

# 演奏者の意図は聴取者にどう伝わるのか： フレーズ表現における視覚情報の影響分析

藤坂 亜南<sup>1</sup> 橋田 光代<sup>2</sup> 片寄 晴弘<sup>1</sup>

概要：本研究では、演奏表現が聴衆にどう解されるのかという課題について取り扱う。特に、フレーズ表現の理解という課題に対して、演奏映像、つまり、視覚的情報がどのように関与するかを調査することを目的とする。フレーズ構造の意図伝達は、演奏者が意図した表現と聴取者が聴き取った表現がどれだけ一致しているかで表すことができることから、本研究では、クラシックピアノ・ソロの楽曲を演奏している音声のみの動画と、映像付きの動画を視聴してもらい、フレーズ構造を記述する被験者実験を実施する方法をとる。しかし、フレーズ構造を記述する難しさ、また昨今で話題になっている COVID-19 感染対策により対面実験が行えなくなったことから、実験用の Web システムを開発し、オンライン化で実験を行なった。視聴実験の結果から、演奏映像をつけることでフレーズ構造の伝達度は低くなり、特にピアノを経験している被験者はその傾向が強いことが示唆された。

## 1. はじめに

演奏を聴く人々は、聴こえてきた音から感情、表現といった演奏者の意図を感じ取り、音楽鑑賞を楽しんでいる。同じ曲でも異なる演奏者によって聴こえ方が違うのは、各演奏者の考え方や意図が異なるからであり、聴取者側はその違いを聴き分けるのも楽しみの一つである。ここで述べる演奏者の意図とは、強弱、演奏テクニック、感情、動作表現など様々な要素が含まれたものである。演奏者はこれらを如何にして観客に伝えるかを追究し、演奏を聴く人々に感動を与えるために音として届けようとしている。

演奏意図の伝達に関する研究は昔から存在している [1], [2], [3], [4], [5], [6]。千住らは、演奏者が込めた「形容詞系」の意図が聴取者にどのように伝達されるのかの検討を行った [7]。一方で、演奏意図の伝達に関するもう一つの切り口としてモダリティがあげられる [8], [9], [10]。聴覚情報に対して、視覚情報が加わることで、表現意図の伝達はどう高まるかということに対する取り組みである。これらの研究のほとんどは、音楽聴取における視覚情報は重要であると言われている。

我々は、演奏表現の一つである「フレーズ構造」に焦点を当てて、演奏者の意図が聴取者にどう伝わるのかの検討を進めている [11]。先行研究では、まず「伝わりやすさ」

をどのように定義すべきかという課題があり、階層的なフレーズ構造と頂点音の類似性指標を求める方策についての議論を実施した。本研究では、次の段階として、視覚情報の有無によるフレーズ構造の伝わりやすさの影響について調べることを目的とする。

## 2. フレーズ構造に基づく意図の伝達

「演奏者の意図」とは、広義には、演奏者が「このように演奏したい」という考えや事柄を指す。意図そのものは聴取者が目で直接捉えられるものではないが、奏でられた音や、演奏者の身体動作といった演奏表現行為を通じて間接的に汲み取っていくものである。演奏者が表出した表現と聴取者が感じ取った表現が合致するほど、演奏者の意図は正確に伝わっているとみなすことができる。

フレーズ構造は、生成音楽理論 (A Generative Theory of Tonal Music : GTTM)[12] のサブ理論となるグルーピング構造に準拠したものである。演奏意図の伝達メディアとしてフレーズ構造を扱う最大のメリットは、楽譜に書かれた音符にマーキングをすることでシンボリックに記述可能であるという点である。これに加えて本研究では保科理論 [13] の頂点を参考とし、各フレーズの中で最も印象に残る音を「頂点音」と呼び、フレーズ構造の一要素として取り扱うこととする (図 1)。

### 2.1 演奏者のフレーズ構造の記述

演奏におけるフレーズ構造の伝達という問題を取り扱う

<sup>1</sup> 関西学院大学理工学部  
Kwansei Gakuin University

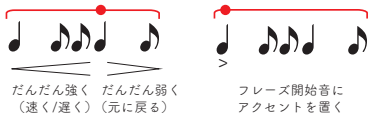
<sup>2</sup> 福知山公立大学情報学部  
The University of Fukuchiyama



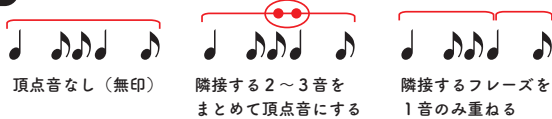
図 1 階層的フレーズ構造の一例．丸印はフレーズの頂点を示す（文献 [11] より抜粋）．

