督は、音楽の音源を画面上に登場させて、映像と調和しない音楽が存在する状況をつくり出したと思われる。そういった状況設定のもとで、映画がつくり出す物語の世界の中でそれほど違和感を与えず、音と画の対位法の効果が発揮されるのであろう。

## 第5章 全体的考察

### 5.1 本研究の概要

本研究では、視聴覚刺激の音（あるいは音楽）と映像の調和感の連続測定実験、音（あるいは音楽）と映像の印象の連続記述実験を実施して、映像作品において調和感が形成される過程を明らかにし、意味的調和の状況およびその効果を探ることを試みた。

一般に、音と映像の主観的な調和感には、音と映像のアクセントを一致させることによって得られる構造的調和と、音と映像のそれぞれが持つ印象を一致させることによって得られる意味的調和があると言われている。音と映像のアクセントの一致の状態を検知して得られる構造的調和は、比較的単純な視聴覚統合の処理によって得られるものと考えられる。これに対して、意味的調和を得るための視聴覚統合処理は、音の印象および映像の印象を感知し、さらに両方の印象を比較することにより印象の一致を捉えるものと考えられる。意味的調和の処理過程は構造的調和のものに比べ、より複雑な過程を含むことになり、その処理には構造的調和を感知するよりも時間がかかるものと考えられる。

2章で示したように、単純な条件下で構造的調和および意味的調和を作り出した視聴覚刺激を用いた音と映像の調和感の連続測定実験により、意味的調和が感知されたことによって得られる音と映像の調和感のピークに達する時間（12秒から20秒程度）が構造的調和による調和感のピークに達する時間（7秒から8秒程度）よりも長いことが示された。意味的調和を形成した視聴覚刺激の音の印象、映像の印象の連続測定実験により、音および映像の印象がピークに達する時間は、意味的調和に基づく調和感がピークに達する時間よりも短く、意味的調和の形成が音の印象、映像の印象を感知したのちになされるものとの考えが支持される結果となった。ただし、意味的調和を意図した視聴覚刺激を複数回視聴した場合、意味的調和のピークに到達する時間が2回目以降に短縮する傾向が見られた。構造的調和の場合にはそのような効果は見られず、視聴覚刺激を複数回視聴しても調和感のピークに達する時間に差はなかった。意味的調和の形成過程が比較的複雑な処理によるため、同じ視聴覚刺激を繰り返し視聴することにより処理が早まることが示唆される。構造的調和の場合、もともと処理が単純なため、処理が早ま

ることではないものと考えられる。

意味的調和はこれまで映画などで多用されてきた手法であると言われ、現在でも映像作品における定石的手法としてその効果が利用され続けている。3章で示したように、音楽に映像を組み合わせるという手法で制作された Disney のアニメーション「FANTASIA」「FANTASIA 2000」の一部を用いて、音楽と映像の調和感の連続測定実験を行い、音楽と映像の印象を連続記述選択実験で測定した結果、意味的調和が形成されている様子が示された。音楽と映像を入れ替えた視聴覚刺激では意味的調和が形成されないことから、映像作品の意味的調和は制作者の意図によるものと考えられる。また、いったん意味的調和が形成されると、作品中でその要因が弱まっても意味的調和の効果は持続し、調和感が保たれることが示された。意味的調和の形成過程に着目すると、調和感は徐々に上昇する傾向が見られ、意味的調和が形成されるにはある程度の時間を要することが確認された。また、意味的調和が形成されるのに必要と考えられる時間が経過した後も、さらに調和感が増し続けていたことから、作品内容の展開によっては、累積的に調和感が増することも示された。本研究に用いた視聴覚刺激では、後半部に迫力のある音楽表現、映像表現が用いられており、制作者が意図した盛り上げ効果が調和感のさらなる上昇を支えていたものと考えられる。

さらに、視聴覚刺激の始まりの場面よりも前から抜粋し前半分であった場面が後半部になるようにした視聴覚刺激や、後半部から始まる視聴覚刺激、前半と後半を入れ替えた視聴覚刺激を用いて、音楽と映像の調和感の形成過程を調和感の連続測定実験により検討した。元の視聴覚刺激の冒頭部で調和感が低いとされた部分も、同じ部分が中間部になった場合には高い調和感が得られていた。元の視聴覚刺激で高い調和感を得た部分でも、その部分を冒頭部にした場合には調和感が低下した。この結果からも、意味的調和の形成にはある程度の時間を要することが確認された。

映像作品では、音楽と映像の印象を類似させ意味的調和を図るような制作手法が一般的とされたが、時代の流れの中で、この原則にとらわれない作風を創造する制作者が現れた。映画監督として国際的な評価を得た黒澤明は、わざと音楽と映像の印象を一致させないことで独特の効果を狙う「音と画の対位法」と呼ばれる手法を多用した。4章で示したように、黒澤明監督作品「野良犬」から音と画の対位法が用いられている場面を抜粋し、音楽と映像の調和感の連続測定実験を行った結果、対位法を用いた部分では音楽と映像の調和感は低く、通常の場合では音楽と映像の調和感は高いことが確認された。さらに、音楽、映像、視聴覚刺激の印象を連続記述選択法で測定したところ、通常の場合



面では音楽と映像の印象は類似しており、対位法の用いられた場面では反対の印象が持たれていた。これらの結果より、対位法は意味的調和を崩す手法であることが示された。

ただし、黒澤作品から抜粋した視聴覚刺激全体の印象評定実験によると、対位法の部分を含むことによりその印象がユニークになり総合的な評価も高い。付加された音楽を入れ替えて意図的に対位法の部分を通常の意味的調和が形成されるようにすると、ある程度高い評価は得られるが、平凡な印象になる。

