

2021年度 博士論文

音の持続時間や立ち上がり時間の増加に伴う Perceptual Attack Time(PAT)の移動に統合の時間窓が与える影響

The Influence of the Auditory Integration Time on Perceptual Attack Time Shifting as the Duration and the Rise Time of a Tone.

2017年度入学

学籍番号 2317917

岡本 悠

Haruka Okamoto

東京藝術大学大学院 音楽研究科 音楽専攻

音楽文化学研究領域 音楽音響創造研究分野

Research Field of Creativity of Music and Sound

Tokyo University of the Arts

提出: 2021年4月28日

修正: 2021年12月24日

# 凡例

- 言語音…言語や音声を用いた音を表す。
- 一般音…言語や音声を用いない音、楽器音や合成音などを表す。
- P-center…言語音の研究領域で用いられる。初めてこの概念が提唱された論文における定義は「心理的な音の発生の瞬間」である。
- Perceptual Attack Time(PAT)…一般音の研究領域で用いられる。

本論での定義は「時間の流れの中に配置された、知覚された音のアタックの瞬間」である。物理的オンセット、知覚的オンセット、最大エネルギーとは異なる。

- 物理的オンセット (physical onset time)…物理的に空気が振動し始めた時刻のことである。
  - 知覚的オンセット (perceptual onset time)…音の強さがその音の最小可聴閾に達した時刻のことである。
  - 最大エネルギー (energy peak time)…音のエネルギーが最大になった時刻のことである。
- SMS 課題 (sensorimotor synchronization task; 感覚運動同期課題)…規則的に提示される刺激に合わせてタッピングする実験課題の総称として用いる。
  - タッピング課題…ある  $A$  音が、メトロノームのように等時間間隔に、数十回繰り返される音列を聴きながら、各音に合うように指でタッピングを行う課題である。SMS 課題のうち、PAT 測定のために行われる、タッピングを用いた課題のことを指す。
  - 同調のズレ量  $Asy(A, i)$ …タッピング課題において算出される。ある  $A$  音が繰り返される音列における、 $i$  回目の  $A$  音へのタップ時刻  $t_{tap}(A, i)$  から  $i$  回目の  $A$  音の物理的オンセット  $t_{onset}(A, i)$  を引いた値のことである。定義式は式 2.3 である。
  - 同調のズレ量の平均値  $\overline{Asy}(A)$ …タッピング課題において算出される。ある  $A$  音が繰り返される音列における、 $i$  回目の  $A$  音へのタップ時刻  $t_{tap}(A, i)$  から  $i$  回目の  $A$  音の物理的オ

ンセット  $t_{\text{onset}}(A, i)$  を引いた値である同調のズレ量  $\text{Asy}(A, i)$  を、タッピングの数  $n$  で平均した値である。定義式は 2.4 である。

- 同時調整課題…ベース音とテスト音がそれぞれ等時間間隔に鳴る音列を聴き、テスト音を「2音が同時に鳴る位置に」動かす課題である。通常、調整前のベース音とテスト音は、聴感上明らかに同時ではない位置に配置される。
- 等時調整課題…ベース音が等時間間隔に鳴り、そのベース音の間にテスト音が鳴る音列を聴き、テスト音を「2音が等時間間隔に鳴る位置に」動かす課題である。通常、調整前のベース音とテスト音は、聴感上明らかに等時間間隔ではない位置に配置される。
- 立ち上がり時間…音の立ち上がり部分の長さ、すなわち物理的オンセットから最大エネルギーまでの時間である。
- 立ち上がり時間の割合…持続時間に占める立ち上がり時間の割合のことである。これは、 $(\text{立ち上がり時間}) / (\text{持続時間}) \times 100[\%]$  で算出される。
- $\Delta\text{AO}_X$ …PAT の数値的指標の一つである。ある  $X$  音における物理的オンセットと PAT との差分を表す。定義式は式 2.1(P.9) である。
- $\Delta\text{PC}_{X,Y}$ …PAT の数値的指標の一つである。ある 2 音 ( $X$  音、 $Y$  音) の  $\Delta\text{AO}$  の差分を表す。定義式は式 2.2(P.9) である。

この数値が  $X$  音の PAT と  $Y$  音の PAT との時間差を、符号が  $X$  音の PAT と  $Y$  音の PAT との時間的な前後関係を表すと解釈できる。

- タッピング課題より算出する  $\Delta\text{PC}_{X,Y}$  の定義式は式 2.7 である。
- 同時調整課題より算出する  $\Delta\text{PC}_{X,Y}(\Delta\text{PC}_{\text{test,base}})$  の定義式は式 2.8 である。
- 等時調整課題より算出する  $\Delta\text{PC}_{X,Y}(\Delta\text{PC}_{\text{test,base}})$  の定義式は式 2.9 である。

# 目次

第1章 序論	1
1.1 研究背景と目的	1
1.2 研究方法	1
1.3 研究射程	2
1.4 本論の構成	2
第2章 先行研究と本論の仮説	5
2.1 “Perceptual Attack Time; PAT”の来歴と定義	5
2.1.1 PATの草分けである言語音領域における“P-center”	5
2.1.2 “P-center”の一般音領域への流入	5
2.1.3 一般音領域におけるPATの定義	6
2.1.4 PATの類似概念との差異	6
2.2 PATの数値化	7
2.2.1 $\Delta AO_X$	8
2.2.2 $\Delta PC_{X,Y}$	9
2.3 PAT測定課題と課題間の差異	10
2.3.1 タッピング課題	10
2.3.2 同時調整課題	17
2.3.3 等時調整課題	19
2.3.4 課題間の差異	20
2.4 PATを変化させる音の要素	23
2.4.1 立ち上がり時間	23
2.4.2 持続時間	27
2.4.3 周波数	28
2.4.4 提示音圧レベル	28
2.4.5 合成音と楽器音	29
2.5 聴覚情報処理における統合の時間窓	33
2.5.1 聴覚情報が保持される「時間範囲」とその範囲分の情報処理による知覚 の変化	34
2.5.2 統合の時間窓の長さ	35

2.6	本論の仮説	42
2.6.1	仮説	43
2.6.2	仮説検証の方法	44
2.7	まとめ	48
<b>第3章</b>	<b>予備実験</b>	<b>49</b>
3.1	予備実験 1: タッピング課題における PAT 測定に必要なタッピングの数の検討	50
3.1.1	実験方法	50
3.1.2	分析	51
3.1.3	結果	53
3.1.4	考察	54
3.2	予備実験 2: 実験課題の比較検討	56
3.2.1	実験方法	56
3.2.2	分析と結果	59
3.2.3	考察	64
3.2.4	本論で用いる PAT 測定課題の検討	67
3.3	予備実験 3: 仮説検証方法の確認並びに本実験条件の検討	68
3.3.1	実験方法	68
3.3.2	分析と結果、考察	69
3.3.3	本実験における刺激音の水準とサンプルサイズの検討	72
3.4	結論	74
<b>第4章</b>	<b>本実験に共通する実験方法と分析方法</b>	<b>76</b>
4.1	実験方法	76
4.1.1	刺激音	76
4.1.2	実験手続き	77
4.1.3	実験装置	78
4.2	分析方法	78
4.2.1	分析の前処理	78
4.2.2	分析方法	79
<b>第5章</b>	<b>本実験 1: 持続時間や立ち上がり時間の割合の増加と PAT の変化との関係と、統合の時間窓</b>	<b>82</b>
5.1	実験方法	82
5.1.1	参加者	82
5.1.2	刺激音	82
5.1.3	実験手続き	83

5.1.4	実験装置	83
5.1.5	実験環境と実験日程	83
5.2	分析と結果	83
5.2.1	生データの確認およびデータ処理	83
5.2.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	89
5.2.3	分散分析	90
5.3	考察	92
5.3.1	生データの確認	92
5.3.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	93
5.3.3	分散分析	93
5.4	結論	94
<b>第 6 章</b>	<b>本実験 2: 統合の時間窓の長さが異なる場合 (音楽訓練の有無) における、持続時間や立ち上がり時間の割合の増加と PAT の変化との関係と、統合の時間窓</b>	<b>96</b>
6.1	実験方法	96
6.1.1	参加者	96
6.1.2	刺激音	97
6.1.3	実験手続き	97
6.1.4	実験装置	97
6.1.5	実験環境と日程	97
6.2	分析と結果	98
6.2.1	生データの確認およびデータ処理	98
6.2.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	104
6.2.3	分散分析	107
6.3	考察	109
6.3.1	生データの確認	109
6.3.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	109
6.3.3	分散分析	110
6.4	結論	110
<b>第 7 章</b>	<b>本実験 3: 統合の時間窓の長さが異なる場合 (異なる周波数) における、持続時間や立ち上がり時間の割合の増加と PAT の変化との関係と、統合の時間窓</b>	<b>112</b>
7.1	実験方法	112

7.1.1	参加者	112
7.1.2	刺激音	113
7.1.3	実験手続き	113
7.1.4	実験装置	114
7.1.5	実験環境と日程	114
7.2	分析と結果	114
7.2.1	生データの確認およびデータ処理	114
7.2.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	121
7.2.3	分散分析	123
7.3	考察	124
7.3.1	生データの確認	124
7.3.2	持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」	124
7.3.3	分散分析	125
7.4	結論	125
<b>第 8 章</b>	<b>全体考察</b>	<b>127</b>
8.1	3つの実験の結果と仮説の検討の結論	127
8.2	「立ち上がり時間の割合」ではなく「立ち上がり時間」の場合の分析結果	132
8.3	持続時間や周波数の変化と PAT の変化との関係	136
8.4	先行研究と本論の実験との比較	140
<b>第 9 章</b>	<b>結論</b>	<b>144</b>
9.1	本論の概略と結論	144
9.2	本論より示された可能性及び今後の研究課題	148
	謝辞	150
	参考文献	151
	付録 A 実験 1: 各参加者の各刺激音における同調のズレ量のバイオリンプロット	154
	付録 B 実験 2: 各参加者の各刺激音における同調のズレ量のバイオリンプロット	157
	付録 C 実験 3: 各参加者の各刺激音における同調のズレ量のバイオリンプロット	160
	付録 D 各実験における測定された $\Delta PC_{A,D100R0}$ の値と近似した 4 パラメータロジスティック回帰モデルのパラメータ値	163

# 第1章 序論

## 1.1 研究背景と目的

音楽においてリズムは重要である。アンサンブルの際にも他者とリズムを合わせる必要であるし、ソロで演奏する際にもリズムは重要である。

まず、アンサンブルで他者とリズムを合わせることを、簡略的に、他者の音と自分の音が鳴ったと知覚された瞬間を合わせることだと考えよう。では、この、他者と合わせる「音が鳴ったと知覚された瞬間」は音のどこにあるだろうか。これは少なくとも、音が物理的に鳴り始めた瞬間ではない。例えば、コントラバスとピアノとで考えてみる。コントラバスの音は立ち上がりがゆっくりであるが、ピアノの音は立ち上がりが急峻である。そのため、コントラバスとピアノの音が物理的に鳴り始めた瞬間を合わせると、「タァタン」とリズムはズレて聴こえ、コントラバスの音の方が遅いと感じられることが想像される。従って、他者の楽器と自分の楽器とで合わせる「音が鳴ったと知覚された瞬間」は、少なくとも音が物理的に鳴り始めた瞬間ではない。同様に、ソロでリズムを演奏することを、自身の「音が鳴ったと知覚された時刻」を演奏したいリズムパターンに配置することだと考えよう。この場合でも「音が鳴ったと知覚された時刻」は音が物理的に鳴り始めた瞬間ではない。このことは、コントラバスのように立ち上がりがゆっくりな楽器の音を想像すれば容易に想像される。

このようなリズムの知覚に関与する「音が鳴ったと知覚された瞬間」は、“Perceptual Attack Time; PAT”と呼ばれる。本論は、PATが音のどこにあるかを調べるものである。

## 1.2 研究方法

このPATを研究するため、本論では心理実験を行う。確かに、音や演奏の研究方法の一つに演奏分析もある。演奏分析では、演奏された楽譜上の音価がどの瞬間に鳴り始め、どの瞬間に鳴り終わったかを測定する。この演奏分析は、演奏において音が鳴り始めた瞬間を明らかにできるため、楽譜上の音価と実際の演奏での音価の違いなどを分析するのに適している。だが、本論のように、知覚されるものを知るためには適さない。知覚は人間の中にしかないため、物理的な分析では明らかにすることができないからだ。音楽のリズムの知覚に関与する「音が鳴った瞬間」すなわちPATを調査するには、やはり心理実験が適している。

## 1.3 研究射程

PAT 知覚は複雑な事象であり、未だ確立した知見が少ない研究領域である。詳細は 2.4 節で述べるが、PAT の先行研究の間でも、音の何の要素が異なる時に PAT が変化するかが一貫して調査されていなかったり矛盾する報告がされたりするからだ。

このように確立した知見が少ない研究領域において複雑な事象から確実な知見を得るには、できるだけ不確定で制御不可能な要素を減らし、確実に制御できる少ない要素を用いて調査する必要がある。すなわち、基礎研究が必要となる。

従って、本論では PAT 知覚に関する基礎研究を行う。具体的には、メトロノームのような等間隔のリズムを用いて、合成音を刺激音として、その持続時間と立ち上がり時間を変化させた時に PAT がどのように変化するかを調査する。

本論により新たに明らかになることは、一つ目に、PAT 知覚において、持続時間と立ち上がり時間が増加するにつれて PAT がどのように変化するかである。これは、持続時間と立ち上がり時間との相互作用を詳細に記述するという点で本論が新規性を持つ事柄である。二つ目に新たに明らかになることは「統合の時間窓」が PAT 知覚に影響を与えているかどうかである。統合の時間窓は聴覚情報処理システムの一つであり、聴覚情報処理システムのうちで PAT の知覚に関与すると考えられるものだ。これは聴覚情報処理システム、特に統合の時間窓と PAT 知覚の関連を調査した研究として本論が新規性を持つ事柄である。

もちろん、基礎研究として実験を単純化することで、検討できないことも生じる。まず、本論では合成音を用いるが、合成音と楽器音との PAT は若干異なる可能性が示唆されている (2.4.5 項参照)。さらに、本論では、実際の音楽ではなく等間隔のリズムを用いて PAT を測定する。だが、等間隔のリズムにおける PAT と実際の音楽のリズムにおける PAT とが異なるかどうかを実際に実験と検証した先行研究はなく、両者の PAT が同じであるという仮説を述べた研究があるのみである (P.18 参照)。以上のような、合成音から楽器音への知見の発展や、等間隔のリズムから実際の音楽で用いられるリズムへの発展は、今後の応用的研究に期待する。

本論ではあくまで基礎研究として、PAT における持続時間と立ち上がり時間との相互作用の詳細と、「統合の時間窓」が PAT 知覚に影響を与えているかどうかを明らかにする。そのため、より音楽に寄った知見や応用研究による知見については今後の研究に期待する。

## 1.4 本論の構成

次章以降の本論の構成は以下の通りである。また、図 1.4.1 に本論の研究フローを示す。

第 2 章では、PAT の先行研究と、聴覚における統合の時間窓の先行研究を述べた。そして、それらの先行研究より本論の仮説を導いた。

第3章では、予備実験を行った。予備実験を通して、PAT測定課題の一つであるタッピング課題を本実験で用いる妥当性を検証した。また、本実験で用いる持続時間や「立ち上がり時間の割合」の水準を検討し、合わせてサンプルサイズも検討した。

第4章では、第5章から第7章で行われる3つの本実験に共通する実験方法と分析方法を述べた。

第5章では、実験1を行った。音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行い、仮説の検証を行った。

第6章では、実験2を行った。仮説に関わる要素である音楽訓練を受けたことがあるかどうか(音楽訓練の有無)を取り上げた。実験1とほぼ同じ刺激音を用いて、音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行い、仮説の検証を行った。

第7章では、実験3を行った。仮説に関わる要素である周波数の違いを取り上げた。周波数を高くした刺激音を用いて、音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行い、仮説の検証を行った。