### 基本（フレーズと頂点）

「これでひとまとまり」と思う範囲をフレーズとする  
1つのフレーズには頂点音を1つだけ置く



### 可能



### 禁止



図 2 階層的フレーズ構造の記述ルールの一例（文献 [11] より抜粋）．

にあたり、まずは、演奏者自身が記述したフレーズ構造のデータが必要不可欠である。これは先行研究において構築を進めている演奏表情データベース (PEDB) 第2版 [14] 内に収録された「演奏者自身が解釈したフレーズ構造」が書き込まれた楽譜とその時収録された映像を利用することで解決できる。PEDBは、ピアノの名演奏を主に対象とした演奏表情データベースであり、第2版では、フレーズ構造情報の提供を目的に、2016年より、フレーズ構造情報の取得に注力されている。

フレーズ構造の記述にあたっては、スラーや強弱記号等のない五線譜にフレーズの範囲と頂点を手書きで記入してもらう必要があった。演奏者間でフレーズの意味や記述方針にブレが出ないように、GTTM [12] と保科理論 [13] での頂点音の説明に準拠する形で、文献 [11] で設けた統一的なルール（図 2）を設け、インタビューを通じて擦り合わせを行った。

## 2.2 視覚情報を伴う聴取実験における課題

視覚情報の有無により聴取者のフレーズ構造の変化を探るにあたり、本研究では、(a) 演奏音のみの音声データと、(b) 映像付きの演奏動画を実験刺激とした聴取実験を行う。この際、聴取者がフレーズ構造を「聴き取る」際には、演奏音を聴くと同時に映像も観る必要が生じる。その上で、フレーズ構造を譜面に「書き出す」作業では、演奏音を聴きながら机上の譜面を視覚的に追跡する必要に迫られる。つまり、被験者は、演奏を聴きながら、映像観察と譜面追



図 3 映像観察と譜面追跡の同時遂行

跡を同時に行う必要がある。そもそも演奏を聴きながら譜面を追跡すること自体に慣れていない人には重いタスクである。さらに動画で演奏者の動きを見るとなると、楽譜と動画を交互に見なければならず、音楽経験者でも集中して聴くことが困難になる。

この問題を解決するには、被験者の負担を極力減らし、スムーズに実験を行うことが必要となる。以下に、視覚情報を伴う聴取実験における要求事項を整理する。

### 2.2.1 聴取者の視線移動の削減

本研究が行う実験形態は、譜面と演奏映像を交互に見ることが必要不可欠である。ここで、実験環境として、例えば図 3 で示すように、演奏映像を机前面のスクリーンに投影するという形態を取ると、映像と譜面との距離が離れて、視線の移動距離が長くなり、実験参加者の負担になりかねない。この理由により、視線移動の距離を最小限に抑えることが重要である。

### 2.2.2 譜面追跡の容易化

譜面は、今どこを演奏しているかを表す地図のような役割があり、演奏を聴取する上でとても参考になるものである。しかし、フレーズ構造を記述する際は映像観察も譜面追跡も中断することになるため、聴取に戻るたびに、演奏中の譜面の位置を常に示すものが必要である。

### 2.2.3 フレーズ構造記述方法の容易化

聴取者に求めるフレーズ構造の記述方法は、2.1 節および図 2 で示したものと同一フォーマットを用いればよい。ただし、聴取者によるフレーズ構造の記述作業においては、前述したように、映像観察も譜面追跡も一時中断する必要に迫られる。このとき、フレーズ構造の記述ルールが複雑になると、記述作業に手間取るため聴取作業に戻るのに時間がかかる。これを解決するためには、極力単純な方法でシンプルにフレーズ構造を記述する仕組みを構築する必要がある。

## 2.3 遠隔操作による聴取実験における課題

先行研究 [11] における聴取実験では、1名から複数の被験者が共通の会場に集まっており、譜面への手書きによるフレーズ構造の記述法について、実験者が直接教示を行いながら作業を進めていた。そのメリットは、複数の被験者間で、一貫したフレーズ構造を記述できているかの確認作

業をその場で行うことができたという点である。

しかし、2021年現在では、COVID-19感染対策のために、実験者、被験者ともに集合しての実験を行うことが困難となっている。また、将来的に、同様の聴取実験を実施するにあたって、対面かつ手書きという記述方法では被験者数の増強が困難となることから、対面型ではなくオンライン上での視聴実験システムを構築する必要がある。

オンラインで実施するフレーズ聴取実験においては、参加者が別々の場所で独立して実験を進められるようなスタイルを取らなければならない。実験実施にはシステムの使い方や実験概要を説明するためのガイダンスも必要であるが、事前に説明用の動画を作ったとしても記述作業にあたって参加者に疑問が生じた際にはリアルタイムでスムーズに対応する必要がある。フレーズ構造の記述ミス为了避免するためにも、リアルタイムにデモンストレーションもしながら解説や詳細な指示を加えていく必要がある。

### 3. オンライン視聴実験システム

第3章では、フレーズ構造を対象として、視覚情報の有無による聴取の影響をオンライン上で調べるためのWebシステムについて述べる。本研究では、オンラインで実験を行うことから、インターネットに接続できればどこでも実験が可能となるような設計にする方が実験の効率が良い。そこで、マルチプラットフォームに対応しているHTML5とJavaScriptで作成したブラウザベースのアプリケーションを作成した。