調和感の低い部分を含みながら高い評価が得られたのは、その部分の映像に音楽の音源が表現されていることにより、自然な感覚で違和感が感じられる音楽が流れる状況を作りだしていることによるものと考えられる。音楽の音源の部分が認識できないように映像を操作した視聴覚刺激に対する印象評定実験によると、対位法を含む視聴覚刺激ではユニークさは保つものの総合的な評価は下がり、音楽をすべて意味的調和が成立するように操作した視聴覚刺激のほうが単調な印象はあるものの総合的な評価は高くなった。音と映像を調和させない音と画の対位法の効果を十分に出すためには、音源が画面上に存在することが必要であることが示唆された。黒澤作品においては、対位法の手法が効果的に利用されていることがわかる。

## 5.2 音と映像の調和感の形成過程のモデル化

2章の実験1、実験2で得られた、単純な視聴覚刺激における音と映像の調和感の形成過程を表わす曲線(図2-8)は、指数関数型と見なすこともできる。この曲線に指数関数を当てはめたとすると、調和感の形成は次式のような近似式によって表わすことができ、調和感の形成過程の特性は時定数 $\tau$ で表現することができる。時定数とは、一般的にはあるシステムに信号が入力された時の、出力の応答の速さを特徴づける定数で、時間の次元を持つもののことを指す(日本音響学会 編, 1988)。本論文における時定数は、刺激が視覚、聴覚を通して入力され、視聴覚の統合を行う情報処理過程において音と映像の調和感が形成される速さを意味することとなる。

$$c(t) = \alpha(1 - e^{-t/\tau})$$

$\tau$  : 音と映像の調和度の時定数

$t$  : 刺激呈示時間

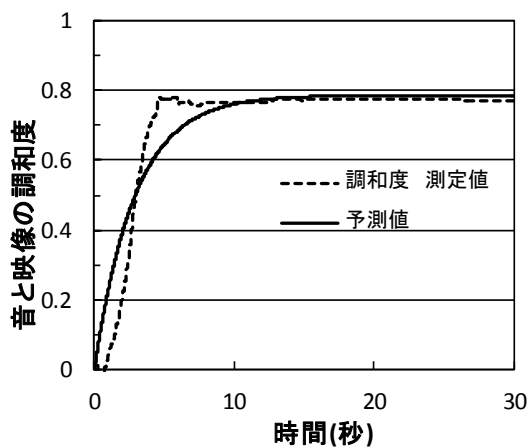
$c(t)$  : 刺激呈示時間  $t$  における音と映像の調和度

$\alpha$  : 定数

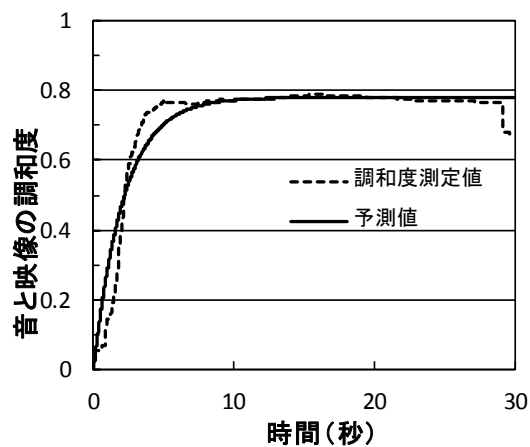
条件ごとの近似曲線への当てはめは、実験 1, 2 で行ったように実験参加者ごとに行い（条件による時定数の差を統計的に検討するため）、実験参加者ごとの実験 1, 実験 2 で得られた調和度 ( $c(t)$ ) と刺激呈示時間 ( $t$ ) から、それぞれの指数近似曲線を最小二乗法によって求め、時定数  $\tau$  を求めた。実験参加者ごとに求めた時定数の平均値を、各条件における時定数とした。検討対象の条件は、2 章と同様に「同期・1 回目」「同期・2 回目以降」「長三和音+笑顔・1 回目」「長三和音+笑顔・2 回目以降」「短三和音+泣き顔・1 回目」「短三和音+泣き顔・2 回目以降」である。なお、明らかに指数関数に当てはまらないと見なせる実験参加者のデータは除外した。これらのデータに近似曲線を当てはめると、いずれも時定数が 30 (秒) 以上となった。除外したのは、「長三和音+笑顔・1 回目」の場合、4 名、「長三和音+笑顔・2 回目以降」の場合、7 名、「短三和音+泣き顔・1 回目」の場合、6 名、「短三和音+泣き顔・2 回目以降」の場合、2 名であった。「同期・1 回目」「同期・2 回目以降」では、除外した実験参加者はいなかった。構造的調和による調和感の形成過程は、全実験参加者のデータとも、指数関数に従うものと考えられるが、意味的調和による調和感の形成過程は、指数関数に従うとは考えにくい実験参加者もいた。

図 5-1 に、実験 1, 2 における条件ごとの平均調和度の推移（図 3-2）とその指数近似曲線を示す。図 5-2 に、各条件における時定数の平均を示す。実験参加者毎に、調和度と近似値のピアソンの積率相関係数を算出し、各条件における平均を求めた。表 5-1 に各条件における調和度と近似値のピアソンの積率相関係数の平均を示す。各条件の相関係数の平均は 0.90 以上であり（すべて  $p < 0.01$ ）、近似曲線の当てはまりの良さが示された。各条件の平均時定数は、「同期・1 回目」で 2.93 秒、「同期・2 回目以降」で 2.60 秒、「長三和音+笑顔・1 回目」で 4.01 秒、「長三和音+笑顔・2 回目以降」で 3.44 秒、「短三和音+泣き顔・1 回目」で 9.45 秒、「短三和音+泣き顔・2 回目以降」で 5.79 秒となった。これらの値から、2 章で求めたピーク到達時間で得られた傾向と同様に、時定数は構造的調和の場合の方が意味的調和の場合よりも短いことが示された。また、同様に、時定数においても、意味的調和形成過程の場合には、2 回目以降の条件で 1 回目よりも短くなる傾向も確認できた。