第8章では、第5章から第7章までの実験結果を踏まえ、先行研究との比較や更なる考察を行った。これを通して、更なる仮説の検証を行った。

第9章では、結論を述べた。

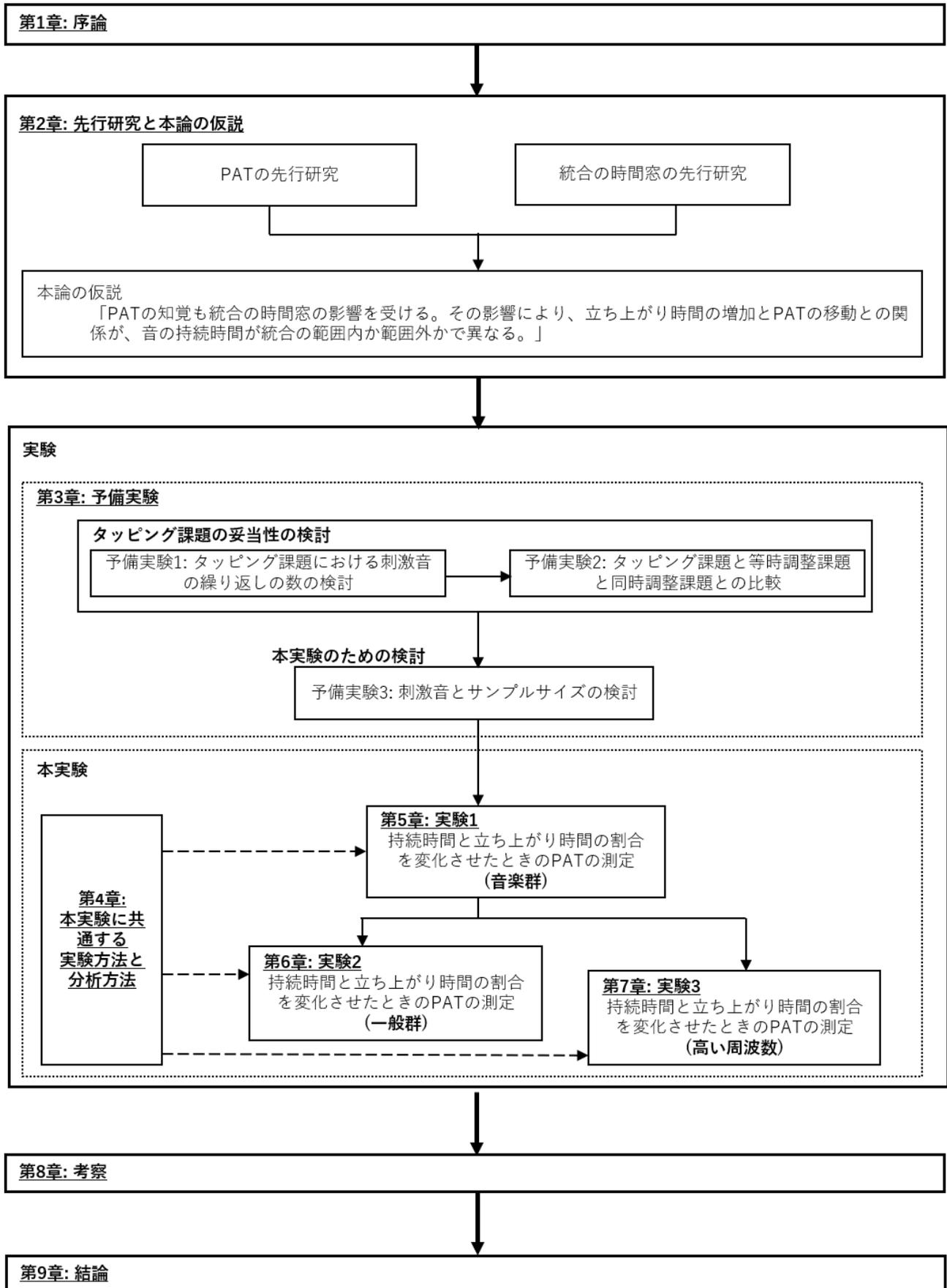


図 1.4.1: 本論の研究フロー

## 第2章 先行研究と本論の仮説

本章では、先行研究について詳述し、本論の仮説を導く。まず、“Perceptual Attack Time; PAT”の先行研究について詳述する。続いて、統合の時間窓の先行研究を詳述する。統合の時間窓は聴覚情報処理システムのうちでPATの知覚に関与すると考えられるものだからだ。その後、本論の仮説を述べる。

### 2.1 “Perceptual Attack Time; PAT”の来歴と定義

本節では、本論で扱う“Perceptual Attack Time; PAT”について述べる。まずは、PATの発端となった研究を述べる。次に、PATの定義を述べる。そして、PATと類似した概念との違いを述べる。

#### 2.1.1 PATの草分けである言語音領域における“P-center”

Morton, Marcus & Frankish[17]は、oneからnineまでの言語音を刺激として、各刺激の時間間隔が知覚的に等間隔に聴こえるように各刺激の位置を調整させた。参加者が調整した後の各刺激の物理的な振幅の開始時刻を確認したところ、各刺激の物理的な振幅の開始時刻は等間隔ではなく、等間隔から最大80ms程度ズレていることを報告した。このことより、「心理的な音の発生の瞬間(“psychological moment of occurrence”)」を“Perceptual center; P-center”と名づけ、P-centerが音のどこにあるのかを更に研究する必要があると主張した。

#### 2.1.2 “P-center”の一般音領域への流入

この“P-center”という用語は、主に、言語音、すなわち言語や音声を用いた音の研究において用いられている。言語音における“P-center”は、後に、一般音、すなわち言語や音声を用いない楽器音や合成音などでも研究されるようになった。

一般音を対象とした研究では、“P-center”ではなく“Perceptual Attack Time; PAT”と呼ばれることが多い。だが、一般音を対象とする研究でもP-centerと称されることがあるため、先行研究では厳密な使い分けはされていない。そこで、本論では言語音を対象とする場合には“P-center”を用い、一般音を対象とする場合には“PAT”を用いる。

### 2.1.3 一般音領域における PAT の定義

一般音の PAT は、言語音の P-center と同じ概念であり、同じく「心理的な音の発生の瞬間」を意味する。だが、PAT の先行研究では、PAT が P-center と同じ概念であると述べられても、PAT の定義として P-center の定義がそのまま引用されることは殆どない。むしろ、リズムや拍節と関連して PAT を定義した、Gordon[8, 9] や Wright[37] による定義が引用されることが多い。

Gordon[8, 9] は、PAT を次のように定義した。

PAT は「音の物理的オンセットと関連して、音のアタックの瞬間や最も目立った拍節的特徴が知覚される時刻のこと」であり、

“the time a tone’s moment of attack or most salient metrical feature is perceived relative to its physical onset”. [8]

「『アタックの瞬間』が知覚された時刻」である。

“as the time its perceptual “moment of attack” occurs” [9]

Wright[37] は、PAT を次のように定義した。

PAT は、「リズム的な配置が知覚された瞬間」であり、

“ PAT is its perceived moment of rhythmic placement.”

「時間の中の滑らかな連続体から、分離した個々の音を知覚する瞬間」である。

“when do we perceive discrete sonic events amidst the smooth continuum of time? This is the issue known as Perceptual Attack Time (“PAT”) or, in the speech community, Perceptual Center (“P-Center”)”

本論でも彼らの定義を参照し、PAT は、「時間の流れの中に配置された、知覚された音のアタックの瞬間」であると定義する。

### 2.1.4 PAT の類似概念との差異

PAT は音の立ち上がり部分にあるとされる。音の立ち上がり部分には、PAT の他にも、物理的オンセット (physical onset time)、知覚的オンセット (perceptual onset time)、最大エネルギー (energy peak time) がある。これらはそれぞれ異なるものである [19, 8, 9, 37, 5, 15, 3]。それぞれの定義を以下にまとめる。音の立ち上がり部分の模式図を図 2.1.1 に示したので、合わせて参照してほしい。ここでは、その他の用語に合わせ、PAT を「知覚的アタック (PAT)」と表記した。

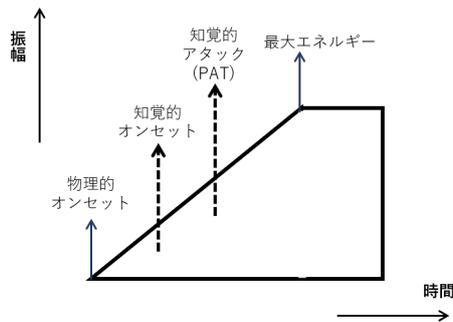


図 2.1.1: 音の立ち上がり部分の模式図 (振幅エンベロープ)

横軸が時間を、縦軸が振幅を表す。音の振幅エンベロープにおける模式図が示された。物理的オンセットと最大エネルギーは、その定義の時刻に実線で示された。一方、知覚的オンセットと知覚的アタック (PAT) は、この例においてこの時刻にあると仮定したため、点線で示された。

- 物理的オンセット (physical onset time) は、物理的に空気が振動し始めた時刻である。生成した合成音においては振幅が 0 以上となった時刻である。録音された音においては暗騒音の振幅を超えた時刻とされることが多い。
- 知覚的オンセット (perceptual onset time) は、音の強さがその音の最小可聴閾に達した時刻である。
- 知覚的アタック (PAT) は、前述の通り、「時間の流れの中に配置された、知覚された音のアタックの瞬間」である。
- 最大エネルギー (energy peak time) は、音のエネルギーが最大になった時刻である。

以上のように、物理的オンセット、知覚的オンセット、PAT、最大エネルギーは異なる概念である。それにもかかわらず、物理的オンセットと知覚的オンセットや、知覚的オンセットと PAT は、文献や研究分野によっては区別されないことがある。例えば、Gordon[8] は、J.Vos & Rasch[34] の用いた “Perceptual onset time” は PAT のことであると指摘した。また、「オンセット」や「アタック」自動検出アルゴリズムはいくつか提案されているが、これらの「オンセット」や「アタック」が物理的なものか知覚的なものかが明記されない場合が多いことも先行研究で指摘された [19, 14]。そのため、これらの研究を参照する際には、その用語が何を示しているかについてよく注意するべきである。

## 2.2 PAT の数値化

PAT を評価したり測定したりするためには、PAT を数値化する必要がある。だが、PAT は人間の知覚の中にあるため、PAT を直接的に測定することはできない。このような PAT の数値化は、理論的に仮定した PAT を別の時刻との関係で表すことで行う。

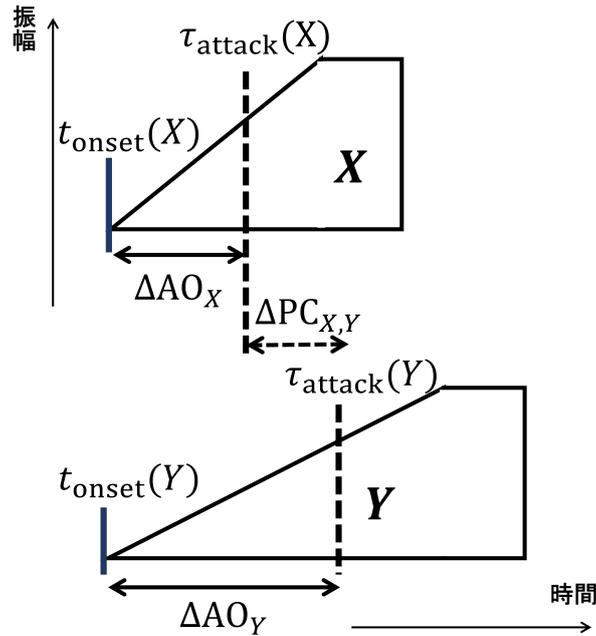


図 2.2.1: X 音と Y 音の模式図 (振幅エンベロープ)

横軸が時間を、縦軸が振幅を表す。音の振幅エンベロープにおける模式図が示された。X 音の物理的オンセットを  $t_{\text{onset}}(X)$  と表記し、Y 音の物理的オンセットを  $t_{\text{onset}}(Y)$  と表記し、図中では各音の左端の実線で示した。また、X 音の音の PAT を  $\tau_{\text{attack}}(X)$  と表記し、Y 音の音の PAT を  $\tau_{\text{attack}}(Y)$  と表記し、図中では破線で示した。この値は理論的に仮定した値であり、直接測定された値ではない。加えて、X 音における物理的オンセットと PAT との差分である  $\Delta AO_X$  と Y 音における物理的オンセットと PAT との差分である  $\Delta AO_Y$ 、並びに、2 音の  $\Delta AO$  の差分である  $\Delta PC_{X,Y}$  も図中に示されている。

本節では、このように PAT を数値化したものとして、次の二つを詳述する。一つ目は、 $\Delta AO_X$  である。これは、ある X 音における物理的オンセットと PAT との差分である。二つ目は、 $\Delta PC_{X,Y}$  である。これは、ある 2 音 (X 音、Y 音) の  $\Delta AO$  の差分である。どちらの場合も差分を表す “ $\Delta$ ” の記号が用いられる。

なお、本節の多くは、Villing, Repp, Ward, & Timoney[32] と、Villing, Ward, & Timoney[33] の論文に依拠している。Villing らは言語音の “P-center” を扱ったため、 $\Delta PC_{X,Y}$  の “PC” は “P-center” の略であると考えられた。本論では一般音の PAT を扱うが、Villing の表記を引用し、 $\Delta PC_{X,Y}$  と表記した。

### 2.2.1 $\Delta AO_X$

先述したように、 $\Delta AO_X$  は、ある X 音における物理的オンセットと PAT との差分である。これについて詳述する。

図 2.2.1 の上段のように、ある X 音があるとする。まず、この X 音の物理的オンセットを、 $t_{\text{onset}}(X)$  と表記し、図中では左端の実線で表した。次に、この音の PAT を  $\tau_{\text{attack}}(X)$  と表記し、破線で表した。破線で表したのは、この  $\tau_{\text{attack}}(X)$  は理論的に仮定した値であり、直接測定された値ではないためである。そして、物理的オンセットと PAT との差分が、差分を表す  $\Delta$  を

用いて、 $\Delta AO_X$  となる。すなわち、 $\Delta AO_X$  は以下の式となる。

$$\Delta AO_X = \tau_{\text{attack}}(X) - t_{\text{onset}}(X) \quad (2.1)$$

### 2.2.2 $\Delta PC_{X,Y}$

先述したように、 $\Delta PC_{X,Y}$  は、ある 2 音 ( $X$  音と  $Y$  音) の  $\Delta AO$  の差分である。これについて詳説する。

図 2.2.1 の全体のように  $X$  音と  $Y$  音との 2 音があるとする。この 2 音の  $\Delta AO$  の差分を、 $\Delta PC_{X,Y}$  と表す。すなわち、 $\Delta PC_{X,Y}$  は以下の式となる。

$$\begin{aligned} \Delta PC_{X,Y} &= (\tau_{\text{attack}}(X) - t_{\text{onset}}(X)) - (\tau_{\text{attack}}(Y) - t_{\text{onset}}(Y)) \\ &= \Delta AO_X - \Delta AO_Y \end{aligned} \quad (2.2)$$

詳細は 2.3 章で詳述するが、この  $\Delta PC_{X,Y}$  は、PAT を測定した実験の結果より算出することができる。

$\Delta PC_{X,Y}$  は、 $X$  音の PAT と  $Y$  音の PAT の関係によって、正の数とも負の数ともなりうる。

$\Delta PC_{X,Y}$  が負の数の場合には (図 2.2.1 の状態)、 $X$  音の  $\Delta AO_X$  よりも  $Y$  音の  $\Delta AO_Y$  の方が大きい ( $\Delta AO_X < \Delta AO_Y$ )。この場合、もし 2 音の物理的オンセット  $t_{\text{onset}}$  が同時であれば ( $t_{\text{onset}}(X) = t_{\text{onset}}(Y)$ )、 $X$  音の PAT が  $Y$  音の PAT よりも時間的前方に位置することになる。言い換えると、もしこの 2 音を物理的オンセットを揃えて鳴らした場合、知覚上は、 $X$  音が  $Y$  音よりも早く聴こえる。

一方、 $\Delta PC_{X,Y}$  が正の数の場合には、 $Y$  音の  $\Delta AO_Y$  よりも  $X$  音の  $\Delta AO_X$  の方が大きい ( $\Delta AO_Y < \Delta AO_X$ )。この場合、もし 2 音の物理的オンセット  $t_{\text{onset}}$  が同時であれば ( $t_{\text{onset}}(X) = t_{\text{onset}}(Y)$ )、 $Y$  音の PAT が  $X$  音の PAT よりも時間的前方に位置することになる。言い換えると、もしこの 2 音を物理的オンセットを揃えて鳴らした場合、知覚上は、 $Y$  音が  $X$  音よりも早く聴こえる。

以上のことを踏まえると、もし 2 音の PAT が同時に知覚されるように 2 音を鳴らしたいならば、 $\Delta PC_{X,Y}$  が正の数であれば  $X$  音を  $\Delta PC_{X,Y}$  の値の分だけ時間的前方へ移動させれば良く、逆に、 $\Delta PC_{X,Y}$  が負の数であれば、 $Y$  音を  $\Delta PC_{X,Y}$  の値の分だけ時間的前方へ移動させれば良い。このようにすると、知覚上、同時に聴こえるようになる。従って、 $\Delta PC_{X,Y}$  は、数値が  $X$  音の PAT と  $Y$  音の PAT との時間差を、符号が  $X$  音の PAT と  $Y$  音の PAT との時間的な前後関係を表すと解釈できる。

## 2.3 PAT 測定課題と課題間の差異

本節では、先行研究で主に用いられる、PAT を測定する実験課題について詳説する。

PAT は知覚されるものであり、人間の心の中にしか存在しないものである。そのため、先行研究で用いられている PAT 測定方法の殆どは間接的な方法である [27, 37]。先行研究で主に用いられる PAT 測定課題は三つある。それらは、タッピング課題、同時調整課題、等時調整課題である<sup>1</sup>。

以下では、まず、各課題の詳細と問題点を述べる。その後、課題間の差異に関して明らかになっていることを述べる。

### 2.3.1 タッピング課題

#### 方法詳細

タッピング課題とは、ある  $X$  音が、メトロノームのように等時間間隔で数十回繰り返される音列を聴きながら、各音に合うように指でタッピングを行う課題である。

この課題の手続きは、SMS 課題 (sensorimotor synchronization task; 感覚運動同期課題) と呼ばれる課題と同一である [23]。参加者のタッピングする時刻が「時間の流れの中」での「その音が鳴った瞬間」、すなわち音の PAT であるため、後述する分析を行うことで PAT を測定することができる。SMS 課題にはいくつかのバリエーションが存在する [23]。例えば、提示される感覚刺激の種類の変異に、聴覚刺激や視覚刺激、触覚刺激などがある。また、刺激の提示回数とタップの回数との関係の変異に、1 回提示された刺激に対し 1 回タップを行う場合や、1 回提示された刺激に対し 2 回タップを行う場合などがある。PAT 測定のためのタッピング課題では、音刺激を用い、1 回提示された音刺激に対し 1 回タップを行う課題が用いられる。

以後、規則的に提示される刺激に合わせてタッピングする課題を総称する場合には「SMS 課題」と称し、そのうちの PAT 測定を目的とした課題のみを表す場合には「タッピング課題」と称する。

#### SMS 課題とタッピング課題における特徴

ここでは、SMS 課題の特徴として先行研究で述べられることのうち、本論に関係した次の三つを述べる。一つ目は、「同調のズレ量」という値とその性質についてである。同調のズレ量は、