システムに含まれる要素は大きく分けて3つあり、(1)視聴対象となる音声データや映像データの提示部、(2)楽譜の譜面が表示されたスペース、そして(3)フレーズ構造を記入するスペースである。以上の事項を取り入れたシステムの画面を図4に示す。

#### 3.1 演奏提示部

演奏提示部では、映像内の演奏者の身体動作ができるだけ明確に見えるように、映像の表示寸法をなるべく大きくする必要がある。ここでは、画面左側に表示された演奏動画を視聴しながら、画面右側でフレーズ構造を入力できるように配置した。演奏動画は参加者のPCマウス操作により再生、停止を繰り返し視聴することが可能である。

#### 3.2 譜面表示部

譜面表示部に関しては、フレーズ構造を記述する際の目印になり、どこからどこまでを範囲とするか、どの音を頂点と置くのかを手助けする役割がある。また、2.2.2節で述べたように、今どこを演奏しているのかを一目で判断できるようにするためのスクロールバーを配置している。これにより、被験者は譜面追跡の手間を省くことができる。

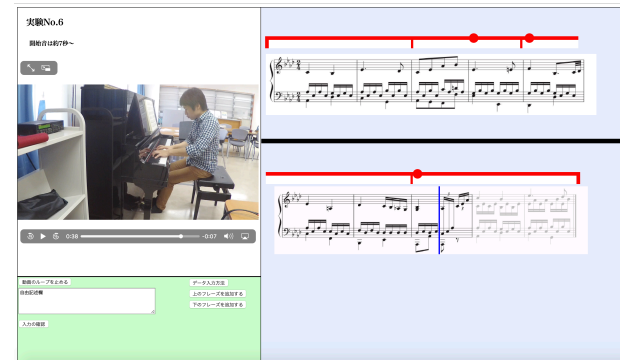


図4 Webシステム画面。左側上：演奏提示部、右側：譜面表示部、フレーズ構造記述部、左側下：その他の操作・入力部

#### 3.3 フレーズ構造記述部

フレーズ構造を記入するスペースは、限りなく譜面に近い場所にあり、かつ譜面を読む時に障害しない位置でなければならない。本実験で使用した譜面はピアノ譜であり、基本的に上段が主旋律であるため、上段の少し上の部分にフレーズ構造を書くスペースがあると理想である。また、画面上での視線移動を減らす必要がある(2.2.1節)ため、映像データと楽譜も極力近くなるような配置にしなければならない。

そこで、参加者がフレーズ構造を入力する場所は、譜面表示部内の各五線譜の上部とした。ここで参加者が行える操作は、フレーズの範囲を決定する「区切り(前後のフレーズの境目)」と、各フレーズの「頂点」をマウス操作で描画(作成、移動、削除)することである(詳細は3.5節)。区切りは赤い線、頂点は赤い丸として表示される。

#### 3.4 その他の操作および入力部

最後に、画面左下には、フレーズ構造の階層数を変化させるボタンや、被験者が演奏を視聴した時に被験者が気づいたこと、感じたことなどを記入する自由記述欄を設けている。また、記入方法を忘れた場合のための説明資料を閲覧できるボタンを設置している。さらに、演奏動画の自動ループをオン、オフさせるボタンを設置している。演奏動画は、再生が終了すると自動的に動画の最初へ戻り、ふたたび再生するように設定している。これは、動画が終了した後に被験者が手作業で動画を繰り返し再生させることによる手間を省くためである。

#### 3.5 フレーズ構造の記入方法

2.2.3節で述べたように、フレーズ構造の記述作業は、参加者の負担にならないよう限りなく少ない操作をとるようにしたい。そこで、フレーズの区切りと頂点の入力と変更・削除に関して、マウスのクリック操作とDeleteキーのみでできるように設計した。以下、フレーズ構造と頂点の記述方法をそれぞれ説明する。

### 3.5.1 区切り

区切りを作成する場合、赤線の少し下部分をマウスをクリックすることで赤い線が新たに表示される。これにより、前方のフレーズの終了音と後方のフレーズの開始音が決定する。区切りの位置を変更したい場合、作成された区切りをもう一度クリックすると、ピンク色に表示され、編集可能な状態となる。この状態からマウスを移動させると、マウスと同方向に区切りを移動することが可能である。移動した区切りの位置を決定させたい場合、その区切りをもう一度クリックすると赤く変色し、固定される。作成した区切りが必要なくなった場合は、同様にクリックして編集状態にし、Delete キーを押すことで選択した区切りを削除することが可能である。

### 3.5.2 頂点

頂点を作成する場合、区切り同様クリック操作になるが、クリックする位置を赤線上にすることで頂点を作成される。同じフレーズ内で頂点の位置を変更したい場合、別の位置の赤線上をクリックすると新たに頂点の位置が変更される。これにより、一つのフレーズに頂点が2つ以上作成されることはなくなる。作成した頂点が必要ないと判断した場合、その頂点をクリックすることでピンク色に変色される。その状態から Delete キーを押すと、頂点が削除される。