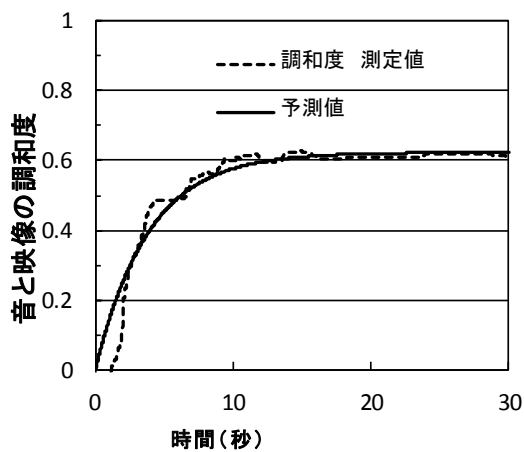
各条件間の値に統計的に有意な差があるかどうかを確認するため、各条件を要因として Non parametric 検定である Kruskal-Wallis の検定を行った。その結果、各条件間に有意な差が見られた ( $H(5) = 19.92, p < 0.01$ )。そこで、多重比較として、Mann-Whitney の U 検定を Bonferroni の不等式による修正を行ったうえで実施した。「短三和音+泣き顔・



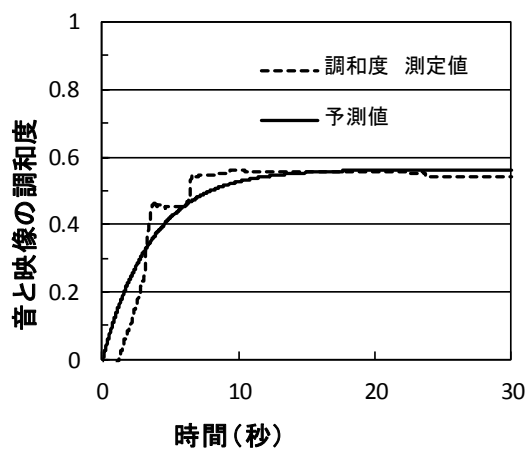
(a) 「同期・1回目」



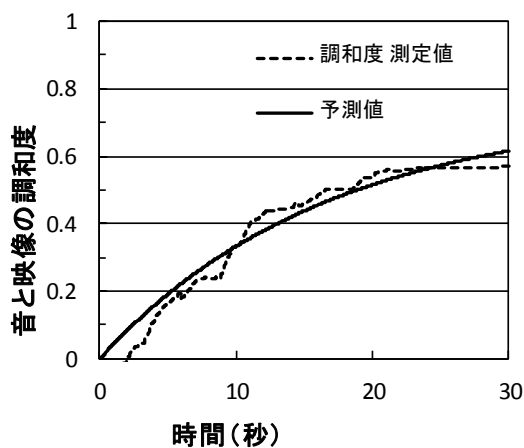
(b) 「同期・2回目以降」



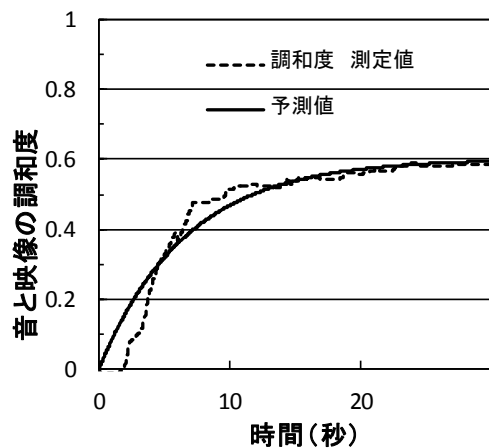
(c) 「長三和音+笑顔・1回目」



(d) 「長三和音+笑顔・2回目以降」



(e) 「短三和音+泣き顔・1回目」



(f) 「短三和音+泣き顔・2回目以降」

図 5-1 実験 1,2 における条件ごとの平均調和度の推移 (図 3-2) とその指数近似曲線

1 回目」「短三和音+泣き顔・2 回目以降」の時定数の平均値は、いずれも「同期・1 回目」「同期 2 回目以降」の時定数の平均値よりも有意に長い (いずれも  $p<.05$ )。しかし、「長三和音+笑顔・1 回目」「長三和音+笑顔・2 回目以降」と「同期・1 回目」「同期 2 回目以降」の時定数の平均の間には有意な差は見られなかった。「短三和音+泣き顔・1 回目」と「短三和音+泣き顔・2 回目以降」の時定数の平均値の差、「長三和音+笑顔・1 回目」と「長三和音+笑顔・2 回目以降」の時定数の平均値の差も有意なものではなかった。

実験 1, 2 で得られた音と映像の調和感の形成過程に対して指数関数を当てはめることにより、各条件の時定数の比較を行った。得られた傾向は、2 章で行ったピーク到達時間による傾向と同様のものではあったが、条件間の統計的な有意差に関しては多少異なる部分もあった。

単純な視聴覚刺激の場合、構造的調和における調和度の時定数は 3 秒程度、意味的調和における調和度の時定数は 3 秒から 9 秒程度であった。これは、200 ms 程度とされている音の大きさの時定数よりもかなり長い (Scharf, 1978)。視聴覚統合過程が関わるとされる調和感形成過程は、音の大きさのような神経の興奮に関わる処理のレベルよりもより高次の処理過程であるために、かなり長い時定数を有する特性を持つものと考えられる。