<sup>1</sup>それぞれの課題を使用した論文の参考文献番号を以下にまとめる。

- タッピング課題を使用した論文…一般音: [5, 15, 35]、言語音 [12, 32]
- 同時調整課題を使用した論文…一般音: [1, 2, 5, 8, 9, 15, 19, 37]、言語音: —
- 等時調整課題を使用した論文…一般音: [3, 8, 9, 15, 34]、言語音: [17, 27, 32, 33]

SMS 課題において分析の際に算出されるものである。本論のタッピング課題でもこれを算出するため詳述する。二つ目は、同調のズレ量の標準偏差の性質についてである。同調のズレ量の標準偏差は、本論でも算出するため、詳述する。そして三つ目は、SMS 課題のバリエーションの一つである “event onset shift” についてである。これは、Villing ら [32] が PAT を測定するために用いた方法であるため詳述する。

以上の三つは主に、Repp のレビュー論文 [23] と Repp & Su のレビュー論文 [24] を参照した。SMS 課題では様々な感覚刺激で行われることがあるが、以下では音刺激に限って述べる。

まず、一つ目に、「同調のズレ量」という値とその性質について述べる。「同調のズレ量 “asynchrony”」は、ある  $A$  音が繰り返される音列における、 $i$  回目の  $A$  音へのタップ時刻  $t_{\text{tap}}(A, i)$  から  $i$  回目の  $A$  音の物理的オンセット  $t_{\text{onset}}(A, i)$  を引いた値である。本論では “ $\text{Asy}(A, i)$ ” と表記する。式で表すと以下である。

$$\text{Asy}(A, i) = t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i) \quad (2.3)$$

通常、この同調のズレ量をタップした回数  $n$  で平均した「同調のズレ量の平均値  $\overline{\text{Asy}}(A)$ 」を算出し、分析に用いる。

$$\overline{\text{Asy}}(A) = \frac{1}{n} \sum_{i=1}^n (t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i)) \quad (2.4)$$

この同調のズレ量の平均値は、通常、負の数になる。このことは「負のズレ “negative mean asynchrony”」と称され、SMS 課題の大きな特徴として知られている。同調のズレ量の平均値が負の数になるということは、参加者の意識としては音に合わせてタップしていても、実際のタップ時刻は音の物理的オンセットよりも時間的に前方にあることを意味する。また、この同調のズレ量の平均値は、多くの場合、参加者間平均をとると  $-30$ – $-20$  ms 程度であると言われる。だが、参加者ごとに算出すると、参加者によって  $-100$  ms から  $0$  ms 程度まで異なり個人差が大きいことが知られる。さらに、高度な音楽訓練を積んだ音楽家は同調のズレ量の平均値の絶対値が小さいこと、すなわち、音楽訓練によって負のズレが小さくなることも知られている。

二つ目に、同調のズレ量の標準偏差の性質について述べる。1 回提示された音刺激に対し 1 回タップを行う場合、同調のズレ量の標準偏差は、音刺激が繰り返される時間間隔 (inter-onset interval; IOI) の数%となることが知られている。また、標準偏差は音楽訓練によって小さくなることも知られている。音楽訓練を受けた人の標準偏差は IOI の 2–4% 程度であり、また、音楽訓練を受けた人の中で最も標準偏差が小さいのは打楽器奏者であり、打楽器奏者の標準偏差は IOI の 0.5% 程度になると報告される。

三つ目に、SMS 課題のバリエーションの一つである “event onset shift” について述べる。これは、“Perturbation” という音刺激を繰り返す時間間隔を「変動 “variation”」させて提示しタッ

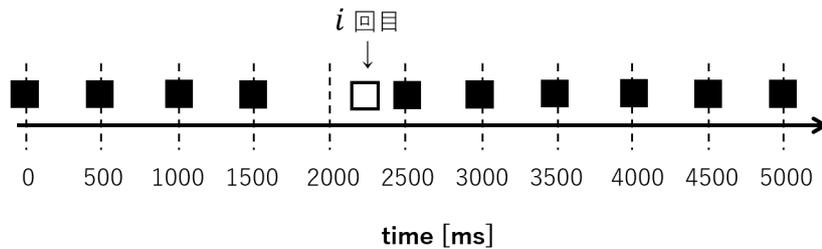


図 2.3.1: “event onset shift”の例

横軸は時間を表している。繰り返される音が黒塗りの四角で示された。変動させる  $i$  回目の音が白抜き四角で示された。この例では、音 (黒塗りの四角) が 500 ms ごとに繰り返される中で、 $i$  回目の音 (白抜き四角) の物理的オンセットを時間的後方へ約 200 ms 程度移動させられた。この移動に伴い、 $i-1$  回目の音 (黒塗りの四角) と、 $i$  回目の音 (白抜き四角) との時間間隔が 700 ms 程度となり、同様に、 $i$  回目の音 (白抜き四角) と、 $i+1$  回目の音 (黒塗りの四角) との時間間隔が 300 ms 程度となった。

ピングさせる課題のひとつである。“event onset shift”は、Villing ら [32] が PAT を測定するために用いた方法である。この方法は、同じ時間間隔で音が繰り返される中で、 $i$  回目の音の物理的オンセットだけを時間的前方あるいは後方へ移動させる方法である。従って、もし  $i$  回目の音の物理的オンセットを時間的後方へ移動させたならば、 $i-1$  回目の音と  $i$  回目の音との時間間隔が長くなり、 $i$  回目の音と  $i+1$  回目の音との時間間隔は短くなる。図 2.3.1 に、この “event onset shift” の例を示した。図中では、 $i$  回目の音のみ白抜き四角で表され、その他の音は黒塗りの四角で表された。

### PAT 測定のためにタッピング課題を用いた研究

ここからは、どのように PAT 測定のためにタッピング課題が用いられるようになったかを述べる。

最初にタッピング課題を PAT 測定に用いたのは、P.Vos, Mates & van Kruysbergen [35] だ。P.Vos らは「ターゲット音の PAT が移動した分、タッピング課題における同調のズレの平均量も変化する」という仮説のもと、タッピング課題を用いた。その後、Janker [12] も、P.Vos らの仮説のもと、PAT 測定にタッピング課題を用いた。

そして、Villing, Repp, Ward, & Timoney [32] が実験結果より P.Vos らの仮説の確からしさを報告した。Villing らは、言語音の P-center を測定するため、等時調整課題と、“event onset shift”を用いたタッピング課題を行った。後者の課題で提示された音列は、音がある時間間隔 (IOI) で繰り返されるが、何音かに一つの音は、そのオンセットが時間的に移動させられていた。彼らはこの方法を “phase correction response; PCR” と称した。オンセットを移動させる時間は、音が繰り返される時間間隔 (IOI) の  $\pm 15\%$  以内とされた。そのため、オンセットを移動させる時間が IOI の 0% である場合は通常のタッピング課題になる。この場合に得られた結果が等時調整課題の結果と一貫していたことから、Villing らは、P.Vos らや Janker の仮説が間接的に証

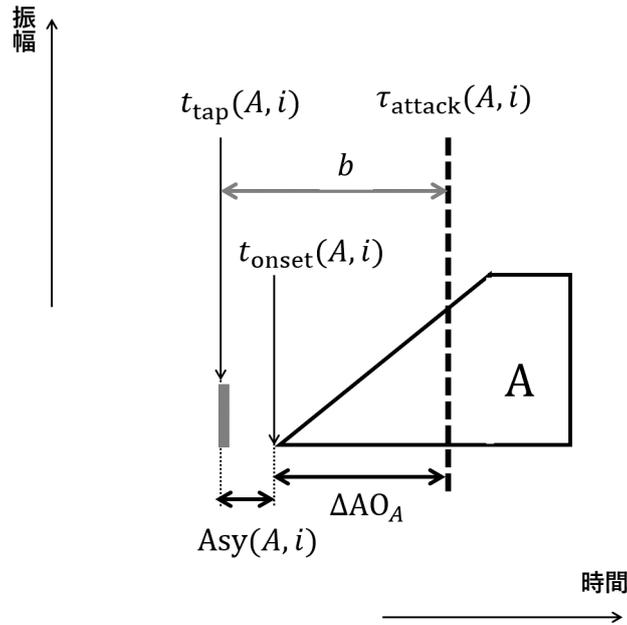


図 2.3.2: ある  $A$  音が繰り返された音列における、 $i$  回目の  $A$  音の同調のズレ量  $Asy(A, i)$  と  $i$  回目の  $A$  音の PAT との関係図 (振幅エンベロープ)  
 横軸が時間を、縦軸が振幅を表す。タップが灰色の長方形のシンボルで示された。 $i$  回目の  $A$  音とタップとの振幅エンベロープにおける模式図に、タップ時刻  $t_{\text{tap}}(A, i)$ 、物理的オンセット  $t_{\text{onset}}(A, i)$ 、同調のズレ量  $Asy(A, i)$ 、PAT  $\tau_{\text{attack}}(A, i)$ 、バイアス  $b$ 、 $\Delta AO_A$  が示された。

明されたと報告した。

### $\Delta PC_{X,Y}$ の算出方法

タッピング課題から  $\Delta PC_{X,Y}$  を算出する方法を述べる [32, 33]。ここからは、刺激音からある  $A$  音と別の  $B$  音を抽出した場合の、 $A$  音と  $B$  音との  $\Delta PC_{A,B}$  を算出することを考える。

まず、ある  $A$  音が繰り返された音列における  $i$  回目の  $A$  音の同調のズレ量  $Asy(A, i)$  と、その  $i$  回目の  $A$  音の PAT との関係を図 2.3.2 に、 $Asy(A, i)$  と PAT との関係図を示したので、合わせて参照してほしい。この図では、タップを灰色の長方形のシンボルで示している。

先述した通り、タッピング課題では、参加者は PAT  $\tau_{\text{attack}}(A, i)$  を狙ってタップする。それにもかかわらず、同調のズレ量  $Asy(A, i)$  は負の値となることが知られている。

このことより、「ある  $A$  音が繰り返される音列における  $i$  回目の  $A$  音へのタップ時刻  $t_{\text{tap}}(A, i)$  は、 $i$  回目の  $A$  音の物理的オンセット  $t_{\text{onset}}(A, i)$  よりも、時間的前方にある」といえる。

$$t_{\text{tap}}(A, i) < t_{\text{onset}}(A, i)$$

さらに、理論的に、この物理的オンセット  $t_{\text{onset}}(A, i)$  は、 $A$  音の PAT  $\tau_{\text{attack}}(A, i)$  よりも時間的

に前方にあることが明らかである。

$$t_{\text{onset}}(A, i) < \tau_{\text{attack}}(A, i)$$

そのため、上記のことを言い換えると、「ある  $A$  音が繰り返される音列における  $i$  回目の  $A$  音へのタップ時刻  $t_{\text{tap}}(A, i)$  は、タップしようとしたその  $A$  音の PAT  $\tau_{\text{attack}}(A, i)$  よりも時間的前方にある」といえる。

$$t_{\text{tap}}(A, i) < \tau_{\text{attack}}(A, i)$$

そうであれば、タップ時刻  $t_{\text{tap}}(A, i)$  は、タップしようとした PAT  $\tau_{\text{attack}}(A, i)$  にバイアス  $b$  が付与されたものであると考えることができる。

$$t_{\text{tap}}(A, i) = \tau_{\text{attack}}(A, i) - b$$

そして、この式は以下のように展開することができる。

$$\begin{aligned} t_{\text{tap}}(A, i) &= \tau_{\text{attack}}(A, i) - b \\ t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i) &= \tau_{\text{attack}}(A, i) - b - t_{\text{onset}}(A, i) \\ t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i) &= \tau_{\text{attack}}(A, i) - t_{\text{onset}}(A, i) - b \\ \text{Asy}(A, i) &= \Delta\text{AO}_A - b \end{aligned} \tag{2.5}$$

すなわち、ある  $A$  音が繰り返された音列における  $i$  回目の同調のズレ量  $\text{Asy}(A, i)$  が、その  $A$  音の  $\Delta\text{AO}_A$  にバイアス  $b$  が付与されたものと考えることができる。

そして、ある  $A$  音が  $n$  回繰り返される音列の中で、繰り返された  $n$  回の  $A$  音へのバイアス  $b$  が一定であると仮定すると、以下のように、同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  もその  $A$  音の  $\Delta\text{AO}_A$  にバイアス  $b$  が付与されたものと考えることができる。

$$\begin{aligned} \overline{\text{Asy}}(A) &= \frac{1}{n} \sum_{i=1}^n (t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i)) \\ &= \frac{1}{n} \sum_{i=1}^n (\tau_{\text{attack}}(A, i) - t_{\text{onset}}(A, i) - b) \\ &= \tau_{\text{attack}}(A) - t_{\text{onset}}(A) - b \\ &= \Delta\text{AO}_A - b \end{aligned} \tag{2.6}$$

先述した通り、繰り返された  $n$  回の  $A$  音へのバイアス  $b$  が一定であると仮定した。さらに、同一の参加者であれば、 $n$  回繰り返される  $A$  音の音列におけるバイアス  $b$  と、 $n$  回繰り返される別の  $B$  音の音列におけるバイアス  $b$  がどちらも一定であると仮定すると、 $A$  音と  $B$  音の PAT の差である  $\Delta\text{PC}_{A,B}$  を以下のように求めることができる。図 2.3.3 に、 $A$  音の同調のズレ量の平

均値  $\overline{\text{Asy}}(A)$  と  $B$  音の同調のズレ量の平均値  $\overline{\text{Asy}}(B)$  と  $\Delta\text{PC}_{A,B}$  との関係図を示したので、合わせて参照してほしい。この図でも、タップを灰色の長方形のシンボルで示している。

$$\begin{aligned}\overline{\text{Asy}}(A) - \overline{\text{Asy}}(B) &= (\Delta\text{AO}_A - b) - (\Delta\text{AO}_B - b) \\ &= \Delta\text{AO}_A - \Delta\text{AO}_B \\ &= \Delta\text{PC}_{A,B}\end{aligned}\tag{2.7}$$

すなわち、ある  $A$  音が  $n$  回繰り返された音列に対する同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  と、別の  $B$  音が  $n$  回繰り返された音列に対する同調のズレ量の平均値  $\overline{\text{Asy}}(B)$  との差より、 $\Delta\text{PC}_{A,B}$  が算出される。

先述した通り、PAT は直接的に測定できず、理論的に仮定し、他の時刻との関係で表すことしかできない。だが、刺激音の各音の PAT  $\tau_{\text{attack}}$  を知ることはできなくとも、 $A$  音と  $B$  音との PAT の差分、すなわち  $\Delta\text{PC}_{A,B}$  を知ることはできる。また、2 音以上の刺激音の PAT の関係を知りたいければ、ある音を基準音  $R$  として、基準音  $R$  以外の全ての音  $A$  との  $\Delta\text{PC}_{A,R}$  を算出すれば、全ての音の PAT の関係を知ることができる。

## 問題点

タッピング課題の問題点は、以下の四つが挙げられる。

一つ目に、タッピングの測定に、タッピングを記録する装置の、時間分解能 (時間の計測における精度) や遅延時間、感度が影響することが挙げられる。この問題点は、Wright[37] によって指摘された。

二つ目に、同調のズレ量の個人差が大きいことが挙げられる。先述した通り、同調のズレ量の平均値は個人差が大きく、参加者間で同調のズレ量の平均値が  $-100-0\text{ms}$  程度まで異なることが知られている。この問題点は、Wright[37] によって指摘された。

三つ目に、同調のズレの平均量を算出する際にも、 $\Delta\text{PC}_{A,B}$  を算出する際にも、個人内のタッピングに含まれるバイアス  $b$  が一定であると仮定しているが、このバイアス  $b$  が一定であると証明できないことが挙げられる。この問題点は、Wright[37] によって指摘された。

四つ目に、先行研究では指摘されていないが、PAT 測定のためのタッピング課題の問題点として、ある  $A$  音の PAT を得るために必要な、音列内で  $A$  音が繰り返される数 (タッピングの数) が明らかとなっていないことが挙げられる。先行研究ではこれを検討せず、London, Nymoan, Langerød, Thompson, Code, & Danielsen[15] はタッピングの数を 33 回として行い、P.Vos ら [35] はタッピングの数を 110 回として行った。この問題は、本論の 3.1 節の予備実験で検討される。

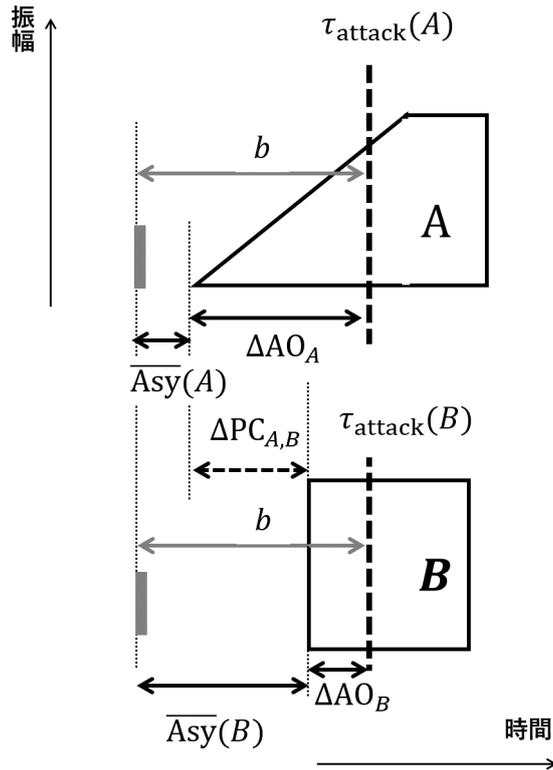


図 2.3.3: A 音の同調のズレ量の平均値  $\overline{Asy}(A)$  と B 音の同調のズレ量の平均値  $\overline{Asy}(B)$  と  $\Delta PC_{A,B}$  との関係図 (振幅エンベロープ)