## 4. 一般聴取者に向けての視聴実験

大学生の男女 61 名を対象に、3 章で述べた Web システムを利用した聴取実験を行った。実験の実施にあたり、フレーズ構造の記述方法の教示や質問にリアルタイムで対応するために、セミナーやミーティングをオンラインで開催するために開発されたアプリ Zoom[15] を併用した。Zoom の特徴として画面共有機能がついており、こちらでシステムを使ったデモンストレーションを行うことで、被験者にわかりやすく説明することが可能である。また、質問の対応方法として Zoom のコメント機能を利用する。被験者は、Zoom を起動しつつ Web システムによる視聴実験を行い、質問等があれば Zoom のチャットにてコミュニケーションを取れるような形態をとった。

### 4.1 楽曲

本研究では、2.1 節で述べた演奏表情データベース PEDB[14] に保存されている楽曲の中から Beethoven の Piano Sonata No.8 Op.13 「悲愴」第 2 楽章の冒頭 8 小節を用いることにした。この曲は Beethoven の三大ピアノソナタの一つとしても数えられるほど有名な楽曲である。ピアノ経験者であれば演奏する機会も多く、演奏者の意図を比較的盛り込みやすいという特徴があり、多様に解釈されやすい曲でもある。

### 4.2 演奏動画

被験者に視聴させる演奏動画は、PEDB 収録の際に撮影されたビデオを用いる。ビデオは、それぞれ演奏者の横側全身から撮影されており、顔の表情や手の動き、足のペダル制御が見えるものである。

対象とする演奏者は、PEDB 所収のうち、フレーズ構造の類似性が低く、かつ身体動作の違いが分かりやすいピアニスト 5 人とした。収録ビデオの音声では、ピアノの音だけでなく、鍵盤を押した時の打音や環境音も含まれているため、ノイズ除去処理を施した録音音源とビデオを合成させ、本実験の提示刺激として用いた。

### 4.3 手続き

被験者が行う作業は、Web システム上で、演奏動画を視聴し、聴き取ったフレーズ構造を逐次記入するというものである。提示した演奏動画は、5 人分のピアニストによる「映像なし(音声のみ)」「映像あり」の 2 種類の動画\*1、全 10 点である。自身が納得するまで演奏動画は繰り返し視聴することができ、記入が終われば次の動画に進んでもらうようにした。平均的に、一つの動画に対する視聴回数は 3 回程度であった。

### 4.4 演奏者と聴取者間のフレーズ構造の類似度

演奏者の意図したフレーズ構造が伝達できたかどうかの度合いは、演奏者と聴取者がそれぞれ同じフォーマットを用いて記述したフレーズ構造を比較することで調べることができる。ここでは、先行研究 [11] で用いた手法に沿って、各フレーズの開始音、頂点音、終了音のそれぞれに対して、演奏者と聴取者がとらえた双方の音符集合適合率の調和平均として与えられる類似性指標を用いて計算する。具体的には、まず演奏者  $p_1$  と聴取者  $a_1$  の捉えたそれぞれのフレーズ構造から、連続した二つの階層のフレーズ開始音から構成される音符の集合を抜き出し、その集合に対して、類似性指標が最大になるような階層の組み合わせ  $N_{p_1}$ 、 $N_{a_1}$  を求める。得られた階層の組み合わせにおいて、開始音、頂点音、終了音のそれぞれの音符集合に対して適合率  $P$  を計算し、その調和平均を求めることで、演奏者と聴取者のフレーズ構造の一致度とする。

$$\text{一致度}(p_1, a_1) = \text{調和平均} \left( P(N_{p_1} \rightarrow N_{a_1}), P(N_{a_1} \rightarrow N_{p_1}) \right)$$

その後、聴取者人数分のフレーズ構造の一致度の調和平均を求めて、演奏者  $p_1$  に対する聴取者群の類似度とした。

### 4.5 実験結果と考察

この実験で得た結果を図 5 に表す。図 5 の左列は演奏者の姿が見えない「映像なし条件」時による演奏データ、右列は「映像あり条件」時による演奏データを指す。上か

\*1 「音声のみ」動画の場合、映像部分はブラックアウトされている。

P1-映像無					P1-映像有				
ALL	開始音(S)	頂点音(P)	終了音(E)	ALL	開始音(S)	頂点音(P)	終了音(E)		
全員	0.25	0.37	0.13	0.25	0.42	0.09	0.25		
EXP	0.27	0.34	0.20	0.27	0.34	0.43	0.22		
MID	0.24	0.36	0.13	0.22	0.25	0.45	0.06		
NON	0.23	0.39	0.10	0.21	0.21	0.41	0.08		

P2-映像無					P2-映像有				
ALL	開始音(S)	頂点音(P)	終了音(E)	ALL	開始音(S)	頂点音(P)	終了音(E)		
全員	0.45	0.67	0.00	0.67	0.46	0.85	0.00		
EXP	0.43	0.64	0.00	0.64	0.47	0.71	0.00		
MID	0.45	0.67	0.00	0.67	0.48	0.73	0.00		
NON	0.45	0.67	0.00	0.67	0.44	0.66	0.00		