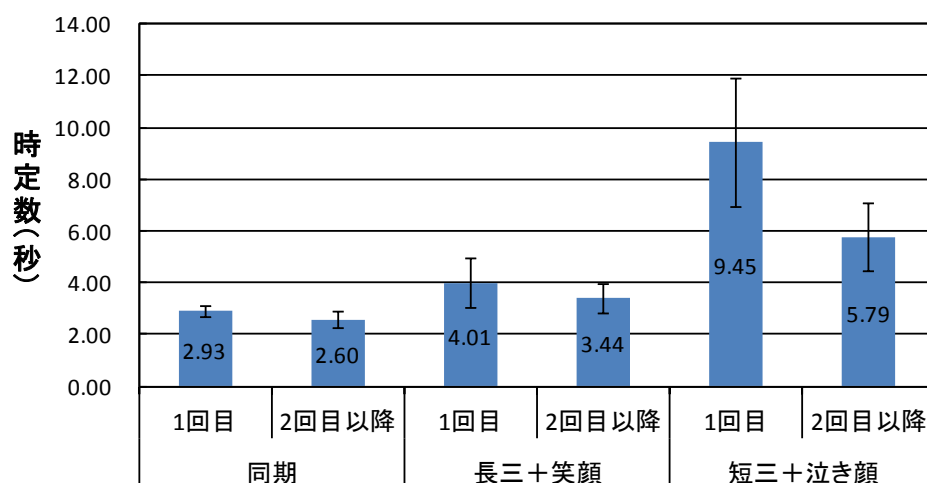


図 5-2 実験 1, 2 の条件ごとの平均調和度の推移の時定数 (平均と標準誤差)

表 5-1 実験 1,2 の条件ごとの調和度と近似値の相関係数

	同期		長三+笑顔		短三+泣き顔	
	1回目	2回目以降	1回目	2回目以降	1回目	2回目以降
相関係数 (平均)	0.94	0.92	0.91	0.90	0.94	0.92

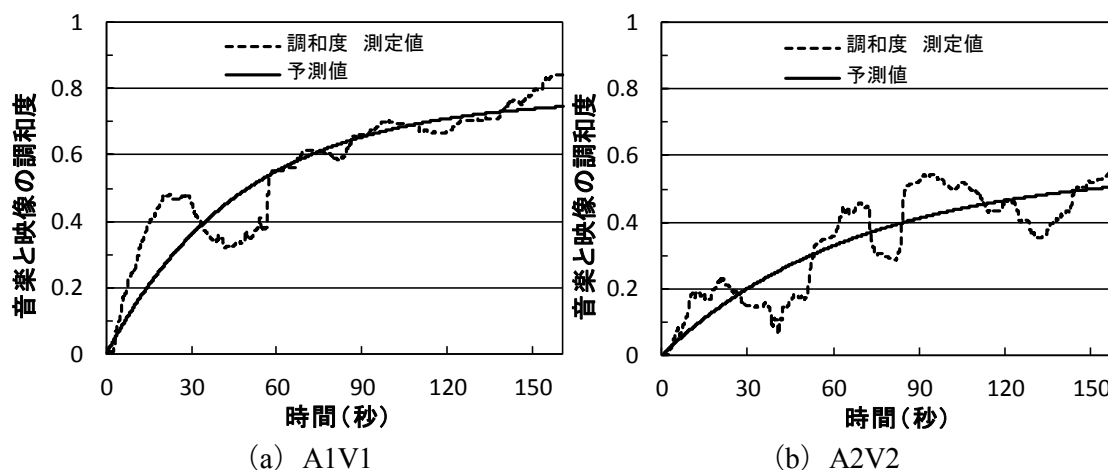
(すべて  $p < .01$ )

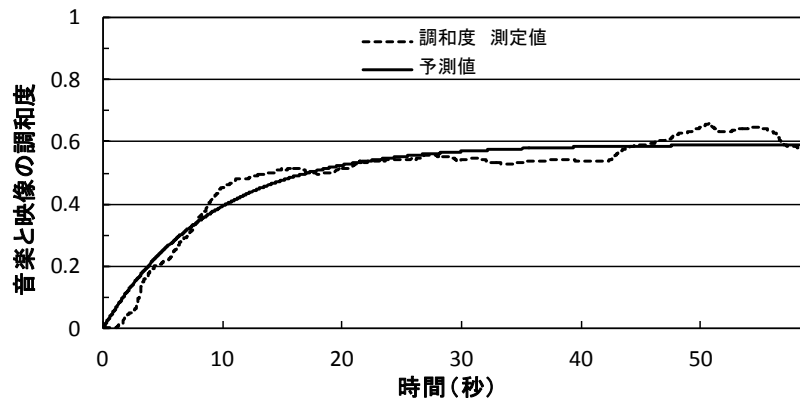
図 5-3 実験 3 の A1V1, A2V2 の平均調和度の推移とその指数近似曲線

市販の映像作品を用いた 3 章の実験 3 により得られた音楽と映像の調和感の形成過程に関しても同様に指数関数を当てはめた。このデータに関しては、実験参加者ごとのデータではなく、平均調和度のデータに対して (3 章でも、平均調和度で検討したため)、指数関数近似を行った。図 5-3 に A1V1, A2V2 の平均調和度の推移 (図 3-2) とその指数近似曲線を示す。平均調和度と近似値の相関係数は、A1V1 が 0.91, A2V2 が 0.87 であった (共に  $p < .01$ )。A1V1 の時定数は 48.00 秒, A2V2 の時定数は 68.20 秒となった。調和度のピークへの到達時間を求めて 3 章で指摘したように、後半に「迫力のある」印象が高まる音楽や映像を用いて展開に盛り上がりがある場合、2 章のように単純な視聴覚刺激の場合に比べ、より長い時定数となる。