横軸が時間を表し、縦軸が振幅を表す。タップが灰色の長方形のシンボルで示された。A 音と B 音の振幅エンベロープにおける模式図が、A 音の PAT  $\tau_{\text{attack}}(A)$  と B 音の PAT  $\tau_{\text{attack}}(B)$  を揃えて示された。この模式図に、A 音の  $\Delta AO_A$ 、A 音の同調のズレ量の平均値  $\overline{Asy}(A)$ 、A 音へのバイアス  $b$  と、B 音の  $\Delta AO_B$ 、B 音の同調のズレ量の平均値  $\overline{Asy}(B)$ 、B 音へのバイアス  $b$ 、そして  $\Delta PC_{A,B}$  が示された。

## 2.3.2 同時調整課題

### 方法詳細

同時調整課題とは、ベース音とテスト音がそれぞれ等時間間隔に鳴る音列を聴き、テスト音を「2音が同時に鳴る位置に」動かす課題である。通常、調整前のベース音とテスト音は、聴感上明らかに同時ではない位置に配置される。

参加者は、2音のPATが同時になるように調整する。

### $\Delta PC_{X,Y}$ の算出方法

同時調整課題より、 $\Delta PC_{X,Y}$  を算出する方法を述べる。ここでは、ペアで調整させたベース音 base とテスト音 test の  $\Delta PC_{\text{test,base}}$  を算出することを考える。

同時調整課題では、ベース音とテスト音の PAT  $\tau_{\text{attack}}$  が同時になるように調整される。従って、調整後のベース音の PAT  $\tau_{\text{attack}}(\text{base})$  とテスト音の PAT  $\tau_{\text{attack}}(\text{test})$  は同じ時刻であると考えられる。

$$\tau_{\text{attack}}(\text{base}) = \tau_{\text{attack}}(\text{test})$$

これを用いると、 $\Delta PC_{\text{test,base}}$  は、その定義式より、以下のように展開することができる。

$$\begin{aligned}\Delta PC_{\text{test,base}} &= \Delta AO_{\text{test}} - \Delta AO_{\text{base}} \\ &= (\tau_{\text{attack}}(\text{test}) - t_{\text{onset}}(\text{test})) - (\tau_{\text{attack}}(\text{base}) - t_{\text{onset}}(\text{base})) \\ &= (\tau_{\text{attack}}(\text{test}) - t_{\text{onset}}(\text{test})) - (\tau_{\text{attack}}(\text{test}) - t_{\text{onset}}(\text{base})) \\ &= t_{\text{onset}}(\text{base}) - t_{\text{onset}}(\text{test})\end{aligned}\tag{2.8}$$

すなわち、調整後のベース音の物理的オンセットとテスト音の物理的オンセットとの差分が、 $\Delta PC_{\text{test,base}}$  となる。

### 問題点

同時調整課題の問題点は、以下の五つが挙げられる。

一つ目の問題点は、マスキングや音の融合が生じることである。調整していく過程でテスト音とベース音の時間間隔が短くなると、マスキングが生じて音のラウドネスが異なったり、音の融合が生じてスペクトルや音色が異なったり一つの音に聴こえたりする。また、音の融合は更なる問題を生じさせる。持続時間が短い音と長い音のペアを同時になるように調整させると、調整後は、理論的には、持続時間が短い音よりも長い音の方が時間的前方に位置するはずである。だが、Wright[37] は、調整後を確認したら、持続時間が長い音よりも短い音が時間的前方

に位置し、音によっては、短い音が鳴り終わってから長い音が鳴り始めるように調整されたものもあったと報告した。このような矛盾が生じた理由を、Wright は実験参加者が音の融合を避けた結果であると推察した。

二つ目の問題点は、ある 2 音のうち、ベース音をどちらの音にするかによって、得られる  $\Delta PC_{\text{test,base}}$  が異なることである。A 音と B 音があったとすると、どちらの音の  $PAT_{\tau_{\text{attack}}}$  も他の音や状況によって変化しないはずであり (これは PAT の「文脈独立性仮説」と呼ばれる<sup>2</sup>)、 $\Delta AO_A$  も  $\Delta AO_B$  も他の音や状況によって変化しないはずである。従って、理論的には、以下の式のように、A 音と B 音のペアを調整させる場合に、ベース音を B 音とした時に得られる  $\Delta PC_{A,B}$  と、ベース音を A 音とした時に得られる  $\Delta PC_{B,A}$  は、符号は逆転するが、絶対値は同じになる。

$$\begin{aligned}\Delta PC_{A,B} &= \Delta AO_A - \Delta AO_B \\ &= -(\Delta AO_B - \Delta AO_A) \\ &= -\Delta PC_{B,A}\end{aligned}$$

それにもかかわらず、Danielsen, Nymoen, Anderson, Câmara, Langerød, Thompson, & London[5] は、クリック音とノイズ音のペアにおいて、ベース音をクリック音にするかそれともノイズ音にするかによって、得られた  $\Delta PC_{\text{test,base}}$  に 20 ms の差があり、その差が有意であったと報告した。

三つ目の問題点は、ベース音によって調整の難易度と  $\Delta PC_{\text{test,base}}$  の標準偏差が異なることである。ベース音に、クリック音やドラム音のような立ち上がりが急峻な音を用いると、立ち上がりが緩やかな音を用いるよりも  $\Delta PC_{\text{test,base}}$  の標準偏差が大きく、加えて、より調整が難しいという参加者の意見があったと報告された [8, 17]。また、Wright[37] は適切なベース音について調査し、最も得られる結果のばらつきが小さいのは、ベース音がテスト音とスペクトルが似ている短い音 (“Spectrally Matched Click”) の時であったと報告した。

四つ目の問題点は、テスト音の持続時間や立ち上がり時間によって、得られる  $\Delta PC_{\text{test,base}}$  の標準偏差が異なることである。例えば、Camp, Kilchenmann, Volken, & Senn[2] は、音の立ち上がり時間と得られた  $\Delta PC_{\text{test,base}}$  の標準偏差の相関係数が  $r = .92, p < .001$  であったことから、立ち上がり時間が長ければ長いほど PAT の標準偏差が増えると報告した。Nymoen, Danielsen, & London[19] は、持続時間が長く立ち上がり時間も長い音の方が、持続時間が短く立ち上がり時間も短い音よりも  $\Delta PC_{\text{test,base}}$  の標準偏差が大きかったと報告した。Polfreman[22] は持続時間が長い音の方が持続時間が短い音よりも標準偏差が大きかったと報告した。

五つ目の問題点は、調整中の音の知覚の変化により、2 音の PAT が同時か否かの判断が難しくなることである。この調整中の音の知覚の変化は、時間間隔の長さの変化と、2 音全体の印

<sup>2</sup> PAT の「文脈独立性仮説 “context independence hypothesis”」とは、「P-center は文脈に依存しない、すなわち、時間的に前後にある音が持つタイミングや音響的特徴などの影響を受けない。各音に固定されたものである」という仮説である。Morton らによって提唱された [17]。これはまだ、実際の実験により明らかにされてはいない。

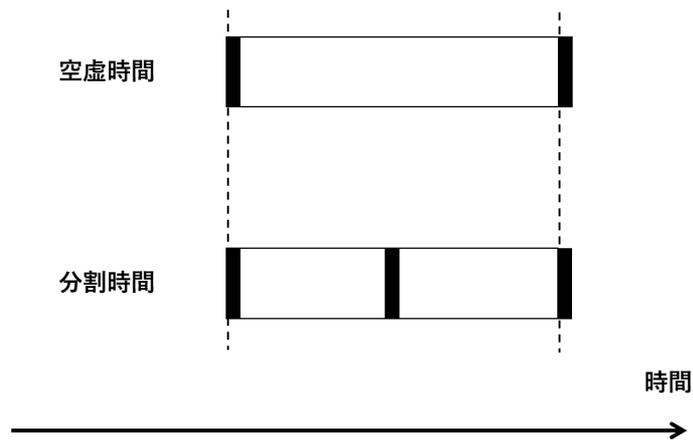


図 2.3.4: 時間間隔の長さの知覚

象の変化が考えられる。まず、調整中に参加者の注意がどこに向くかによって、2音の時間間隔の長さの知覚が変化する。図 2.3.4 に示した通り、2音の時間間隔の長さの知覚には「空虚時間」と「分割時間」の2種類がある。「空虚時間」とは、2音によって両端が区切られ、その間に刺激音が含まれない場合の時間間隔である。「分割時間」とは、2音によって両端が区切られるが、その間に刺激音が含まれる場合の時間間隔である。物理的には同じ時間間隔でも、知覚上は、後者の方が時間間隔が長いことが知られている [41]。同時調整課題では、調整前の「ベース音-テスト音-ベース音」の組み合わせを一つとして注意を向けると分割時間になるが、「テスト音-ベース音」を一つとして注意を向ければ空虚時間になる。このように注意の向け方によって時間間隔が異なって知覚される。次に、調整に伴い2音の時間間隔が変化すると、2音全体の印象が変化する。調整前にはベース音とテスト音との時間間隔がある程度あるが、同時に調整する過程で、2音の時間間隔は徐々に縮まっていく。音と音の時間間隔によって知覚が変化することが知られており、時間間隔が数百 ms 以上あると2音それぞれの音の要素を同定でき順序を知覚することができるが、数百 ms から 20 ms までは全体的な印象に基づいた順序判断しかできなくなり、2 ms 以下になれば単音か2音かの弁別さえもできなくなると言われている [10, 11, 36]。以上のように、調整中に、時間間隔の長さの知覚が変化したり、2音全体の印象の知覚が変化したりすることで、その知覚に気を取られ、2音の PAT が同時か否かの判断が難しくなることも問題である。これらのことを Gordon[9] も示唆している。

### 2.3.3 等時調整課題

#### 方法詳細

等時調整課題とは、ベース音が等時間間隔に鳴り、そのベース音の間にテスト音が鳴る音列を聴き、テスト音を「2音が等時間間隔に鳴る位置に」動かす課題である。通常、調整前のベース音とテスト音は、聴感上明らかに等時間間隔ではない位置に配置される。

参加者は、2音の PAT が等時間間隔に聴こえるように調整する。

## $\Delta PC_{A,B}$ の算出方法

等時調整課題より、 $\Delta PC_{X,Y}$  を算出する方法を述べる。ここでは、ペアで調整させたベース音 base とテスト音 test の  $\Delta PC_{\text{test,base}}$  を算出することを考える。

等時調整課題では、ベース音とテスト音の PAT  $\tau_{\text{attack}}$  は、等時間間隔になるように調整される。従って、調整後のベース音の PAT  $\tau_{\text{attack}}(\text{base})$  とテスト音の PAT  $\tau_{\text{attack}}(\text{test})$  は、ベース音とベース音の時間間隔の半分、すなわち、ベース音とテスト音の理論的な時間間隔 (idealIOI) の分だけ異なると考えられる。

$$\tau_{\text{attack}}(\text{test}) - \tau_{\text{attack}}(\text{base}) = \text{idealIOI}$$

これを用いると、 $\Delta PC_{\text{test,base}}$  は、その定義式より、以下のように展開することができる。

$$\begin{aligned} \Delta PC_{\text{test,base}} &= \Delta AO_{\text{test}} - \Delta AO_{\text{base}} \\ &= (\tau_{\text{attack}}(\text{test}) - t_{\text{onset}}(\text{test})) - (\tau_{\text{attack}}(\text{base}) - t_{\text{onset}}(\text{base})) \\ &= (\tau_{\text{attack}}(\text{test}) - t_{\text{onset}}(\text{test})) - (\tau_{\text{attack}}(\text{test}) - \text{idealIOI} - t_{\text{onset}}(\text{base})) \\ &= -t_{\text{onset}}(\text{test}) - (-\text{idealIOI} - t_{\text{onset}}(\text{base})) \\ &= t_{\text{onset}}(\text{base}) - t_{\text{onset}}(\text{test}) + \text{idealIOI} \end{aligned} \quad (2.9)$$

すなわち、調整後のベース音の物理的オンセットとテスト音の物理的オンセットの差分に、理論的なベース音とテスト音の時間間隔を加えたものが、 $\Delta PC_{\text{test,base}}$  となる。

## 問題点

等時調整課題の問題点は、調整中に参加者の注意がどこに向くかによって、2音間の時間間隔の長さの知覚が変化することにより2音の PAT の時間間隔が等しいか否かの判断が難しくなることである。これは、先述した同時調整課題の問題点の五つ目として挙げたものである。先述したことと同様であるため、本項では割愛する。

### 2.3.4 課題間の差異

今まで述べてきたように、先行研究で PAT を測定するために用いられる実験課題は、タッピング課題、同時調整課題、等時調整課題である。本項では、それぞれの課題から得られる PAT を比較した研究を取り上げる<sup>3</sup>。以下では、まず、課題間で得られる値が統計的に異なるのか

<sup>3</sup>本項で述べた研究を、それらの研究が比較した課題ごとに列挙する。

タッピング課題と同時調整課題との比較研究

– Danielsen, Nymoen, Anderson, Câmara, Langerød, Thompson, & London[5]

についての報告を述べる。次に、統計的有意ではなく差異がある傾向が示されたという報告を述べる。

#### 2.3.4.1 課題間の有意差

課題間で得られる値が統計的に異なるのかについて報告されたことを述べる。

まず、タッピング課題と等時調整課題とで、得られる値が有意に異なることが Villing, Repp, Ward, & Timoney[32] によって報告された。Villing らは、本論の 12 ページで述べた通り、等時調整課題と “event onset shift” を用いたタッピング課題を行った結果、通常のタッピング課題と同様になる条件の結果と等時調整課題との結果に有意差がなかったことを報告した。

次に、等時調整課題と同時調整課題とで、得られる値が有意に異なることが、London, Nymoen, Langerød, Thompson, Code, & Danielsen[15] と Danielsen, Nymoen, Anderson, Câmara, Langerød, Thompson, & London[5]、そして Nymoen, Danielsen, & London[19] によって報告された。London らは実験 1 にて、タッピング課題、同時調整課題、等時調整課題の比較検討を行った。分析にて、同時調整課題と等時調整課題から  $\Delta PC_{\text{test,base}}$  を、タッピング課題からは同調のズレ量  $\overline{Asy}(\text{base})$  を算出した<sup>4</sup>。それらの値を比較した結果、等時調整課題と同時調整課題の値の差は有意ではなかったと報告した。そして、この London らの実験結果を詳細に分析した Danielsen ら [5] も、同様の報告をした。さらに、Nymoen ら [19] は、等時調整課題と同時調整課題との比較実験を小規模に行った結果、両者より得られた値に有意差がなかったと報告した。

---

#### タッピング課題と等時調整課題との比較研究

- Villing, Repp, Ward, & Timoney[32]

#### 同時調整課題と等時調整課題との比較研究

- Nymoen, Danielsen, & London[19](下記の London ら [15] の実験 1 よりタッピング課題と同時調整課題に限り詳細な分析を行い、更なる実験を行った研究)
- Gordon[8, 9]

#### タッピング課題と同時調整課題と等時調整課題 (3つの課題) の比較研究

- London, Nymoen, Langerød, Thompson, Code, & Danielsen[15]

<sup>4</sup>London らや Danielsen らが 3つの課題より算出した値 (“the P-center location”) の算出式は掲載されていない。だが、以下の記述より、同時調整課題と等時調整課題では、ベース音の物理的オンセットとテスト音の物理的オンセットの差、すなわち  $\Delta PC_{\text{test,base}}$  を算出したと推察された。

“probe locations are reported in milliseconds relative to the physical onset of the stimulus.” London[15]P.5

一方、タッピング課題では、以下の記述より、タップ時刻と刺激音の物理的オンセットの差、すなわち「同調のズレ量」を算出したと推察された。

“For each registered tap, the time difference between its detected onset and the first zero crossing of the closest stimulus sound was calculated.” London[15]P.5

一方、タッピング課題と同時調整課題とで得られる値が有意に異なると London ら [15] と Danielsen ら [5] が報告した。London らは、先述した通り、実験 1 にて、タッピング課題、同時調整課題、等時調整課題の比較検討を行った。分析にて、同時調整課題と等時調整課題から  $\Delta PC_{\text{test,base}}$  を、タッピング課題からは同調のズレ量  $\overline{Asy}(\text{base})$  を算出し、比較した結果、タッピング課題と 2 つの調整課題とから算出された値に有意な差があったと報告した。

だが、London らと Danielsen らは、タッピング課題における同調のズレ量  $\overline{Asy}(\text{base})$  と、同時調整課題と等時調整課題における  $\Delta PC_{\text{test,base}}$  という異なる値を比較していた。そのため、彼らの有意差があるという報告は、タッピング課題と 2 つの調整課題より得られる PAT が異なることの根拠にはならないと考えられる。むしろ、他の先行研究で報告されたように、タッピング課題と等時調整課題とで、また、等時調整課題と同時調整課題とで得られる値に有意差がないのであれば、タッピング課題と同時調整課題とで得られる値に有意差がない可能性が考えられる。

以上のように、先行研究より、3 つの課題から得られる PAT に有意差がない確証は得られていない。だが、3 つの課題を  $\Delta PC_{\text{test,base}}$  という一つの値で比較すれば、3 つの課題間に有意差がない可能性が高いと考えられる。これについては、本論の 3.2 節の予備実験にて検討する。