P3-映像無					P3-映像有				
ALL	開始音(S)	頂点音(P)	終了音(E)	ALL	開始音(S)	頂点音(P)	終了音(E)		
全員	0.39	0.39	0.39	0.40	0.41	0.40	0.42		
EXP	0.51	0.49	0.53	0.52	0.46	0.42	0.52		
MID	0.38	0.38	0.38	0.39	0.41	0.42	0.38		
NON	0.36	0.35	0.36	0.38	0.40	0.38	0.42		

P4-映像無					P4-映像有				
ALL	開始音(S)	頂点音(P)	終了音(E)	ALL	開始音(S)	頂点音(P)	終了音(E)		
全員	0.30	0.61	0.06	0.23	0.34	0.61	0.15		
EXP	0.33	0.60	0.13	0.26	0.33	0.59	0.10		
MID	0.29	0.62	0.04	0.21	0.30	0.64	0.04		
NON	0.28	0.61	0.05	0.19	0.28	0.60	0.02		

P5-映像無					P5-映像有				
ALL	開始音(S)	頂点音(P)	終了音(E)	ALL	開始音(S)	頂点音(P)	終了音(E)		
全員	0.32	0.40	0.34	0.23	0.39	0.42	0.36		
EXP	0.39	0.41	0.45	0.30	0.38	0.44	0.29		
MID	0.32	0.41	0.34	0.21	0.37	0.42	0.34		
NON	0.29	0.40	0.30	0.16	0.38	0.42	0.39		

映像なし条件

映像あり条件

図 5 各演奏データに対する聴取者のフレーズ構造の類似度。類似度が高いほど緑に、低いほど赤く表示している

ら順に、演奏者をそれぞれ P1~P5 と振り分けている。表内の各数字はフレーズの開始音 (S), 頂点音 (A), 終了音 (E), それら全て (ALL) を対象とした、ピアニストと被験者群との類似度を表しており、1.0 で完全に一致 (全員が演奏者と同じ音を対象音として聴取した) となる。図 5 では、類似度が高いほど濃い緑へ、低いほど赤くなるよう色分けしている。各表の最左列は被験者群の分類を表しており、楽器経験総年数が 10 年以上の被験者を EXP 群、10 年未満の被験者を MID 群、学校教育での授業以外に特に経験のない者を NON 群とした。

各ピアニストにおいて「映像なし」「映像あり」の各条件での類似度を比較すると、両者で顕著な差は特に見られなかった。理由としては、フレーズ構造のみに集中して鑑賞することに慣れておらず、特に楽器経験の浅い被験者にはタスク自体が難しかったのではないかと推測される。また、映像が追加されたからといって、クラシックピアニストの動きを一つ一つ理解することは、同じようにクラシック音楽を経験したピアノ経験者でないと難しいと感じるかもしれない。今回の実験では、半数以上が楽器未経験者であり、楽器経験、特にピアノ経験の被験者が比較的少なかったため、正確なデータとして現れなかったと考えられる。

## 5. 楽器経験者に向けた聴取実験

4 章の実験では、楽器経験の有無にかかわらず被験者を集めていたことを受けて、楽器経験者を対象に、同様の手続きで再度実験を実施した。

### 5.1 手続き

前章の実験に参加していない、かつ何らかの楽器演奏経験のある男女 23 名に参加してもらった。実験の流れは前

回の実験と同様であり、Web システムを使ってデータ入力をしてもらった。また本実験の最後に、アンサンブル演奏経験、コンクール経験など、フレーズ表現に関連したいくつかの項目の聞き取りを実施した。使用した楽曲は前回でも使用した「悲愴」に加え、Chopin の夜想曲第 2 番「ノクターン」、Beethoven の「エリーゼのために」の計 3 曲とした。追加した 2 曲の演奏動画は、「悲愴」と同様に PEDB で所有されている横向きの動画を採用し、録音音源と合成して実験刺激として用いた。

### 5.2 実験結果

今回の実験で得られた結果を図 6 に表す。表の見方は図 5 同様で、左列は「映像なし条件」時によるフレーズ構造解釈、右列は「映像あり条件」時による解釈である。演奏者をそれぞれ P1~P5 と振り分けており、フレーズの開始音 (S), 頂点音 (P), 終了音 (E) それぞれの一致度が高いほど濃い緑色に、低いほど赤色で示している。各表の最左列は被験者群の分類を表しており、楽器経験者の中でもピアノの経験がある者 (9 名) と他の楽器の経験者 (14 名) で分けた類似度を求めている。

まず、楽曲ごとで見ていくと、「ノクターン」と「エリーゼのために」ではどの演奏データに対しても頂点の意図が伝わりにくく、「ノクターン」においてはどの演奏データでも開始音と終了音の意図が伝わりやすいという結果が得られた。「悲愴」については、「映像なし」「映像あり」の間で大きな差はあまり見られない。しかし、演奏者 P2, P3 におけるピアノ経験者群の頂点音のように、「映像あり」になると「映像なし」よりも類似度が下がったものが見られる。