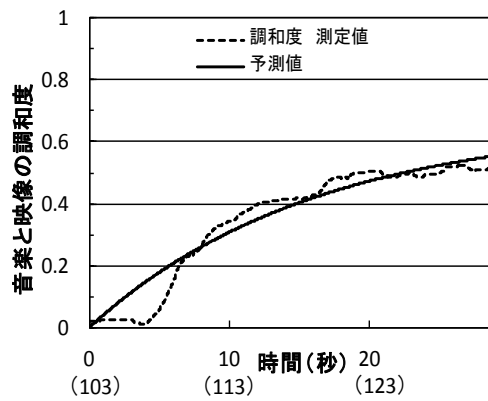
さらに、4 章の実験 6 で得られた「野良犬」の、音楽と映像の印象が一致して意味的

調和が形成されている AV-①, A'V-②, A'V-③, A'V-④の刺激に対する平均調和度のデータ (図 4-2) に関しても, 指数関数を当てはめた。平均調和度と近似値の相関係数は AV-①が 0.97, A'V-②が 0.97, A'V-③が 0.97, A'V-④が 0.98 であった (すべて  $p < .01$ )。図 5-4 に AV-①, A'V-②, A'V-③, A'V-④の平均調和度の推移とその指数近似曲線を示す。AV-①の時定数は 9.15 (秒), A'V-②の時定数は 16.05 (秒), A'V-③の時定数は 6.89 (秒), A'V-④の時定数は 55.43 (秒) であった。④をのぞき, これらの場面は映像作品の途中の場面であり, 場面の最後に向かって盛り上げる工夫などがなされていない。

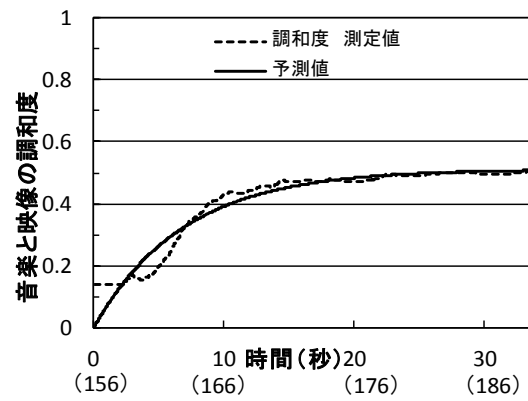
AV-①はオリジナル素材で短調の音楽が組み合わせられている。条件としては, 実験 2 の「短三和音+泣き顔・1 回目」と同様と考えられる。時定数もいずれも 9 秒程度と同程度で, 同じような音楽と映像の調和感の形成過程が得られたものと考えられる。A'V-②は, 組み合わせ素材ではあるが, 意味的調和のとれた視聴覚刺激である。時定数は AV-①よりも長めではあるが, 構造的調和の場合よりも時定数が長い。A'V-③も, 組み合わせ素材ではあるが, 意味的調和のとれた視聴覚刺激である。A'V-③は A'V-②の場面の延長であり, 同じ曲が使われている。A'V-③の時定数は A'V-②の時定数よりも短い, これは A'V-②の視聴時に意味的調和が形成されたことにより, そのことの記憶が A'V-③の意味的調和の形成の時定数を早めたものと考えられる。この効果は, 2 章で示された同じ視聴覚刺激を視聴したときに意味的調和の形成過程が早まる効果と同様のものと考えられる。A'V-④も, 組み合わせ素材であるが, 映像は映画のラストシーンである。実験参加者はそのことを承知していたわけではないが, 結末にむかう展開により調和感が上昇を続け, 50 秒以上の長い時定数になったものと考えられる。



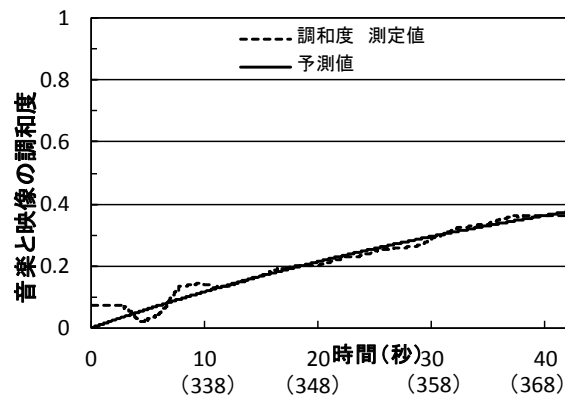
(a) AV-①



(b) A'V-②



(c) A'V-③



(d) A'V-④

図 5-4 実験 6 の AV-①, A'V-②, A'V-③, A'V-④ の平均調和度 (図 4-2) の推移とその指数近似曲線

### 5.3 視聴覚融合過程のモデル化

さらに、本研究で得られた実験結果をもとに、映像作品視聴時の視聴覚情報の統合過程をモデル化すると図 5-5 のように表すことができる。

視覚情報と聴覚情報は、視聴覚情報の統合処理の部分に伝達され、構造的調和と意味的調和を感知する過程に伝達されるものとする。構造的調和は、聴覚的アクセントと視覚的アクセントの同期によって得られるが、視覚情報と聴覚情報がほぼ直接利用されて同期しているかを判断して、得られる。一方、意味的調和は、視覚的印象と聴覚的印象を比較して得られるが、その前段階に聴覚的印象および視覚的印象を感知する過程が存在する。そのため、構造的調和に比べて意味的調和の形成に時間を要するものと考えられる。