#### 2.3.4.2 課題間の差異傾向

次に、課題間で得られる値に、差異がある傾向が示されたとの報告をまとめる。

傾向として報告された差異は次の四つである。

一つ目に、課題間で得られる標準偏差が異なる傾向があると報告された [15, 5, 9]。London らは、実験 1 として 3 つの課題の同時比較を行った結果、同時調整課題で得られる  $\Delta PC_{\text{test,base}}$  の標準偏差が最も大きく、タッピング課題で得られる同調のズレ量の平均値  $\overline{Asy}(\text{test})$  の標準偏差が最も小さかったと報告した。Danielsen らも London らの実験 1 を再度詳細に分析し、同様のことを報告した。Gordon は、同時調整課題と等時調整課題を行った。その結果、両課題から得られる  $\Delta PC_{\text{test,base}}$  の標準偏差が僅かに異なると報告した。

二つ目に、立ち上がり時間が長い音の PAT が、課題間で異なる傾向があると報告された [9]。Gordon [9] は、同時調整課題と等時調整課題を行った。その結果、立ち上がり時間が長い音 (約 100ms 以上) とベース音との  $\Delta PC_{\text{test,base}}$  が、等時調整課題よりも同時調整課題の方が大きい傾向があったと報告した。

三つ目に、持続時間が増加した時の PAT の変化量と、立ち上がり時間が増加した時の PAT の変化量との、どちらがより大きいか課題間で異なる傾向があると報告された [5]。Danielsen ら [5] は、London ら [15] の実験 1 におけるタッピング課題と同時調整課題をより詳細に分析した。その結果、同時調整課題では、持続時間が増加した時の  $\Delta PC_{\text{test,base}}$  の変化量よりも立ち上

がり時間が増加した時の  $\Delta PC_{\text{test,base}}$  の変化量の方が大きい傾向があるが、タッピング課題では、立ち上がり時間が増加した時の同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  の変化量よりも持続時間の増加が増加した時の同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  の変化量の方が大きい傾向があると報告した。

四つ目に、持続時間が長い音と立ち上がり時間が長い音との、どちらが標準偏差が大きいかが課題間で異なる傾向があると報告された [5]。Danielsen ら [5] は、London らの実験 1 におけるタッピング課題と同時調整課題との詳細な分析を行った。その結果、同時調整課題では、持続時間が長い音よりも立ち上がり時間が長い音の方が  $\Delta PC_{\text{test,base}}$  の標準偏差が大きい傾向があるのに対し、タッピング課題では、持続時間が長い音の方が同調のズレ量の平均値  $\overline{\text{Asy}}(\text{base})$  の標準偏差が大きい傾向があると報告した。

五つ目に、得られた値の確率密度分布が課題間で異なる傾向が報告された。今まで述べてきた PAT はすべて「PAT はある一瞬の時点にあるものである」という考え方に従っている。だが、PAT はある一時点のものではなく許容幅をもつ、すなわち、「PAT が何らかの確率密度分布に従う」という考え方もある。その考えに基づいた研究では、実験結果より PAT の分布をカーネル密度推定で算出した [9, 8, 37, 2, 5]。算出された PAT の分布は、多くの場合は単峰分布となると報告された (正規分布に従うものから、尖度や歪度があるものも含む)。だが、同時調整課題と等時調整課題を用いた先行研究では、音によって双峰分布になるものもあったと報告する研究と [9, 8, 37, 5]、全て単峰分布であったと報告する研究 [2] がある。一方、タッピング課題を用いた先行研究では、全ての刺激音が単峰分布であったと報告された [5]。

以上のように、先行研究より、課題間に統計的有意差はないものの、課題間に僅かな差異がある傾向が指摘されている。

## 2.4 PAT を変化させる音の要素

PAT の先行研究において主な問いは、音の何の要素を変化させると PAT が変化するかである。それらの検討の結果、一般音 (言語音ではなく、楽器音や合成音のこと) においては、報告数の多い順に、立ち上がり時間、持続時間、周波数、提示音圧レベルを変化させた時に、PAT が異なると報告される。また、合成音と楽器音とで、PAT を変化させる要因が異なる可能性も報告される。以下では、それらについて詳述する。

### 2.4.1 立ち上がり時間

立ち上がり時間が変化すると PAT も変化するという報告について述べる。それらの報告より、立ち上がり時間が長い方が PAT が時間的に後方に位置することが一貫して報告される。

Gordon[8, 9] は、16 種類の楽器音を用いて、同時調整課題と等時調整課題を行った。参加者は、コンピューター音楽に精通していた。刺激の楽器音<sup>5</sup> は、録音された後、持続時間が 320 ms に、周波数が中央の Es(約 311 Hz) に、音圧レベルが最大で 90 dB になるよう調整された。立ち上がり時間は調整されなかった。楽器音ごとの立ち上がり時間は 45–122 ms であった<sup>6</sup>。分析にて、 $\Delta PC_{\text{test,base}}$  を従属変数とし、刺激音(楽器音)の種類を独立変数として、一要因参加者内分散分析を行った結果、刺激音の種類の主効果が有意であったと報告した。Gordon はこの結果から、実験前に各刺激音の持続時間、周波数、音圧レベルは統一されたため、それぞれの楽器音の立ち上がり時間や周波数スペクトラムの時間変化の違いが PAT を変化させるが、より重要なのは立ち上がり時間(及び立ち上がり部分の傾き)であると考察した。

London ら [15] は、実験 1 にて、8 種の楽器音とクリック音を用いて、タッピング課題、同時調整課題、等時調整課題を行った。参加者は音楽訓練を受けたことがあった。刺激音である楽器音<sup>7</sup> は、持続時間、立ち上がり時間、周波数がそれぞれ異なった。分析にて、同時調整課題と等時調整課題では  $\Delta PC_{\text{test,base}}$  を、タッピング課題では同調のズレ量  $\overline{\text{Asy}}(A)$  を従属変数とし、課題の種類、刺激音の種類を独立変数として、二要因の分散分析を行った。その結果、刺激音の種類的主効果が有意であったと報告した。

そして、Danielsen ら [5] は実験 1 として、London らの実験 1 におけるタッピング課題と同時調整課題を、さらに詳細に分析した。London らの用いた 8 種の刺激音は、持続時間 2 水準(長・短)、立ち上がり時間 2 水準(長・短)、周波数 2 水準(高・低)に分類されていた。Danielsen らは、同時調整課題では  $\Delta PC_{\text{test,base}}$  を、タッピング課題では同調のズレ量  $\overline{\text{Asy}}(A)$  を算出した。それぞれの課題において、これらに従属変数として、持続時間、立ち上がり時間、周波数を独立変数として、三要因分散分析を行った。その結果、タッピング課題でも同時調整課題でも、持続時間の主効果、立ち上がり時間の主効果、周波数の主効果、そして持続時間と周波数との交互作用が有意であったと報告した。また、持続時間と立ち上がり時間との交互作用はタッピング課題でのみ有意であったと報告した。

---

<sup>5</sup>Gordon[8, 9] が用いた 16 種類の楽器音は、以下である。コール・アングレ (English horn)、ファゴット、トランペット、フレンチホルン、3つのボーイングによるチェロ(同じ奏者が異なるボーイングを行った; sul ponticello, muted sul tasto, normal bowing)、トロンボーン、2人の異なる奏者によるオーボエ、フルート、Eb 管クラリネット、Bb 管クラリネット、ソプラノサクソ、音量が 2 水準 (mf, p) で異なるアルトサクソであった。

<sup>6</sup>Gordon[9] は、2 種類の立ち上がり時間を報告した。それは、物理的オンセットから最大エネルギーまでの範囲において振幅が “0–100%” まで変化した時の時間を立ち上がり時間としたものと、同範囲において振幅が “10–90%” まで変化した時の時間を立ち上がり時間としたものである。本論では、前者の “0–100%” を立ち上がり時間としたものを引用した。

<sup>7</sup>London ら [15] が用いた 8 種の楽器音とクリック音は以下である。なお、以下の持続時間は London らが報告する MIR toolbox より算出された値を引用し、周波数は MIR toolbox より算出された spectral centroid の値を引用した。

クリック音(持続時間 1 ms、周波数 3 kHz)、キックドラム(持続時間 123 ms、立ち上がり時間 5 ms、周波数 780 Hz)、スネアドラム(持続時間 38 ms、立ち上がり時間 2 ms、周波数 2831 Hz)、ダークピアノ(持続時間 371 ms、立ち上がり時間 18 ms、周波数 623 Hz)、ライトピアノ(持続時間 211 ms、立ち上がり時間 19 ms、周波数 893 Hz)、コントラバス(持続時間 70 ms、立ち上がり時間 31 ms、周波数 538 Hz)、カバサ(持続時間 61 ms、立ち上がり時間 22 ms、周波数 8199 Hz)、シンセベース(持続時間 220 ms、立ち上がり時間 33 ms、周波数 781 Hz)、フィドル(持続時間 89 ms、立ち上がり時間 43 ms、周波数 1581 Hz)であった。

また、Danielsen らは、実験2にて、ノイズ音を用いて、タッピング課題と同時調整課題を行った。参加者の殆どは音楽訓練を受けたことがあった。刺激音であるノイズ音は、持続時間を2水準(100, 400 ms)、立ち上がり時間を2水準(3, 50 ms)、中心周波数を2水準(100, 700 Hz)で変化させた。この6個の刺激音に、クリック音(持続時間1 ms、立ち上がり時間0 ms、中心周波数3000 Hz)を加えた、計7個の刺激音が用いられた。分析にて、同時調整課題より  $\Delta PC_{\text{test,base}}$  を算出し、タッピング課題より同調のズレ量の平均値  $\overline{A_{\text{sy}}}(A)$  を算出した。それぞれの課題において、これらを従属変数とし、持続時間、立ち上がり時間、周波数を独立変数として、三要因分散分析を行った。その結果、主効果も交互作用も全て有意ではなく、多重比較を行った結果、タッピング課題でも同時調整課題でも、クリック音を除いて、有意となった組み合わせはなかったと報告した。だが、考察において、傾向として、立ち上がり時間がより長い音の方がPATが時間的後方へ位置したこと、すなわち、有意にはならなかったが、合成音を用いた実験2でも楽器音を用いた実験1と同様の傾向が見られたと報告した。

Camp, Kilchenmann, Volken, & Senn [2] は、8種の楽器音を用いて、同時調整課題を行った。参加者は音楽訓練を受けていた。刺激音である楽器音<sup>8</sup> は、ラウドネスのみ統制された。持続時間(2–4700 ms)、立ち上がり時間(0.07–1666 ms)、周波数(77–465 Hz など) はそれぞれ異なった。様々な分析を行い、最終的に、PATの変化により重要なのは、立ち上がり時間であると結論付けた。

Collins[3] は、サイン波を用いて、等時調整課題を行った。参加者は音楽訓練を受けたことがあった。刺激音であるサイン波は、持続時間を200 msで固定し、立ち上がり時間を5水準(0, 10, 20, 50, 100 ms)で変化させ、周波数を5水準(約130, 261, 523, 1046, 2093 Hz)で変化させた。提示音圧レベルは参加者間で固定であった。分析にて、 $\Delta PC_{\text{test,base}}$  を従属変数とし、立ち上がり時間、周波数、繰り返しを独立変数として、三要因分散分析を行った結果、立ち上がり時間の主効果のみ有意であったと報告した。

J.Vos & Rasch[34](P.Vosら [35]ではない) は、実験1にて、サイン波を用いて、等時調整課題を行った。参加者は音楽訓練を受けたことがあった。刺激音であるサイン波は、立ち上がり時間を5水準(5, 20, 40, 60, 80 ms)で変化させ、周波数を400 Hzで固定した。この5種類のサイン波は、ベース音とテスト音に用い、ベース音の立ち上がり時間がテスト音の立ち上がり時間より短い組み合わせのみに調整させた。刺激音の提示音圧レベルは82 dBで参加者間で固定であった。通常の等時調整課題ではテスト音の位置を調整させるが、J.Vosらは、テスト音の持

<sup>8</sup>Campら [2] が用いた8種の楽器音は以下である。なお、立ち上がり時間は、暗騒音レベルを超えた時刻から最大エネルギーの時刻までとされた。

クリック音(持続時間2 ms、立ち上がり時間0.07 ms)、ハイハットシンバル(持続時間260 ms、立ち上がり時間5.76 ms、周波数0–22000 Hz)、ライドシンバル(持続時間4700 ms、立ち上がり時間1.77 ms、周波数260–22000 Hz)、ピアノ(持続時間2850 ms、立ち上がり時間27.71 ms、周波数234 Hz)、ピッチカートで奏されたコントラバス(持続時間3950 ms、立ち上がり時間37.69 ms、周波数77 Hz)、アルコで奏されたヴィオラ(持続時間2700 ms、立ち上がり時間534.56 ms、周波数441 Hz)、アルコで奏されたチェロ(持続時間3170 ms、立ち上がり時間810.93 ms、周波数330 Hz)、スル・タストで奏されたヴァイオリン(持続時間2630 ms、立ち上がり時間1666.21 ms、周波数465 Hz)

続時間の長さを調整させた。分析にて、 $\Delta t$ を算出した。この値は、算出方法としては、本論の  $\Delta PC_{\text{base, test}}$  ( $\Delta PC_{\text{test, base}}$  ではない) と同様である<sup>9</sup>。だが、J.Vos らは持続時間を調整させる方法を用いたため、この  $\Delta t$  には、テスト音とベース音の立ち上がり時間だけでなく、調整後のテスト音の持続時間も影響していると考えられる。このことに注意するべきであるため、 $\Delta PC_{\text{base, test}}$  と言い換えずに、 $\Delta t$  と称した。J.Vos らは、分析にて、ベース音の立ち上がり時間のそれぞれの水準において、この  $\Delta t$  を従属変数として、テスト音の立ち上がり時間を独立変数とした一要因分散分析を行った。また、テスト音の立ち上がり時間のそれぞれの水準において、ベース音の立ち上がり時間を独立変数とした一要因分散分析を行った。すなわち、ベース音が 5 ms の時の、テスト音 5 水準 (5, 20, 40, 60, 80 ms) を独立変数とした一要因分散分析を行い、次にベース音が 20 ms の時の、テスト音 4 水準 (20, 40, 60, 80 ms) での分散分析を行う、ということを繰り返した。また、同様に、テスト音が 20 ms の時の、ベース音 2 水準 (5, 20 ms) を独立変数とした一要因分散分析を行い、次にテスト音が 40 ms の時の、ベース音 3 水準 (5, 20, 40 ms) を独立変数とした一要因分散分析を行う、ということを繰り返した。このように行ったのは、先述した通り、実験にて、ベース音の立ち上がり時間がテスト音の立ち上がり時間より短い組み合わせのみに調整させたからであった。以上の一要因分散分析の結果、殆どの分散分析にて立ち上がり時間の主効果が有意であったと報告した。

最後に、Wright[37] と、Polfreman[22] は、実験結果を考察し、立ち上がり時間が長い音 (“non-percussive sounds” / “slow attacks”) は、立ち上がり時間が短い音 (“percussive sounds”) よりも、PAT が時間的後方に位置することを推察した。だが、これらの音の定義は述べられておらず、具体的に持続時間や立ち上がり時間がどのくらいの長さの音を指すのかも報告されていない。

一方、以上の報告とは異なり、Bechtold & Seen[1] は、立ち上がり時間と PAT との相関が低いことから、立ち上がり時間は PAT を変化させないと結論づけた。Bechtold らは、楽器音を用いて、同時調整課題を行った。参加者は音楽訓練を受けていた。刺激音である楽器音は、アーティキュレーション (“articulation” / “strength of the tongue”) とダイナミクスを変化させて演奏させたものであった。刺激音ごとに、持続時間 (平均 564 ms、標準偏差 77 ms)、立ち上が

<sup>9</sup> J.Vos らの報告によると、“ $\Delta t$ ”は、テスト音とベース音の理論的時間間隔を “ $T$ ” (本論の  $\text{idealIOI}$ ) とし、ベース音の物理的オンセットを “ $t_a$ ” (本論の  $t_{\text{onset}}(\text{base})$ ) とし、テスト音の物理的オンセットを “ $t_b$ ” (本論の  $t_{\text{onset}}(\text{test})$ ) とした場合の、“ $\Delta t = T - (t_b - t_a)$ ” である。これを、本論の表記に合わせると  $\Delta t = T - (t_b - t_a) = \text{idealIOI} - (t_{\text{onset}}(\text{test}) - t_{\text{onset}}(\text{base}))$  となる。これを式変形すれば、以下のように、 $\Delta t = \Delta PC_{\text{base, test}}$  であることが導ける。

$$\begin{aligned}
 \Delta t &= \text{idealIOI} - (t_{\text{onset}}(\text{test}) - t_{\text{onset}}(\text{base})) \\
 &= -\text{idealIOI} + (t_{\text{onset}}(\text{test}) - t_{\text{onset}}(\text{base})) \\
 &= t_{\text{onset}}(\text{test}) - t_{\text{onset}}(\text{base}) - \text{idealIOI} \\
 &= t_{\text{onset}}(\text{test}) - t_{\text{onset}}(\text{base}) - \text{idealIOI} \\
 &= -\Delta PC_{\text{test, base}} \\
 &= \Delta PC_{\text{base, test}}
 \end{aligned}$$

り時間 (60–173 ms)、音圧レベル (RMS loudness –17.1––36.9 dB) が異なった。刺激音全体の提示音圧レベルも参加者それぞれが好むレベルに調整させた。分析として、 $\Delta PC_{\text{test,base}}$  と立ち上がり時間との相関係数を算出したところ、相関が低かったと報告した。

以上のように、立ち上がり時間が異なると PAT も異なることが多くの先行研究より報告された。それらの研究では、一貫して、立ち上がり時間が長い方が PAT が時間的に後方に位置することが報告された。

だが、これらの研究では、立ち上がり時間をいくつかの水準で変化させたときの PAT の変化が統計的に有意かどうかのみが議論されており、立ち上がり時間がどのくらい長くなると PAT がどのくらい移動するのか、という両者の関係は明らかになっていない。