そこで、被験者が回答したアンケート結果と、実験で得られたフレーズ構造の類似度の結果が関連するかを調べるため、2 要因分散分析を行なった。その結果、演奏者 P2 の「悲愴」に対し、演奏データ (映像なし, 映像あり)-ピアノ経験 (あり, なし) 間において、視聴覚の主効果に有意差が見られた (図 7)。この演奏では、ピアノ経験の有無にかかわらず、「映像あり」の方が「映像なし」よりも有意に類似度が低いことが示された。

### 5.3 考察

有意差ありと出た「悲愴」の P2 の演奏においては、「映像あり」の方が「映像なし (音声のみ)」よりも類似度が低い結果となった。これは、演奏者の動きが見えることで、意図がかえって伝わりにくくなってしまったということを表す。演奏者 P2 の身体動作を見てみると、他の 4 人の演奏者に比べて、全体的に演奏時の上体が大きく前後に揺れていたことが確認できた。しかし、その上体の動きとフレーズ構造が呼応しているかということ、その限りではないように見受けられる。例えば、演奏者 P2 が記述した頂点音のある一つでは、体を前に倒して縮こまっており、そ

悲愴

P1-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P1-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.23	0.45	0.12	0.10	全員	0.22	0.46	0.12	0.08
ピアノ経験者		0.26	0.51	0.12	0.14	ピアノ経験者	0.22	0.51	0.12	0.02
他楽器経験者		0.20	0.42	0.12	0.07	他楽器経験者	0.23	0.43	0.13	0.13

P2-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P2-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.47	0.55	0.27	0.58	全員	0.43	0.54	0.19	0.56
ピアノ経験者		0.46	0.54	0.23	0.63	ピアノ経験者	0.40	0.54	0.09	0.56
他楽器経験者		0.47	0.56	0.29	0.55	他楽器経験者	0.45	0.54	0.26	0.56

P3-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P3-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.42	0.42	0.41	0.42	全員	0.47	0.48	0.45	0.48
ピアノ経験者		0.41	0.43	0.36	0.43	ピアノ経験者	0.36	0.43	0.22	0.43
他楽器経験者		0.41	0.41	0.45	0.41	他楽器経験者	0.54	0.51	0.60	0.51

P4-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P4-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.26	0.64	0.07	0.06	全員	0.25	0.62	0.06	0.06
ピアノ経験者		0.23	0.64	0.01	0.03	ピアノ経験者	0.23	0.58	0.06	0.05
他楽器経験者		0.28	0.65	0.11	0.09	他楽器経験者	0.26	0.64	0.06	0.07

P5-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P5-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.30	0.47	0.33	0.11	全員	0.26	0.44	0.26	0.09
ピアノ経験者		0.28	0.51	0.20	0.14	ピアノ経験者	0.28	0.46	0.26	0.11
他楽器経験者		0.32	0.44	0.42	0.09	他楽器経験者	0.25	0.43	0.26	0.07

映像なし条件

映像あり条件

ノクターン

P1-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P1-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.47	0.63	0.15	0.63	全員	0.46	0.63	0.11	0.63
ピアノ経験者		0.47	0.65	0.11	0.65	ピアノ経験者	0.47	0.64	0.14	0.64
他楽器経験者		0.47	0.62	0.19	0.62	他楽器経験者	0.44	0.62	0.09	0.62

P2-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P2-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.52	0.64	0.31	0.63	全員	0.55	0.65	0.35	0.65
ピアノ経験者		0.48	0.64	0.16	0.64	ピアノ経験者	0.55	0.66	0.31	0.67
他楽器経験者		0.55	0.64	0.40	0.62	他楽器経験者	0.55	0.64	0.37	0.63

P3-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P3-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.51	0.68	0.20	0.66	全員	0.56	0.68	0.33	0.67
ピアノ経験者		0.53	0.74	0.13	0.73	ピアノ経験者	0.59	0.68	0.39	0.68
他楽器経験者		0.50	0.64	0.24	0.62	他楽器経験者	0.55	0.68	0.29	0.66

P4-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P4-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.53	0.68	0.22	0.68	全員	0.54	0.68	0.24	0.68
ピアノ経験者		0.51	0.68	0.16	0.68	ピアノ経験者	0.58	0.74	0.25	0.74
他楽器経験者		0.54	0.68	0.28	0.68	他楽器経験者	0.51	0.64	0.23	0.64

P5-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P5-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.52	0.66	0.23	0.67	全員	0.51	0.67	0.19	0.67
ピアノ経験者		0.53	0.66	0.26	0.66	ピアノ経験者	0.54	0.71	0.19	0.71
他楽器経験者		0.51	0.66	0.20	0.67	他楽器経験者	0.49	0.63	0.19	0.63

映像なし条件

映像あり条件

エリーゼのために

P1-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P1-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.30	0.44	0.02	0.44	全員	0.32	0.45	0.04	0.46
ピアノ経験者		0.35	0.52	0.01	0.52	ピアノ経験者	0.32	0.47	0.03	0.47
他楽器経験者		0.27	0.39	0.03	0.38	他楽器経験者	0.31	0.44	0.06	0.45