意味的調和を感知する過程は、記憶とのつながりがあるものと考えられる。そして、視聴覚融合過程での処理結果は記憶にも蓄えられるものと考えられる。意味的調和は、蓄えられた記憶の影響を受け、一度感じた調和感は、同じ視聴覚情報を受け取ったときには感知しやすくなる。記憶の情報が調和感の形成に一種の促進効果を持つのであろう。

視聴覚融合過程では、映像作品の展開を解釈しようとしているが、音や映像の迫力感などにより、そこから盛り上がりを感じることもある。そのような盛り上がりを感じたときには、調和感をさらに上昇させることもある。また、映像作品の自然な展開の中では調和感も持続しやすく、調和を感じる要因が弱まってから後にも、調和感は維持される。

一般に、音楽と映像が調和している場合には、映像作品の評価も高いことが示されている（岩宮，2011）。しかし、黒澤作品のように例外的に音と映像の調和を崩した部分があっても、評価が高く保たれることもある（4章参照）。映像作品の展開を見守って整合性（つじつま）をチェックしている部分から、音と映像の調和を崩した部分があっても、その音がその作品に存在する必然性が感じられ、ユニークな印象も生ずることもあって、評価は高くなるのであろう。



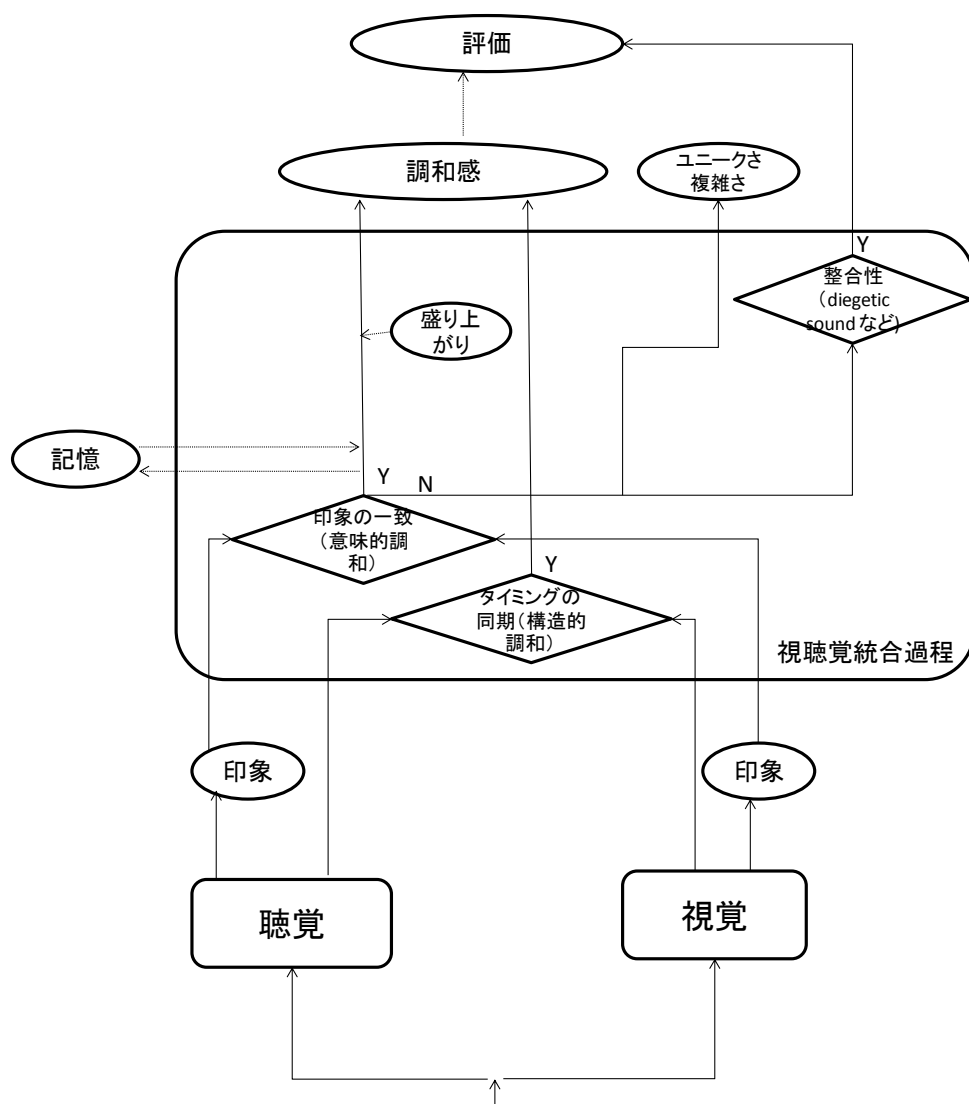


図 5-5 映像作品視聴時の視聴覚情報統合過程のモデル化

## 第6章 結論

一般に、音と映像の主観的な調和感には、音と映像のアクセントを一致させることによって得られる構造的調和と、音と映像のそれぞれが持つ印象を一致させることによって得られる意味的調和があると言われている。本研究では、視聴覚刺激の音（あるいは音楽）と映像の調和感の連続測定実験、音（あるいは音楽）と映像の印象の連続記述実験などを実施して、映像作品において音楽と映像の調和感が形成される過程を明らかにし、市販の作品における意味的調和の状況およびその効果を探ることを試みた。

本研究では、音と映像の同期や印象の類似を単純な形で構成した視聴覚素材や、市販の映像作品の一部を用いることにより、さまざまな条件下で音（あるいは音楽）と映像の調和感の形成過程を探った。