## 2.4.2 持続時間

次に、持続時間が変化すると PAT も変化するという報告について述べる。それらより、一貫して、持続時間が長い方が、PAT が時間的に後方に位置することが報告された。

P.Vos ら [35](J.Vos ら [34] ではない) は、合成音を用いて、タッピング課題を行った。刺激音である合成音は、持続時間を 5 水準 (1, 2, 50, 100, 300 ms) で変化させ、立ち上がり時間ではなく持続時間に占める立ち上がり時間の割合を 3 水準 (0, 40, 80 %) で変化させた。また、音を繰り返す時間間隔 (テンポ) も 3 水準 (500, 700, 900 ms) で変化させた。周波数は 500 Hz で固定であった。分析は、以下の二つに分けて行われた。まず、音を繰り返す時間間隔 (テンポ) が 700 ms のデータのみを絞り、分析を行った。同調のズレ量の平均値を従属変数とし、持続時間、立ち上がり時間の割合を独立変数として、二要因参加者内分散分析を行った。その結果、持続時間の主効果、立ち上がり時間の主効果、そして持続時間と立ち上がり時間との交互作用が有意であったと報告した。次に、全てのデータに対して分析を行った。同調のズレ量の平均値を従属変数とし、テンポ、持続時間を独立変数として、二要因参加者内分散分析を行った。その結果、テンポの主効果は有意ではなく、持続時間の主効果のみ有意であったと報告した。

Danielsen ら [5] は、実験 1 として、2.4.1 項で述べた通り、London ら [15] のタッピング課題と同時調整課題に絞り、さらに詳細に分析を行った。分散分析の結果、どちらの課題でも、持続時間の主効果が有意であったと報告した。また、タッピング課題でのみ、持続時間と立ち上がり時間との交互作用が有意であったと報告した。

また、Danielsen らは、実験 2 にて、2.4.1 項で述べた通り、ノイズ音を用いてタッピング課題と同時調整課題を行った。分散分析の結果、有意となったものはなかったと報告した。だが、考察において、傾向として、立ち上がり時間がより長い音の方が PAT が時間的に後方へ位置したこと、すなわち、合成音を用いた実験 2 でも楽器音を用いた実験 1 と同様の傾向が見られたと報告した。

一方、以上の報告とは異なり、Bechtoldら [1] は、持続時間と PAT との相関が低いことから、持続時間は PAT を変化させないと結論づけた。Bechtold らは、2.4.1 項で述べた通り、楽器音を用いて、同時調整課題を行った。分析として、 $\Delta PC_{\text{test,base}}$  と持続時間との相関係数を算出したところ、相関が低かったと報告した。

以上のように、持続時間が異なると PAT も異なることが報告された。それらの研究では、一貫して、持続時間が長い方が PAT が時間的に後方に位置することが報告された。加えて、これらのことを報告した先行研究では同時に、持続時間と立ち上がり時間との交互作用も報告された。持続時間が長い音の方が、立ち上がり時間が異なる時の PAT の変化が大きいことが示された。

だが、これらの研究では、持続時間をいくつかの水準で変化させたときの PAT の変化が統計的に有意かどうかのみが議論されている。すなわち、持続時間がどのくらい長くなると PAT がどのくらい移動するのかという関係は明らかになっていない。

### 2.4.3 周波数

続いて、周波数が変化すると PAT が変化するという報告について述べる。そして、周波数が異なっても PAT は異ならないと報告する研究について言及する。

Danielsen ら [5] は、2.4.1 項で述べた通り、London ら [15] のタッピング課題と同時調整課題を詳細に分析した。分散分析の結果、周波数の主効果が有意であったと報告した。周波数が低い音の方が、PAT が時間的に後方に位置すると報告した。

一方、Danielsen ら [5] は、2.4.1 項で述べた通り、ノイズ音を用いて実験 2 を行った。分散分析の結果、有意となったものはなく、周波数が異なっても PAT が殆ど変化しなかったと報告した。

Collins [3] は、2.4.1 項で述べた通り、サイン波を用いて、等時調整課題を行った。立ち上がり時間、周波数、繰り返しの三要因分散分析の結果、立ち上がり時間の主効果のみ有意であり、周波数は有意ではなかったと報告した。

以上のように、周波数が低い音の方が PAT が時間的に後方に位置すると報告するものと、周波数が異なっても PAT は変化しないと報告するものがある。従って、周波数が PAT を変化させるかどうかは未だ明らかではない。

### 2.4.4 提示音圧レベル

そして、提示音圧レベルが変化すると PAT も変化するという報告について述べる。

J.Vosら [34] は、実験2にて、サイン波を用いて、等時調整課題を行った。参加者は音楽訓練を受けたことがあった。刺激音であるサイン波は、立ち上がり時間を3水準 (5, 40, 80 ms) で変化させ、周波数を 400 Hz で固定した。また、刺激音の提示音圧レベルを3水準 (37, 57, 77 dB) で変化させた。これらのサイン波をベース音とテスト音に用い、全ての組み合わせに調整させた。通常の等時調整課題ではテスト音の位置を調整させるが、J.Vosらは、2.4.1項で述べた実験1と同様に、実験2でも、テスト音の持続時間の長さを調整させた。分析にて、 $\Delta t$  を算出した ( $\Delta t$  については 26 ページの J.Vos の実験1の詳述を参照してほしい)。 $\Delta t$  を従属変数とし、テスト音の立ち上がり時間、ベース音の立ち上がり時間、提示音圧レベル、繰り返しを独立変数として、四要因参加者内分散分析を行った。その結果、ベース音の立ち上がり時間の主効果、テスト音の立ち上がり時間の主効果、ベース音の立ち上がり時間と提示音圧レベルとの交互作用、テスト音の立ち上がり時間と提示音圧レベルとの交互作用が有意であったと報告した。再度注記しておくが、J.Vosらは持続時間を調整させる方法を用いたため、この結果には、テスト音とベース音の立ち上がり時間だけでなく、調整後のテスト音の持続時間も影響していると考えられる。

Bechtoldら [1] は、2.4.1項で述べた通り、楽器音を用いて同時調整課題を行った。分析として、 $\Delta PC_{\text{test,base}}$  を従属変数とし、アーティキュレーション、ダイナミクスを独立変数として、二要因の分散分析を行った。その結果、アーティキュレーションの主効果、ダイナミクスの主効果、アーティキュレーションとダイナミクスとの交互作用が有意であったと報告した。下位検定の結果より、ダイナミクスが小さい音では、アーティキュレーションが異なると PAT が大きく変化する可能性と、ダイナミクスが中庸な音よりもダイナミクスが大きい音の方が PAT が時間的前方に位置する可能性があることを報告した。

以上のように、提示音圧レベルが PAT を変化させることは、主に、提示音圧レベルとその他の音の要素との交互作用で報告された。具体的に提示音圧レベルとその他の音の要素との交互作用がどのようになっているのか、提示音圧レベルが PAT をどのように変化させるかについては更なる研究が必要である。

## 2.4.5 合成音と楽器音

最後に、今まで述べてきた音の何の要素が PAT を変化させるのかが、合成音と楽器音とで異なる可能性が報告された。すなわち、合成音の知見を楽器音にそのまま転用できない可能性が報告された。

Collins[3] は、2.4.1項と 2.4.3項で述べた実験以外に、多種の楽器音を用いても等時調整課題を行い、実験から得られた PAT と既存の PAT モデル (J.Vosら [34] や Gordon[8] が提唱したモデルなど) より予測される PAT を比較した。モデルの各パラメータは手動で探索した。その比

較の結果、合成音では、モデルの予測値と実験で得た実測値とが近かったが、楽器音では大きく異なったことを報告した。

Danielsen ら [5] は、2.4.1 項と 2.4.2 項、2.4.3 項で述べた通り、実験 1 として、楽器音を用いた実験への詳細な分析を行い、実験 2 として、合成音を用いた実験と分析を行った。分散分析の結果、実験 1 の楽器音においては、持続時間の主効果、立ち上がり時間の主効果、周波数の主効果が有意であったと報告した。一方、実験 2 の合成音においては、全て有意ではなく、傾向としては持続時間がより長い音の方が PAT が時間的後方へ位置し、立ち上がり時間がより長い音の方が PAT が時間的後方へ位置したことを報告した。

だが、Danielsen らの楽器音を用いた実験 1 では、分散分析における独立変数の各水準が綺麗に分割されていないという問題がある。Danielsen らは 8 種の楽器音を、各要因ごとに 2 つのカテゴリ (立ち上がり時間を例にすると、早い音と遅い音) に分け、分散分析を行った。報告された各水準の値の範囲を見ると、立ち上がり時間において「早い」水準は 2–19 ms であり、「遅い」水準は 22–43 ms であった。持続時間では、「短い」水準は 61–123 ms であり、「長い」水準は 89–371 ms であった。周波数においては、「低い」水準は 538–781 Hz であり、「高い」水準は 893–2831 Hz であった。特に、持続時間では 2 つの水準間で持続時間の範囲が重なっている部分があり、立ち上がり時間と周波数も連続値を 2 つに分割した形に近く、各水準があまり綺麗に分割されていないことが明らかである。

以上のように、合成音と楽器音に差異がある可能性が挙げられる。だが、Collins の報告はモデルへの当てはめからの考察であり、実際の精密な実験の結果ではない。また、Danielsen らも、先述した通り、精密に調査されたとは言い難い。そのため、合成音と楽器音に差異がある可能性はあるが、結論づけるためには、より詳細な研究が必要であると考えられる。

## まとめ

先行研究より、PAT を変化させるものとして、持続時間、立ち上がり時間、周波数、提示音圧レベルが報告された。また、合成音の知見を楽器音にそのまま転用できない可能性も報告された。

明らかとなったことを以下に箇条書きで示す。今まで挙げた先行研究で用いられた刺激音と分散分析の結果を表 2.4.1 にまとめた (なお、時間範囲の列における「統合の時間窓」については次節で述べる)。また、この表における分散分析の結果の列を、行列で表したものが表 2.4.2 である。先行研究において何が調査され、その分散分析の結果がどうだったのかを把握するために必要に応じて参照してほしい。

- 立ち上がり時間が異なると PAT も変化することが最も多くの先行研究より報告された。これらの研究では、一貫して、立ち上がり時間が長い音の方が PAT が時間的に後方に位置

することが報告された。だが、立ち上がり時間と PAT との相関係数は低い可能性も示唆された。

だが、立ち上がり時間がどのくらい長くなると PAT がどのくらい移動するのか、という両者の関係は明らかになっていない。

- 持続時間が異なると PAT も変化することが先行研究より報告された。それらの研究では、一貫して、持続時間が長い方が PAT が時間的に後方に位置することが報告された。一方、持続時間と PAT との相関係数は低い可能性も示唆された。

だが、持続時間がどのくらい長くなると PAT がどのくらい移動するのか、という両者の関係は明らかになっていない。

- 持続時間と立ち上がり時間との交互作用が先行研究より報告された。それらの研究では、持続時間が短い音よりも長い音の方が、立ち上がり時間が異なる時の PAT の変化が大きいことが報告された。
- 周波数が低い音の方が PAT が時間的に後方に位置するという報告と、周波数が異なっても PAT は変化しないという報告があった。従って、周波数が低い音の方が PAT が時間的に後方に位置する可能性があるが、結論づけるためには更なる研究が必要である。
- 提示音圧レベルとその他の音の要素との交互作用がある可能性が報告された。だが、結論づけるためには更なる研究が必要である。また、具体的に提示音圧レベルが PAT をどのように変化させるかについても更なる研究が必要である。
- 音の何の要素が PAT を変化させるのかが、合成音と楽器音とで異なる可能性が報告された。だが、結論づけるためには、より詳細な研究が必要である。

表 2.4.1: PAT の先行研究で用いられた刺激音と分散分析の結果

\*1 ... 「論文」では、各論文の名称として、第一著者と出版年を記載した。その論文中で2つ以上の実験を行っているものは、どの実験かも記載した。

\*2 ... 「課題」では、PAT 測定課題を以下の方法で略記した。

タッピング課題は「タッピング」、同時調整課題は「同時調整」、等時調整課題は「等時調整」と記された。

\*3 ... 「刺激音」では、タッピング課題においてはすべての刺激音を、等時調整課題と同時調整課題では、テスト音を記載した。

\*4 ... 「時間範囲」では、統合の時間窓 (150–300ms) を 200 ms とした場合に、用いられた持続時間が統合の時間窓の範囲内か範囲外か、それとも範囲内から範囲外までかを示した。

「内」の表記は用いられた持続時間が 200 ms 以内であることを、「外」の表記は持続時間が 200 ms 以上であることを、「内外」の表記は持続時間が 200 ms 以内から 200 ms 以上までであることを表す。

\*5 ... 「分散分析の結果」では、分散分析の各要因を略記した。略記の仕方は計画の列に記載したものをを用いた。また、各要因の主効果については要因の略記のみとし、交互作用は2つの要因を“×”でつないで略記した。そして、有意は「SIG.」有意でないは「N.S.」と略記した。論文中に何も報告されなかった要因は有意ではなかったと推測し、「(N.S.)」と記載した。

例) 立ち上がり時間 “R” の主効果が有意であった場合、「“R” SIG.」と記載した。

立ち上がり時間 “R” と持続時間 “D” の交互作用が有意であった場合、「“R”×“D” SIG.」と記載した。

論文*1	課題*2	刺激音*3	時間範囲*4	立ち上がり時間	持続時間	周波数	提示音圧レベル	値	分散分析の計画	分散分析の結果*5
P.Vos1995[35] 実験2	タッピング	合成音	内外	0–80% (0–240 ms)	1–300 ms	[500 Hz]	参加者が調整	同調のズレ量 目視: –55–100 ms	二要因 (立ち上がり時間 “R”、 持続時間 “D”)	“R” SIG. “D” SIG. “R×D” SIG.
Danielsen2019[5] 実験2	タッピング	合成音	内外	0–50 ms	1–400 ms	100–3000 Hz	参加者が調整	同調のズレ量 –26.71–26.48 ms	三要因 (立ち上がり時間 “R”、 持続時間 “D”、 周波数 “F”)	“R” N.S. “D” N.S. “F” N.S. (全て N.S. で傾向止まり)
	同時調整	合成音	内外	0–50 ms	1–400 ms	100–3000 Hz	参加者が調整	ΔPC 0.42–34.46	同上	同上
Collins2006[3]	等時調整	合成音	内	0–100 ms	[200 ms]	130–2093 Hz	[記載なし]	ΔPC 目視: –10–80 ms	三要因 (立ち上がり時間 “R”、 周波数 “F”、 繰り返し “Rep”)	“R” SIG. “F” N.S. “Rep” N.S. “R”×“F” N.S. 全ての交互作用 N.S.
J.Vos1981[34] 実験1	等時調整 (持続時間を調整)	合成音	—	5–80 ms	調整値	[400 Hz]	[82 dB]	Δt (≈ΔPC) –1.6–34.9 ms	一要因 (立ち上がり時間の ペアの種類 “Rpair”)	“Rpair” SIG.
J.Vos1981[34] 実験2	等時調整 (持続時間を調整)	合成音	—	5–80 ms	調整値	[400 Hz]	37, 57, 77 dB	Δt (≈ΔPC) –53.3–58.1 ms	四要因 (ベース音の立ち上がり時間 “Rbase”、 テスト音の立ち上がり時間 “Rtest”、 音量 “L”、 繰り返し “Rep”)	“Rbase” SIG. “Rtest” SIG. “Rbase”×“L” SIG. “Rtest”×“L” SIG. その他は記載なし (“L” N.S.)
Danielsen2019[5] 実験1 (London2019[15] 実験1)	タッピング	楽器音	内外	0–43 ms	1–487 ms	538–8199 Hz	参加者が調整	同調のズレ量 –21.57–16.89 ms	三要因 (立ち上がり時間 “R”、 持続時間 “D”、 周波数 “F”)	“R” SIG. “D” SIG. “F” SIG. “D”×“F” SIG. “D”×“R” SIG. その他は記載なし
	同時調整	楽器音	内外	0–43 ms	1–487 ms	538–8199 Hz	参加者が調整	ΔPC –0.50–23.44 ms	同上	同上
Gordon1987[8] (Gordon1984[9])	同時調整、 等時調整	楽器音	外	45–122 ms	[320 ms]	[311 Hz]	[–90 dB]	両課題平均 ΔPC –21–28 ms	一要因 (楽器の種類 ≈ 立ち上がり時間 “R”)	“R” SIG.
Bechtold2018[1]	同時調整	楽器音 (奏法変化)	外	60–173 ms	平均 564ms ±SD 77ms	[311 Hz; Eb4]	参加者が調整 (刺激音の RMS: –17.1––36.9 dB)	ΔPC –23––83 ms	— 相関係数 (立ち上がり時間 “R” と ΔPC、 持続時間 “D” と ΔPC)	— “R” と ΔPC 弱い “D” と ΔPC 弱い
Nymoen2017[19]	同時調整	楽器音	—	記載なし	記載なし	記載なし	記載なし	言及ないが多分 ΔPC 目視: 0–20 ms	—	—
Polfreman2013[22]	同時調整 (ベース音も 調整可能)	楽器音、 正弦波	—	記載なし	記載なし	[130 Hz; C4]	参加者が調整 (実験中も 調整可能)	ΔPC –0.46–51.21 ms	—	—
Camp2011[2]	同時調整	楽器音	内外	0.07–1666 ms	2–4700 ms	77–465 Hz, 22000 Hz	記載なし	ΔPC –6.2–46.5 ms	様々な分析	—
Wright2008[37]	同時調整	楽器音	外	ばらばら (記載なし)	[300 ms]	[311 Hz; Eb4]	参加者ごとの 再生環境	ΔPC の中央値 –14.0–40.2 ms	—	—