P2-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P2-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.35	0.45	0.16	0.43	全員	0.35	0.44	0.19	0.43
ピアノ経験者		0.36	0.47	0.14	0.47	ピアノ経験者	0.35	0.44	0.19	0.44
他楽器経験者		0.34	0.43	0.17	0.41	他楽器経験者	0.35	0.45	0.18	0.43

P3-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P3-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.27	0.38	0.04	0.39	全員	0.33	0.47	0.04	0.48
ピアノ経験者		0.27	0.39	0.01	0.39	ピアノ経験者	0.33	0.46	0.05	0.48
他楽器経験者		0.27	0.37	0.06	0.38	他楽器経験者	0.33	0.48	0.04	0.48

P4-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P4-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.33	0.46	0.10	0.44	全員	0.33	0.43	0.14	0.42
ピアノ経験者		0.33	0.45	0.08	0.45	ピアノ経験者	0.35	0.47	0.12	0.47
他楽器経験者		0.33	0.46	0.12	0.43	他楽器経験者	0.31	0.40	0.16	0.38

P5-映像無	ALL	開始音(S)	頂点音(P)	終了音(E)	P5-映像有	ALL	開始音(S)	頂点音(P)	終了音(E)	
全員		0.35	0.42	0.18	0.44	全員	0.37	0.47	0.19	0.46
ピアノ経験者		0.40	0.48	0.24	0.49	ピアノ経験者	0.41	0.50	0.22	0.50
他楽器経験者		0.31	0.39	0.14	0.40	他楽器経験者	0.35	0.45	0.18	0.42

映像なし条件

映像あり条件

図 6 楽器経験のある聴取者を対象としたフレーズ構造の類似度

の音の発音に対して緊張を持って臨んでいるように感じられる。これに対して別の頂点音では、上体は逆に後ろに反らしている様子が確認できた。つまり全ての頂点音に対して、常に同じような身体動作をしているわけではないということである。一方で、演奏音だけを聴くと、強弱の変化を感じることができ、演奏者が頂点音と記述した音符は若

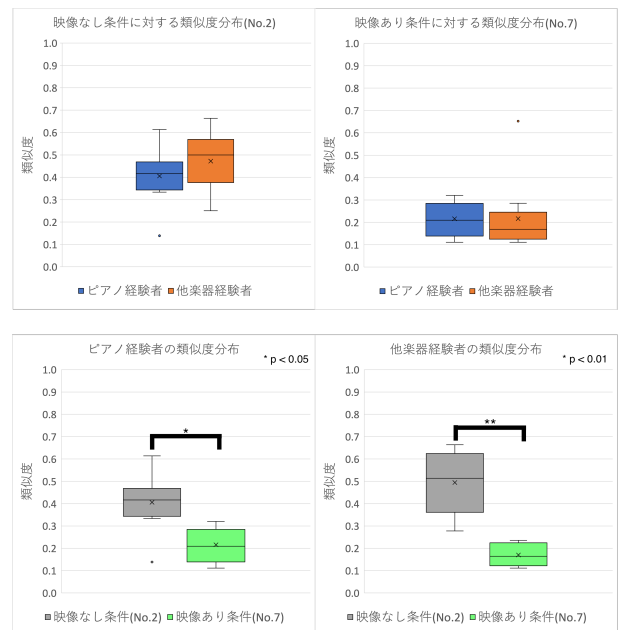


図 7 P2 の「演奏映像の有無」条件における類似度分布

干音量が大きい。この演奏者 P2 の場合、身体動作は、フレーズ構造とは異なる「演奏意図」を意識した結果のものであり、結果として、被験者は演奏者とは異なるフレーズ構造を捉えることになったのかもしれない。

図 6 「悲愴」における演奏者 P3 の頂点音の比較では、ピアノ経験者では「映像あり」の方が類似度が低くなった (0.36 0.22) 一方で、他楽器経験者はむしろ高くなっている (0.45 0.60) ことが確認できる。P3 の身体動作は、演奏中小さく前後に揺れていることが確認でき、P3 が記した頂点音の時は必ず、呼吸をしているためか、上体の特に肩が上がる行動が見られる。しかし、この行動は頂点音以外でもされているため、肩の上昇と頂点音の表現には一貫性がないと言える。しかし、ピアノ経験のある被験者の多数が、肩の上昇と一致するところに頂点音をマークしていることから、ピアノ経験者は演奏者 P3 の身体動作を「きちんと見た」がために、逆に類似度が下がることになったと考えられる。さらに、ピアノ経験者は、自身の経験によって、身体動作と演奏表現の法則性が独自に存在することも一般的に考えられ、それがバイアスとなり、ピアニストの意図とは異なる解釈をしてしまった可能性がある。