意味的調和を得るための視聴覚情報の統合処理は、音の印象および映像の印象を感知し、さらに両方の印象を比較することにより印象の類似を捉えるものと考えられる。意味的調和の処理過程は、音と映像のアクセントの同期を監視するのみの構造的調和の処理過程に比べより複雑な過程を含むことになるため、その処理には構造的調和を感知するよりも時間を要することが示された。

さらに、いったん意味的調和が形成されると、作品中でその要因が弱まっても意味的調和の効果は持続し、調和感が保たれる傾向が得られた。市販の作品においては、その展開に応じて、さらに累積的に調和感が上昇することもあることが示された。

一般的には、映像作品の制作者は意味的調和の調和を図るが、わざと音楽と映像の印象を類似させないこともある。黒澤明監督の作品においてしばしば用いられた「音と画の対位法」はその代表的なものである。黒澤明監督作品から対位法が用いられている場面では、実験により音楽と映像の調和感は低いことが示された。

ただし、印象評定実験により、対位法が使われた視聴覚刺激の総合的な良さの評価は、全般的に音楽と映像の意味的調和がとれた視聴覚刺激と同程度かそれ以上に高いことが示され、黒澤作品では対位法の手法が効果的に使われているものと考えられる。さらに、対位法が効果的にあるためには、映像の印象と対立する音楽が整合性を持って映像内に存在する必要があることも確認された。黒澤監督は、そのために音楽の演奏風景を挿入したのと考えられる。

本研究で実施した一連の実験により、意味的調和の形成過程についての処理過程が明

らかになるとともに、市販の作品における意味的調和の状況が示された。さらに、意味的調和を崩す手法を効果的に用いる手法が成立しうることも明らかになった。映像作品の制作者は、意味的調和の効果をうまく活用しているものと考えられる。本研究によって得られた実験結果により、映像作品における音楽の効果を論じるための知見、作品制作面での参考資料を示すことができた。

1章で述べたように、音（あるいは音楽）と映像の相互作用の研究、音（あるいは音楽）と映像の調和に関する研究などは、これまでも多く行われている。しかし、音（あるいは音楽）と映像の調和感が形成される過程を対象としたものはこれまでになく、本研究がこの分野に果たした貢献は大きいと思われる。また、音楽と映像の調和をわざと崩すような手法にまで科学的手法で切り込んだ研究も、これまでにないものである。映像作品における音楽と映像が融合する過程を科学的に解明するためには、まだまだ研究すべき課題は少なくないが、本研究もこの分野の発展の一步に貢献したものと考えられる。

## 参考文献

- 秋山裕和, フリー音楽素材 H/MIX GALLERY, <http://www.hmix.net/music/k/k7.mp3>, 参照 2012-08-01.
- Bolivar, V. J., Cohen, A. J., & Fentress, J. C., Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action, *Psychomusicology*, 13, 28-59, 1994.
- Boltz, M. G., Schulkind, M. & Kantra, S., Effects of background music on the remembering of filmed events, *Memory and Cognition*, 19, 593-606, 1991.
- Chion, M., *Le Son au cinema* (映画にとって音とは何か), 勁草書房, 1993.
- Chion, M., *Le musique au cinema* (映画の音楽), みすず書房, 2002.
- Cook, N., *Analyzing musical multimedia*, Oxford University Press, 1998.
- Crowder, R. G., Perception of the major/minor distinction: I. Historical and theoretical foundations, *Psychomusicology*, 4, 3-12, 1984.
- Disney, W. (Producer), *FANTASIA* [DVD], 2000.
- Disney, R. E. & Ernst, D. W. (Producer), *FANTASIA 2000* [DVD], 2000.
- Ekuman, P. & Friesen, W. V., Constants across cultures in the face and emotion, *Journal of Personality and Social psychology*, 17, 124-129, 1971.
- Gorbman, C., *Unheard melodies: Narrative film music*, Indiana University Press, 1987.

Gregory, D., Using computers to measure continuous music responses, *Psychomusicology*, 2, 127-134, 1989.

Hevner, K., Experimental studies of the elements of expression in music, *American Journal of Psychology*, 48, 246-268, 1936.

岩宮眞一郎, オーディオ・ヴィジュアル・メディアを通しての情報伝達における視覚と聴覚の相互作用に及ぼす音と映像の調和の影響, *日本音響学会誌*, 48, 649-657, 1992.

岩宮眞一郎, 視覚と聴覚の相互作用に及ぼす音響再生系の音質の影響—オーディオ信号に帯域制限を加えた場合—, *JAS journal*, 33, 29-35, 1993.

Iwamiya, S., Interaction between auditory and visual processing when listening to music in an audio visual context: 1. Matching 2. Audio quality, *Psychomusicology*, 13, 133-154, 1994.

岩宮眞一郎, 音楽と映像のマルチモーダル・コミュニケーション改訂版, 九州大学出版会, 2011.

岩宮眞一郎, 林克明, 色彩が音楽の印象に与える影響, *芸術工学研究*, 1, 63-68, 1999.

岩宮眞一郎, 上月裕, 菅野禎盛, 高田正幸, 音楽の調性及びテンポと映像の速度及び密度が映像作品の印象に及ぼす影響, *音楽知覚認知研究*, 8, 53-64, 2002.

神宮英夫, 印象測定論: 2. SD法を考える, *東京学芸大学紀要*. 第1部門, 教育科学, 43, 99-104, 1992.

Kalinak, K., *Settling the score*, University of Wisconsin Press, 1992.

金森慎弥, 米田涼, 山田真司, ゲーム音楽とゲーム映像の調和感に関する研究, *音楽知覚認知学会平成24年度春季研究発表会資料*, 75-78, 2012.

金基弘, 岩宮眞一郎, 藤丸沙由美, テロップの書体と効果音の印象の類似の効果, 音楽知覚認知研究, 11, 2, 73-90, 2005.