表 2.4.2: PAT の先行研究の分散分析の結果

表 2.4.1 の分散分析の列にてまとめたものを行列に変換したものである。

分散分析が行われ、有意であったと報告された要因には“SIG.”と、有意ではないと報告された要因には“N.S.”と表記した。分散分析が行われたが論文中に何も報告されなかった要因は、その要因は有意ではなかったと推測し、“(N.S.)”と記載した。

統制された要因は“[-]”と表記した。また、提示音圧レベルを参加者が調整した場合には“好み”と表記した。J.Vos らが用いた課題では持続時間の長さを調整したため、彼らの持続時間のセルに“調整”と記載した。

検討されなかった交互作用は“(空白)”とした。

\*1 ... 「論文」では、各論文の名称として、第一著者と出版年を記載した。その論文中で2つ以上の実験を行っているものは、どの実験かも記載した。

論文*1	刺激音	立ち上がり時間 (R)	持続時間 (D)	周波数 (F)	提示音圧レベル (L)	D×R	D×F	F×L	R×F	R×L	F×L
P.Vos ら [35]	合成音	SIG.	SIG.	[-]	好み	SIG.					
Danielsen 実験 2[5]	合成音	N.S.	N.S.	N.S.	好み	N.S.	N.S.		N.S.		
Collins[3]	合成音	SIG.	[-]	N.S.	[-]					N.S.	
J.Vos ら 実験 1[34]	合成音	SIG.	調整	[-]	[-]						
J.Vos ら 実験 2[34]	合成音	SIG.	調整	[-]	(N.S.)						SIG.
London ら [15] / Danielsen 実験 1[5]	楽器音	SIG.	SIG.	SIG.	好み	SIG.	SIG.			(N.S.)	
Gordon[9, 8]	楽器音	SIG.	[-]	[-]	[-]						

## 2.5 聴覚情報処理における統合の時間窓

音の知覚は聴覚情報処理を経て形成されるため、PAT の知覚も聴覚情報処理の何らかの影響を受けるはずである。

今まで述べてきた PAT の先行研究では実験で刺激音の持続時間を変化させており、その多くは持続時間を 1 ms から 400 ms あるいは 500 ms 程度まで変化させている。これは、32 ページにある表 2.4.1 の PAT の先行研究の刺激音からも明らかである。この時間範囲は、「統合の時間窓」の時間範囲と重なる。詳細は後述するが、「統合の時間窓」とは、聴覚情報が保持される時間範囲の一つである。この時間範囲で保持された情報はまとめて統合処理される。例えば、この処理により生じる現象としては、統合の時間窓の範囲内であれば、持続時間が増加するにつれ最小可聴閾が低下し、ラウドネスが上昇して知覚される現象が生じることだ。この現象は統合の時間窓の範囲外では生じず、持続時間が増加してもラウドネスも最小可聴閾も変化しないことも知られている。そうであれば、持続時間を変化させた時の PAT の知覚にも、統合の時間窓が影響する可能性が考えられる。どのように PAT の知覚に統合の時間窓が影響するかについては次節で検討する。

本節では、統合の時間窓について詳説する。以下では、まず、そもそもなぜ、ある時間範囲分の聴覚情報を保持し、その時間分の情報をまとめて処理する必要があるかについて簡略的に述べる。次に、PAT 知覚に関与するであろう、統合の時間窓の範囲内と範囲外の境界、すなわち統合の時間窓の長さについて詳述する。

## 2.5.1 聴覚情報が保持される「時間範囲」とその範囲分の情報処理による知覚の変化

そもそもなぜ、ある「時間範囲」分の聴覚情報を保持し、その時間分の情報をまとめて処理する必要があるかについて簡略的に述べる。

私たち人間が日常生活で耳にする音の大半は、時間が経過するにつれ、周波数や音圧レベルが変化する。そのような音から得た情報を処理する時に一瞬一瞬の情報を逐次的に処理してしまふと時間変化に対応しにくい。そのため、まず、ある「時間範囲」分の音の情報を保持し、その後、その時間分の情報をまとめて処理する必要がある。

確かにこのようにすることで時間変化する音を効率的に知覚できる。だが、その半面、もともと感覚器官に入ってきた音の物理的な情報と、それらの情報が処理され知覚された結果が異なることも知られている。例えば、一つの「時間範囲」に二つの音が含まれた場合、その二つの音が一つの音として知覚されたり、その二つの音の全体的なラウドネスがそれぞれの音のラウドネスよりも大きく知覚されたりする。また、一つの「時間範囲」に一つの音が含まれた場合には、その音の持続時間分のエネルギーが統合(積分)され、この処理により、先述した統合の時間窓の範囲内であれば持続時間が増加するにつれてラウドネスと最小可聴閾が変化する現象も生じる。

このような現象を引き起こす「時間範囲」は、引き起こされる現象とその「時間範囲」の長さによって、三つに大別される [6, 4, 20, 40, 31, 39]<sup>10</sup> これら三つの「時間範囲」のうち、PAT 知覚に関与すると考えられる「時間範囲」は、統合の時間窓である。本論ではこの統合の時間窓の長さのみを取り上げる。

<sup>10</sup>三つの「時間範囲」について、その値と概略を述べる。

一つ目は、数 ms の時間範囲である。ギャップ検知、振幅変調検知などの実験を行った結果、数 ms の時間的ギャップ(隙間)や音の変化を知覚できることが報告された。これは主に、聴覚が精密な時間分解能をもつシステムであること、すなわち聴覚の時間分解能(temporal resolution)と関連して言及されることが多い。

二つ目は、本論で扱う、150–300 ms の時間範囲である。本文で述べるように、最小可聴閾の測定やラウドネス一致課題などより、この「時間範囲」内であれば持続時間が長くなると、ラウドネスが上昇して知覚され、最小可聴閾が下がるという現象が生じるが、この「時間範囲」外になるとこの現象が生じず、持続時間が長くなってもラウドネスも最小可聴閾も一定となると報告された。この現象における音の強さや持続時間の関係を、時間範囲の中の音のエネルギーを積分する式によって表すことができる。そのため、この時間範囲は「統合の時間窓」とも呼ばれる。「時間積分(temporal integration / temporal summation)」と関連して言及されることが多い。

三つ目は、2–20 s の時間範囲である。これは、騒がしい環境下で聞き取れなかった単語でも、前後の文脈によって情報を補完し、知覚することができることから、少なくとも数秒以上前の情報も保持されていると考えられている [4]。これは、短期記憶と関連して言及されることもある。

この三つの時間範囲では、それぞれ異なった現象に対し、異なった性質の知覚処理が行われる。特に一つ目と二つ目の時間範囲は、どちらも片方の時間範囲で生じる現象を、もう片方の時間範囲の知覚処理では説明できないというパラドクス(“resolution-integration paradox” [31, 20, 6])を孕んでおり、それぞれ知覚処理の方法が異なるのか、それともある1つの知覚処理の異なる側面を反映しているのかについては結論が出ていない [39]

## 2.5.2 統合の時間窓の長さ

先述した通り、統合の時間窓の範囲内であれば、刺激音の持続時間が長くなるにつれ、最小可聴閾が低下し、ラウドネスが増加する現象が生じる。

この現象が知られてから、多くの研究により、この現象における音の持続時間と音の強さの関係が調査されるようになった。そのため、統合の時間窓についての研究は、主に、最小可聴閾やラウドネスの測定を通して音の持続時間と音の強さの関係を記述することで行われる。また、統合の時間窓の長さがどのくらいかという問いも主に最小可聴閾やラウドネスの測定により調査されるが統合の時間窓の長さの調査は、脳波のミスマッチ陰性電位による研究からも行われている。

以下では、統合の時間窓の長さに関する報告を詳説する。まず、最小可聴閾やラウドネスの研究の報告について述べる。次に、脳波のミスマッチ陰性電位の研究の報告について述べる。

### 2.5.2.1 最小可聴閾やラウドネスの研究より報告される、統合の時間窓の長さ

統合の時間窓の長さがどのくらいかという問いは主に最小可聴閾やラウドネスの測定により調査される。このような研究では、持続時間を変化させたときの最小可聴閾やラウドネスを測定し、持続時間の増加と最小可聴閾やラウドネスの変化をプロットする。そして、持続時間が増加するにつれて最小可聴閾やラウドネスが変化する範囲と、持続時間が増加してもそれらが変化しない範囲との「境界の時間」を算出する。この境界の時間は「臨界継続時間 “critical duration”」と呼ばれ [26, 20, 30, 42, 43]、これが統合の時間窓の長さだと考えられる。加えて、いくつかの先行研究では周波数や音圧レベルなどの持続時間以外の音の要素が異なる場合にも臨界継続時間が変化するかどうかも調査されている。

以下では、まず、臨界継続時間の定義を確認する。次に、先行研究より報告される臨界継続時間の値を述べる。最後に、周波数や音圧レベルなどによって臨界継続時間が異なるかどうかについての報告を述べる。

#### 臨界継続時間の定義

多くの先行研究では臨界継続時間の定義についてはあまり言及されていない。だが、先行研究の中で臨界継続時間として報告される値より、その定義には種類があると考えられる。

先述した通り、臨界継続時間は、持続時間が増加するにつれ最小可聴閾やラウドネスが変化する範囲と、持続時間が増加してもそれらが変化しない範囲との境界の時間である。これは、持続時間を変化させたときの最小可聴閾やラウドネスを測定した後、次のように導かれる。まず、ある音の持続時間を変化させた時の (横軸)、その音の最小可聴閾が知覚されるために必要

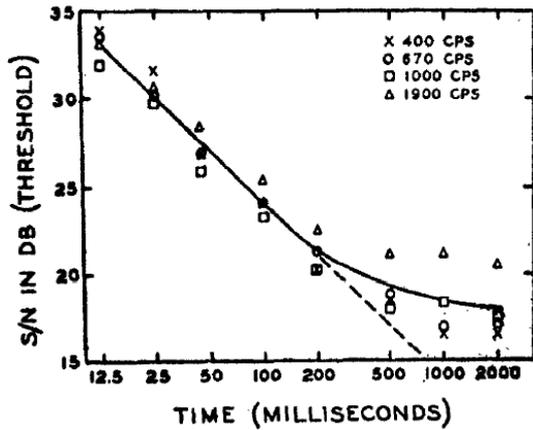


図 2.5.1: Garner ら [7] の図 3

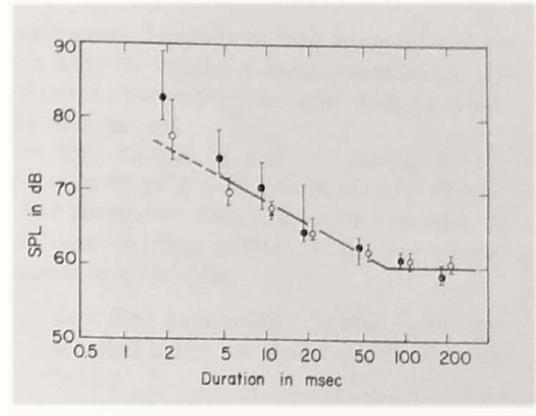


図 2.5.2: Scharf[26] の図 8(P.205)  
Scharf が引用した Port(1963) の図である。  
Port(1963)“Ueber die Lautstärke einzelner  
kurzer Schallimpulse,” Acustica, 13, 212–  
223.

な音の強さ (縦軸)、あるいは、その音がある一定のラウドネスに知覚されるために必要な音の強さ (縦軸) を図示する。次に、多くの場合、この図に、音の持続時間と音の強さの関係を近似した予測線が重ねられる。この予測線は曲線的な 1 本の場合と、傾きの異なる直線 2 本の場合がある。前者の曲線的な 1 本の場合の例として、図 2.5.1 に Garner & Miller[7] が報告した図を、後者の直線 2 本の場合の例として、図 2.5.2 に Scharf[26] がレビューにて報告した図を掲載した。そして、このような図より臨界継続時間が報告される。

報告された値から、その定義は以下の三つに分類されると考えられる。

一つ目の臨界継続時間の定義は、持続時間の増加に対して音の強さが線形に増加する、その最後の時刻である。例えば、この定義での臨界継続時間は、図 2.5.1 では 200 ms となり、図 2.5.2 では 70 ms 付近となる。以下では「線形変化の末の時刻としての臨界継続時間」と称する。

二つ目は、持続時間の増加に対して音の強さが漸近した時刻である。これは、予測線が曲線的な 1 本の場合に意味を持つ。すなわち、曲線的な予測線においてはどこまで持続時間が長くなっても音の強さが完全に一定になることがないため、ある持続時間において音の強さが「漸近した “asymptotic”」と判断することになる。例えば、Scharf[26] は、1 phon 以内のばらつきで漸近したと判断した<sup>11</sup>。図 2.5.1 では、2 dB 以内のばらつきを漸近したと判断すれば、この定義での臨界継続時間は 500 ms 付近となる。以下では、「漸近した時刻としての臨界継続時間」と称する。

三つ目は、持続時間の増加に対して音の強さが線形に増加する最後の時刻から、持続時間の増加に対して音の強さが漸近した時刻になるまでの時間範囲である。すなわち、先述した一つ目の時刻から二つ目の時刻までの時間範囲である。これも、二つ目と同様に、予測線が曲線的

<sup>11</sup>Scharf[26]P.209 より、以下の文章から推察される。

“( Accordingly, loudness level comes within 1 phon of its asymptotic value when duration reaches 180 msec.)”

な1本の場合に意味を持つ。図 2.5.1 では、2 dB 以内のばらつきを漸近したと判断すれば、この定義での臨界継続時間は 200–500 ms 付近となる。以下では、「線形変化の末から漸近した時刻までの時間範囲としての臨界継続時間」と称する。

## 臨界継続時間の値

先行研究より報告される臨界継続時間は大きく異なっており、50 ms 程度から 500 ms 以上にまで及ぶ。そのため、以下では便宜的に、報告される臨界継続時間によって先行研究をグループに分け、詳述する。そのグループは、臨界継続時間が 150 ms 以下、150–500 ms、500 ms 以上とした。最後に、臨界継続時間の値に関してレビューを行った Scharf[26] の報告を述べる。

以下で先行研究を詳述する際には、用いられた刺激音と報告された臨界継続時間のみを述べ、測定実験の詳細や聴覚の「積分時定数 time constant」には触れない。確かに、最小可聴閾の測定とラウドネスの測定における測定方法の種類や測定時の聴取状態などが臨界継続時間に影響を与えることも考えられる。だが、本論ではそれらは扱わない。また、聴覚の積分時定数は、持続時間の増加と音の強さの変化の関係を数式で表した時に用いられる定数であり、臨界継続時間と時定数を同一のデータより算出した場合は時定数の方が小さい値となることが知られている [20]。すなわち、聴覚の積分時定数は臨界継続時間と異なる概念であるため、本論では扱わない。

臨界継続時間が 150 ms 以下と報告した研究を挙げる。

Stevens & Hall[28] は、ホワイトノイズを刺激音として、ラウドネスの測定と最小可聴閾の測定を行った。ラウドネスの測定においては、持続時間を 10 水準 (5, 10, 18, 30, 50, 80, 130, 200, 300, 500 ms) で変化させ、最小可聴閾の測定にはこの 10 水準に 2 水準 (800, 1000 ms) を追加した計 12 水準で変化させた。音の提示音圧レベルも 6 水準 (36, 51, 66, 81, 96, 109 dB SPL) で変化させた。その結果、最小可聴閾での臨界継続時間は 230 ms であったが、閾値以上のラウドネス知覚における臨界継続時間は、提示音圧レベルによらず約 150 ms であったと報告した。彼らの用いた臨界継続時間の定義は、一つ目の「線形変化の末の時刻としての臨界継続時間」と推察された。

Miller[16] は、ホワイトノイズを刺激音として、最小可聴閾の測定とラウドネスの測定を行った。持続時間を 9 水準 (23, 35, 45, 58, 95, 200, 450, 810, 1550 ms) で変化させた。Miller は、この実験の結果とその他の実験の結果を、聴覚処理の仕組みの理論と合わせて考察し、臨界継続時間は約 65 ms であると結論付けた。彼は臨界継続時間の定義を述べていないため、それが上述の三つの定義のどれに相当するかは不明である。そこで、彼らの報告した図より推察すると、一つ目の「線形変化の末の時刻としての臨界継続時間」が音圧レベルによって異なると考えられる。音圧レベルごとの臨界継続時間は 60–145 ms であり、音圧レベルが小さくなるにつれ臨

界継続時間が長くなったといえる。

次に、臨界継続時間が 150–500 ms の間と報告した研究を述べる。

Munson[18] は、サイン波を刺激音として、ラウドネスの測定を行った。持続時間を 5 水準 (5, 10, 40, 100, 200 ms) で変化させた。Munson は、この実験結果より理論的な予測式を導出し、持続時間を 5 ms から 100 秒まで変化させたときのラウドネスの変化を予測し、図に示した。そして、持続時間が 300 ms までは、持続時間が増加するにつれラウドネスも増加するが、持続時間が 300 ms から 100 秒までは、持続時間が増加するにつれラウドネスが僅かに減少し続けると報告した。彼らは臨界継続時間について述べていない。そこで、音の持続時間が 1 秒以上になると聴覚の順応や疲労によりラウドネスが減少すること [26] を考慮し、Munson の報告する図中の 0 秒から 1 秒までに着目すると、持続時間が 300 ms から 1 秒までのラウドネスの減少が僅かであることが明らかである。その僅かなラウドネスの減少を漸近したとみなせば、二つ目の「漸近した時刻としての臨界継続時間」は 300 ms 付近であると推察できる。

Plomp & Bouman[21] は、最小可聴閾の測定を行った。持続時間を 13 水準 (1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, 10000 ms) で変化させ、周波数も 6 水準 (250, 500, 1000, 2000, 4000, 8000 Hz) で変化させた。Plomp らは、周波数が高くなると時定数が小さくなることのみ言及し、臨界継続時間については言及していない。そこで、彼らの報告する図より臨界継続時間を推察すると、三つ目の「線形変化の末から漸近した時刻までの時間範囲としての臨界継続時間」は、周波数によって異なり、周波数が低いと 200–1000 ms 程度であるが、周波数が高くなると 200–500 ms 程度であるといえる。