4 章、5 章の各実験を踏まえた上での全体的な考察として、本研究ではクラシックピアノのソロ曲のみを取り扱ったが、他の楽器種、ジャンル、構成を変更するとどうなるのかは明らかにはなっていない。例えばピアノではなく、フルートやサクソといった管楽器に変わることによって、楽器種ごとの独特の動きも相まって意図伝達のしやすさが変わっていくのかについて調査する必要がある。また、今回実験で使用した映像データは演奏者の全身が映っているものを採用したが、身体の一部のみを映した映像にすることで意図の伝達度合いが変わるかもしれない。例えば、演奏

者 P1 は他のピアニストに比べて顔の表情の変化が大きく、頂点に指定した音を演奏した時に眉が少し上がるといった様子も見られた。このように、全身像よりも身体の一部を取り上げた方が詳細に演奏意図を読み取れる可能性も否定できない。被験者に提示した演奏動画は、実験時間の制約上一動画につき1分弱程度の長さであり、楽曲の一部分のみを取り扱ったが、演奏者の身体動作から意図を汲み取れるための十分な長さであるかについては課題として調査する必要がある。

本研究では演奏者の意図の一つである「フレーズ構造」をシンボリックに表現させることで情報工学的にデータとして扱えるようにしたが、参加した被験者の中にはフレーズを図式化する行為そのものに慣れていない、あるいはフレーズングという概念を意識したことがないが為に、フレーズ構造を具現化することが難しかったと考えられる。

## 6. まとめ

本研究では、演奏者の意図が聴取者に正確に伝わるかどうかを、演奏表現の一つである「フレーズ構造」に焦点を当て、視聴覚の観点から検討した。実験内での記述における課題点の解決や実験のオンライン化に伴い、実験用の Web システムを作成し、映像の有無による聴取実験を行った。楽器経験の有無にこだわらずに集めた一般聴取者を対象とした実験では、映像の有無による差異は特に見られなかった。しかし、音楽経験者を対象とした実験では、特にピアノの経験者は映像をつけることで類似度が下がったことから、演奏映像は意図伝達における阻害要因になったと言える。これは、視聴覚のモダリティに関する研究の報告とは逆の結果になった。今後は、フレーズ構造の意図伝達に寄与する演奏動作の数え上げと一般性について検討していきたい。

謝辞 本研究は科学研究費補助金(18K18491, 16H02917)による支援を受けて行われた。

## 参考文献

- [1] Nakamura, T.: The Communication of Dynamics between Musicians and Listeners through Musical Performance, *Perception & Psychophysics*, Vol. 41, No. 6, pp. 525–533 (1987).
- [2] Gabrielsson, A. and Juslin, P. N.: Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience, *Psychology of Music*, Vol. 24, No. 1, pp. 68–91 (1996).
- [3] Baraldi, F. B., Poli, G. D. and Rodà, A.: Communicating Expressive Intentions with a Single Piano Note, *Journal of New Music Research*, Vol. 35, No. 3, pp. 197–210 (2006).
- [4] Schubert, E. and Fabian, D.: A Taxonomy of Listeners' Judgements of Expressiveness in Music Performance, pp. 283–303 (2014).
- [5] Héroux, I.: Creative Processes in the Shaping of a Musical Interpretation: A Study of Nine Professional Musicians, *Frontiers in Psychology*, Vol. 9 (2018).

- [6] Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H. and Trainor, L. J.: How Live Music Moves Us: Head Movement Differences in Audiences to Live Versus Recorded Music, *Frontiers in Psychology*, Vol. 9 (2019).
- [7] Senju, M. and Ohgushi, K.: How Are the Players Ideas Conveyed to the Audience?, *Music Perception*, Vol. 4, No. 4, pp. 311–323 (1987).
- [8] Davidson, J. W.: Visual Perception of Performance Manner in the Movements of Solo Musicians, *Psychology of Music*, Vol. 21, No. 2, pp. 103–113 (1993).
- [9] 佐久間 真理, 大串 健吾: 打楽器演奏における演奏者の意図の伝達: 視覚と聴覚の相互作用, *日本音響学会誌*, Vol. 50, No. 8, pp. 613–622 (1994).
- [10] Juchniewicz, J.: The Influence of Physical Movement on the Perception of Musical Performance, *Psychology of Music - PSYCHOL MUSIC*, Vol. 36, pp. 417–427 (2008).
- [11] 橋田 光代, 片寄 晴弘: 演奏表情データベース PEDB 2nd Edition を用いたフレーズ構造聴取に関する初期検討, *研究報告音楽情報科学 (MUS)*, Vol. 2019-MUS-122, No. 21, pp. 1–6 (2019).
- [12] Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*, MIT Press (2004).
- [13] 保科 洋: 生きた音楽表現へのアプローチ: エネルギー思考に基づく演奏解釈法, *音楽之友社* (1998).
- [14] 橋田 光代, 兼口 敦音, 中村 栄太, 古屋 晋一, 小川 容子, 片寄晴弘: ピアニストの演奏解釈を記述した演奏表情データベースの構築, *研究報告音楽情報科学 (MUS)*, Vol. 2017-MUS-116, No. 23, pp. 1–6 (2017).
- [15] ZOOM: <https://zoom.us/>.