北村晴朗 (編著), 要説心理学, 昭学社, 1967.

Kracauer, S., *Theory of film: The redemption of physical reality*, Oxford University Press, 1960.

黒澤明, 野良犬 [DVD], 2003.

Kuwano, S. & Namba, S., Continuous judgment of level-fluctuating sounds and the relationship between overall loudness and instantaneous loudness, *Psychological research*, 47, 27-37, 1985.

Lipscomb, S. & Kendall, R., Perceptual judgment of relationship between musical and visual components in film., *Psychomusicology*, 13, 60-98, 1994.

Lipscomb, S. D., Cognition of musical and visual accent structure alignment in film and animation, *Proceedings of 4<sup>th</sup> International Conference of Music Perception and Cognition*, 309-313, 1996.

MacGurk, H. & MacDonald, J., Hearing lips and seeing voices, *Nature*, 264, 746-748, 1976.

Marshall, S. K. & Cohen, A. J., Effects of musical soundtracks on attitudes to geometric figures, *Music Perception*, 6, 95-112, 1988.

丸山欣哉, 講座心理学 3, 感覚, 第 8 章「感覚間相互作用」, 東京大学出版会, 267-297, 1969

Münsterberg, H., *The photoplay: A psychological study and other writings*, Routledge, 1916/2001.

武藤真介, 計量心理学, 朝倉書店, 76-90, 1982.

中村滋延, 現代音楽×メディアアート 音響と映像のシンセシス, 九州大学出版会, 2008.

難波精一郎, 桑野園子, 連続記述選択法による音楽演奏の評価, 京都市立芸大紀要, 18, 6-38, 1988.

難波精一郎, 桑野園子, 日本音響学会編 音響テクノロジーシリーズ4 音の評価のための心理学的測定法, コロナ社, 2008.

Namba, S., Kuwano, S., Hato, T. & Kato, M., Assessment of musical performance by using the method of continuous judgment by selected description, *Music Perception*, 8, 251-276, 1991.

難波精一郎, 桑野園子, 中村敏枝, ピアノ演奏音の解釈—大きさを手がかりとして—, 大阪大学教養部研究集録, 25, 25-43, 1977.

Negel, F., Kopeis, R., Grewe, E. & Altenmuller, E., EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods*, 39, 283-290, 2007.

日本音響学会 編, 音響用語辞典, コロナ社, 1988.

西村雄一郎, 黒澤明 音と映像, 立風書房, 73-90, 1998

Ortony, A., Clore, G. L. & Collins, A., *The cognitive structure of the emotions*, New York: Cambridge University Press, 1988.

Osgood, C. E., Suci, J. G. & Tannenbaum, P. H., *The measurement of meaning*, Illinois Press, 1957.

Palmer, C., Film music. In *New Grove dictionary of music and musicians*, 6, 549-556, 1980.

Percheron, D. & Butzel, M., Sound in cinema and its relationship to image and diegesis, *Yele French Studies*, 60 16-23, 1980.

Peretz, I., Gagnon, L. & Bouchard, B., Music and emotion: perceptual determinants, immediacy, and isolation, after brain damage, *Cognition*, 68, 111-141, 1998.

Ryan, T. A., Interrelation of the sensory systems in perception, *Psychol. Bull.*, 37, 659-698, 1940.

Scharf, B., "Loudness," in *Handbook of Perception*, 1978.

Schubert, E., Continuous response to music using a two dimensional emotion space, *Proceedings of 4<sup>th</sup> International Conference on Music Perception and Cognition*, 263-268, 1996.

Schubert, E., Modeling perceived emotion with continuous musical features, *Music Perception*, 4, 561-585, 2004.

菅野禎盛, 岩宮眞一郎, 映像と音楽の情緒的印象に対する同期要因と速度対応要因の効果, *日本音響学会誌*, 56, 695-704, 2000.

谷口高士, 音楽作品に感情価測定尺度の作成および多面的感情状態尺度との関連の検討, *心理学研究*, 65, 463-470, 1995.

Tannenbaum, P. H., Music background in the judgment of stage and television drama, *Audio-Visual Communication Review*, 4, 92-101, 1956.

和田陽平, *心理学 I*, 形の知覚, 培風館, 106-111, 1967.



Weber, R., The continuous loudness judgment of temporally variable sounds with an “analog” category scale produce, *Contributions to Psychological Acoustics*, 267-294, 1991.

Whitaker, R., 池田博／横川真顕訳, *The language of film (映画の言語)*, 法政大学出版局, 1970／1983.

Yamada, M., The effect of music on the fear emotion in the context of a survival-horror video game, *Proceedings of 10<sup>th</sup> International Conference on Music Perception and Cognition*, 594-597, 2008.

## 謝辞

本研究を進め、本論文を纏めるにあたりまして、九州大学大学院芸術工学研究院 岩宮眞一郎教授には終始温かく適切なお指導を賜りました。ここに深く感謝を申し上げます。ご審査いただいた、九州大学大学院芸術工学研究院 中村滋延教授、金沢工業大学 情報フロンティア学部メディア情報学科 山田真司教授には多くの有意義なお指導、ご助言を賜りました。ここに厚く御礼を申し上げます。また、多方面にわたり多くのご助言を賜りました九州大学大学院芸術工学研究院 高田正幸助教、金基弘博士に感謝を申し上げます。大西英治氏、江間琴音氏、矢萩徹氏、瀧下郁之氏には、実験の準備から解析まで多大なお協力をいただきました。ここに感謝をいたします。最後に、様々な場面で助けていただいた岩宮研究室、高田研究室のみなさまに心から感謝の意を表します。