萩原 [43] は、ノイズを刺激音として、最小可聴閾の測定とラウドネスの測定を行った。持続時間を 8 水準 (5, 10, 20, 50, 100, 200, 500, 1000 ms) で変化させ、周波数を 6 水準 (250, 500, 1000, 2000, 4000, 8000 Hz) で変化させた。その結果、周波数を 250–8000 Hz まで変化させた時、二つ目の「漸近した時刻としての臨界継続時間」が 210–290 ms であったことから、臨界継続時間は周波数によって異ならず、一貫して 200–300 ms の範囲にあったと結論付けた。また、音圧レベルを増加させた時、臨界継続時間が変化したと報告した。音圧レベルが最も低い時、すなわち最小可聴閾では、臨界継続時間は 250 ms であるが、そこから音圧レベルを 60 dB 増加させた時のラウドネス知覚では、臨界継続時間が 100 ms であったと報告した。

Garner & Miller[7] は、純音を刺激音として、最小可聴閾の測定を行った。持続時間を 8 水準 (12.5, 25, 50, 100, 200, 500, 1000, 2000 ms) で変化させ、周波数も 4 水準 (400, 670, 1000, 1900 Hz) で変化させた。その結果、持続時間が 200 ms までは音響エネルギーが線形積分され、持続時間が 200 ms から 2000 ms までは線形ではないものの積分され続けるとのみ報告し、臨界継続時間については言及しなかった。また、周波数の違いについても言及しなかった。そこで、この報告から推察するならば、一つ目の「線形変化の末の時刻としての臨界継続時間」は 200 ms

であり、二つ目の「漸近した時刻としての臨界継続時間」は2000 ms以上であると考えられる。だが、Garnerらが報告する表では、持続時間が200 ms以上から2000 msまで変化するとき、最小可聴閾に必要な音の強さの変化量が4 dB以内であった。このことより推察するならば、二つ目の「漸近した時刻としての臨界継続時間」は、2 dB 差以内のばらつきを漸近したと考えれば500 msとなり、4 dB 差以内と考えれば200 ms付近となるとも推察される。なお、先述した通り、図 2.5.1 は Garner らが報告する図であるため、必要に応じて参照してほしい。

そして、臨界継続時間が500 ms以上と報告した研究を述べる。

Takeshima, Suzuki, Kono, & Sone[30] は、1 kHz の音を刺激音として、ラウドネスの測定を行った。持続時間を8水準(50, 100, 200, 500, 1000, 2000, 5000, 10000 ms)で変化させた。その結果、持続時間が500 msを超えても、持続時間が増加するにつれてラウドネスも増加するとのみ報告し、臨界継続時間については言及していない。そこで、この報告と Takeshima らが報告した図より推測すると、二つ目の「漸近した時刻としての臨界継続時間」は500 ms以上であるといえる。

最後に、Scharf[26] のラウドネス研究のレビュー論文における臨界継続時間の報告を挙げる。Scharf は、レビューを通して、多くの研究室で多くの参加者に実験を行った研究<sup>12</sup>に重きを置き、その他の先行研究を踏まえて考えると、二つ目の「漸近した時刻としての臨界継続時間」が150–300 ms の範囲のどこかにあると考えるのが妥当であると結論づけた<sup>13</sup>。また、音圧レベルや周波数によって臨界継続時間が異なるかどうかについては断定的なことは言えないと報告した。

以上のように、報告される臨界継続時間は、その定義もその値も様々である。最も広く調査を

---

<sup>12</sup>“round robin test”のことであり、最終レポートは以下の論文である。本論では原著が手に入らなかったため、参考にできなかった。

Pedersen, O. J., Lyregaard, P. E., and Poulsen, T. (1977). The round robin test on evaluation of loudness level of impulsive noise. Acoustics Laboratory, Technical University of Denmark.

<sup>13</sup>Scharf[26] は、

「[持続時間の増加に対してラウドネスが一定となるという意味での] ラウドネスが最大値に達する持続時間に関しては先行研究間でよく一致している

“Best agreement has been reached with respect to the question of the duration at which loudness reaches its full value. (In terms of the equal-loudness contour of loudness constant as duration gets longer)” P.209」

と述べたため、Scharf が用いた臨界継続時間は二つ目の「漸近した時刻としての臨界継続時間」であると考えられる。その値について、その続きで、

「その値は時定数にして80 msであり、それを超えるとラウドネスは持続時間に応じて増加するが、それは僅かにである。[臨界継続時間について述べるならば、] ラウドネスは、測定誤差を含め、150–300 ms の間で最大値に達する。

“A time constant of 80 msec summarizes the round-robin data very well and is close to most of the values for the time constant previously proposed. Beyond 80 msec, loudness continues to increase with duration but more gradually. Loudness may be considered, within the error of measurement, to reach full value between 150 and 300 msec.” P.209」

と述べている。なお、ここで言う時定数を超えた後のラウドネスの増加は、漸近したと判断できるくらいの変化(1 phone 以内)であるということの意味すると考えられる。

行った Scharf の報告が最も信頼性が高いと言える。すなわち、統合の時間窓の長さは 150–300 ms の範囲のどこかにあると考えられる。

### 持続時間以外の音の要素が異なるときの臨界継続時間の値

今まで述べた先行研究より、音の音圧レベルや周波数が異なる場合に臨界継続時間が変化するかどうかについての報告を、再度述べる。先行研究の詳細は先述したため、割愛する。さらに、音の種類が異なる場合に臨界継続時間が変化するかについて考える。

音圧レベルが異なる場合に臨界継続時間も異なるかどうかを検討した研究について述べる。

Miller[16] は、最小可聴閾であるか最小可聴閾以上のラウドネスの知覚であるかによっても、ラウドネス知覚における音圧レベルが異なる場合でも、音圧レベルが小さくなるにつれ臨界継続時間が長くなったと報告した。

同様のことを萩原 [43] も報告した。

一方、Stevens ら [28] は、最小可聴閾であるか最小可聴閾以上のラウドネスの知覚であるかによって臨界継続時間が異なるが、ラウドネスの知覚においては音圧レベルが異なっても臨界継続時間は一定であると報告した。

そして、Scharf[26] は、ラウドネス知覚において、音圧レベルが異なる場合に臨界継続時間が異なるか否かは文献ごとに異なっており、断定的なことは言えないと述べた。

周波数が異なる場合に臨界継続時間も異なるかどうかを検討した研究について述べる。

Plomp ら [21] は、臨界継続時間については言及していないが、彼らの報告する図より推察すると、最小可聴閾における、三つ目の「線形変化の末から漸近した時刻までの時間範囲としての臨界継続時間」が周波数によって異なった。その値は、周波数が低いと 200–1000 ms 程度であるが、周波数が高くなると 200–500 ms 程度であったといえる。

一方、萩原 [43] は、周波数によって臨界継続時間は異ならなかったと結論付けた。

また、Scharf[26] は、ラウドネス知覚において、周波数によって臨界継続時間が異なるか否かは文献ごとに異なっており、断定的なことは言えないと述べた。

続いて、音の種類による臨界継続時間の差異に関して考える。先行研究で刺激音としてしばしば用いられるのは、ホワイトノイズや純音、あるいはパルス音である。これらの音の種類により、臨界継続時間が異なるかについて述べた研究は殆どない。

だが、Eddins[6] は、音の持続時間と音の強さの関係 (“time-intensity trade” / “trading relation”) を数式で表すと、サイン波よりもノイズ音の方が複雑であること、そしてサイン波の式をノイズ音にそのまま適用できないことを報告した。この報告より、臨界継続時間においても、サイン波とノイズ音で異なる可能性が考えられるが、これについては今後の研究が必要である。

また、複合音を用いた時の臨界継続時間も殆ど研究されていないため、今後の研究が必要である。

以上のように、音圧レベルが増加するにつれ、また周波数が高くなるにつれ臨界継続時間が短くなる可能性が示された。だが、矛盾する報告もあるため、これらは可能性にすぎず、結論付けるためには今後の研究が必要である。

加えて、音の種類を、サイン波からノイズ音や複合音に変化させた時に臨界継続時間が変化するかについては、殆ど研究されていない。これも、今後の研究が必要である。

### 2.5.2.2 脳波のミスマッチ陰性電位の研究より報告される、統合の時間窓の長さ

ここからは、脳波のミスマッチ陰性電位の研究より報告される、統合の時間窓の長さについて述べる。

ミスマッチ陰性電位 (mismatch negativity; MMN) とは、脳波の「事象関連電位」の一つである。「事象関連電位」とは、ある事象 (刺激、イベント) を受けて、その事象の処理に特化した脳部位が活性化する反応の総称である。その一つであるミスマッチ陰性電位は、先行する刺激と後続する刺激との間の変化が検出され、かつ、それらの刺激が同じ統合の時間窓の範囲内で提示された場合のみ、観測される。例えば、音と同じ時間間隔で繰り返される中である音が欠如した場合、ある音とその一つ前の音との時間間隔、すなわち音の繰り返しの時間間隔が統合の時間窓の範囲内である場合のみ、ミスマッチ陰性電位が観測される。このことから、統合の時間窓がミスマッチ陰性電位を用いて調べられている [40, 39]。

この、ミスマッチ陰性電位による統合の時間窓の研究における、統合の時間窓の長さの報告について述べる。

Yabe, Tervaniemi, Sinkkonen, Huottilainen, Ilmoniemi, & Näätänen[38] は、持続時間が 50 ms で立ち上がり時間が 5 ms で立ち下がり時間が 5 ms の 1000 Hz のサイン波を刺激音とし、その音を 5 水準の時間間隔 (125, 150, 175, 200, 350 ms) ごとに繰り返す音列を提示し、脳波を計測した。その結果、統合の時間窓の長さが 150–175 ms の範囲にあると報告した。

Russeler, Altenmüller, Nager, Kohlmetz, & Münte[25] は、音楽家と非音楽家の間で統合の時間窓の長さが異なるかを調査した。持続時間が 50 ms で立ち上がり時間が 5 ms で立ち下がり時間が 5 ms の 1000Hz のサイン波を刺激音とし、その音を 4 水準の時間間隔 (100, 120, 180, 220 ms) ごとに繰り返す音列を提示し、脳波を計測した。その結果、音楽家は統合の時間窓が 220 ms 以上であるが、非音楽家は統合の時間窓が 120 ms 以下であると報告した。このことより、音楽家の統合の時間窓は非音楽家よりも長く、音楽訓練によって統合の時間窓が長くなる可能性が明らかとなったと結論付けた。

## まとめ

音の知覚は聴覚情報処理を経て形成されるため、PATの知覚も聴覚情報処理の何らかの影響を受けるはずである。

聴覚情報処理では、ある「時間範囲」分を保持し、その時間分の情報をまとめて処理する仕組みがある。この「時間範囲」は三つに大別されるが、そのうちの一つである「統合の時間窓」では、この範囲内であれば持続時間が増加するにつれ、最小可聴閾が低下し、ラウドネスが上昇して知覚される現象が生じるが、この範囲を超えると、この現象は生じず、持続時間が増加しても最小可聴閾やラウドネスは変化しないことが知られている。

今まで述べてきたように、この統合の時間窓の長さは文献ごとに異なる値が報告されるが、本論では、それらの研究をレビューしたScharfの結論に重きを置きつつ、その他の最小可聴閾やラウドネスの研究の結果とミスマッチ陰性電位の研究の結果を参照し、統合の時間窓の長さは150–300 msの範囲のどこかにあると考える。また、音圧レベルが増加するにつれ、また周波数が高くなるにつれ、臨界継続時間が短くなる可能性が示された。加えて、音楽訓練を受けると統合の時間窓が長くなる可能性も示された。

PATの研究でも実験で刺激音の持続時間を1 msから400 msあるいは500 ms程度まで変化させるが、この持続時間の範囲が「統合の時間窓」の時間範囲と重なる。そのため、持続時間を変化させたときのPATの知覚にも、「統合の時間窓」が影響する可能性が考えられる。その影響については、次章にて検討する。

## 2.6 本論の仮説

2.4章で述べたように、PATの先行研究より、持続時間が短い音よりも長い音の方が、また立ち上がり時間が短い音よりも長い音の方が、PATが時間的後方に位置することが報告された。また、持続時間と立ち上がり時間との交互作用も報告された。加えて、周波数や提示音圧レベルが変化すると、PATも変化する可能性が示唆された。だが、持続時間や立ち上がり時間がどのくらい長くなるとPATがどのくらい時間的後方へ移動するのか、すなわち、持続時間の増加や立ち上がり時間の増加とPATの移動との関係は殆ど調査されていない。加えて、周波数や提示音圧レベルがPATをどのように変化させるかについては結論が出ていない。

そして、PATの先行研究ではPATと統合の時間窓との関係が述べられていないが、2.5節で述べたように、PATの先行研究において刺激音の持続時間を統合の時間窓の範囲内から範囲外まで変化させるならば、PATの知覚が統合の時間窓の影響を受ける可能性が考えられる。

本節にて、PATの知覚に統合の時間窓がどのように影響するかを検討し、仮説を導く。

## 2.6.1 仮説

2.5章で述べたように、最小可聴閾やラウドネスの研究より、統合の時間窓の範囲内であれば、持続時間が増加するにつれ最小可聴閾が低下し、ラウドネスが上昇して知覚される。だが、統合の時間窓の範囲外になると、持続時間が増加しても最小可聴閾やラウドネスは変化しない。これを言い換えると、「持続時間の増加と最小可聴閾やラウドネスの変化との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」といえる。このことをPATの知覚に置き換えてみれば、「『PATを移動させる音の要素』の変化とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」といえる。

では、この『PATを移動させる音の要素』は、何だろうか。まず、持続時間と立ち上がり時間が挙げられる。この二つの要素は、2.4節で述べたように、先行研究にて一貫してPATを移動させることが報告されている。そのため、持続時間と立ち上がり時間は確かにPATを移動させる要素であると考えられる。次に、周波数と提示音圧レベルも挙げられる。だが、この二つはPATを移動させる可能性が示唆されているものの、結論は出ていない。むしろ、周波数と提示音圧レベルは、2.5節で述べたように、統合の時間窓の長さを変化させる可能性が示唆されている。このことを考慮すると、周波数と提示音圧レベルはPATを移動させるのではなく、統合の時間窓の長さを変化させ、そして統合の時間窓の長さが変化したことで、その影響を受けたPATの知覚が変化した可能性が考えられる。従って、『PATを移動させる音の要素』に相当するのは、持続時間あるいは立ち上がり時間であると考えられる。

このことから、PATの知覚に統合の時間窓が与える影響についての仮説の候補として、以下の二つが挙げられる。

仮説候補1「持続時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」

仮説候補2「立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」

だが、2.4章で述べたように、先行研究では、音の持続時間と立ち上がり時間との交互作用が報告されている。そうであれば、上記の二つの仮説のうち、より確からしいのは、二つ目の仮説であると考えられる。

以上のことより、本論では、以下を仮説とし、これを検証する。

「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」

## 2.6.2 仮説検証の方法

以上の仮説を検証するため、音の持続時間と統合の時間窓を変化させた時に、立ち上がり時間の増加と PAT の移動との関係がどのように異なるかを調査する。

### 実験方法

実験にて、持続時間と立ち上がり時間を変化させた刺激音の PAT を測定する。

本論では「立ち上がり時間」ではなく、「持続時間に占める立ち上がり時間の割合 (数式では (立ち上がり時間)/(持続時間)×100[%]である)」を変化させる。以後、これを「立ち上がり時間の割合」と略す。これは、持続時間と立ち上がり時間との交互作用を報告した P.Vos[35] の研究を参考としたためである。

刺激音は、合成音とする。確かに、楽器音と合成音とで PAT の知覚に差異がある可能性が示されているが、その結論は出ていない。仮説検証のためには、精密に統制した実験を行う必要がある。だが、楽器音は複雑であるため、統制しきれない要素が出てくる可能性が高い。そのため、本論では、刺激音は合成音とし、合成音の中でも最も単純なサイン波を用いる。

PAT 測定課題には、タッピング課題と同時調整課題、等時調整課題がある。本論では多くの刺激音を用いることになるため、刺激音の数に対する実験時間が短い方が望ましく、音楽訓練を受けていない参加者にも実験を行うため、実験課題はより単純でより労力が少ない方が望ましい。タッピング課題の利点として、判断に推論が含まれないため多くの参加者にとって簡単であることや、刺激音の数に対する実験時間が短いことが挙げられている [15, 32]。ため、本論ではタッピング課題を用いるのが望ましいと考えられる。だが、タッピング課題を用いるには、タッピング課題で得られる結果がその他の PAT 測定課題より得られる結果と有意に異なることが前提となる。従って、これを実際にタッピング課題を用いて PAT を測定する前に検証する必要がある。

加えて、本論では統合の時間窓を直接測定しない。PAT の知覚に統合の時間窓が関係しているのかを確認するため、統合の時間窓がある長さの時の刺激音の PAT と、統合の時間窓の長さが異なる時の刺激音の PAT とを比較する必要がある。そのため、先行研究にて統合の時間窓の長さを変化させる可能性が指摘されている、提示音圧レベル、周波数、音楽訓練を受けたことがあるかどうか (音楽訓練の有無) を変化させたときの PAT を比較する必要がある。

従って、本論では、持続時間と「立ち上がり時間の割合」を変化させた刺激音を用いてタッピング課題を行い、PAT を測定する。また、提示音圧レベルや周波数、音楽訓練を受けたことがあるかどうか異なる場合の PAT も測定する。

## 分析方法

続いて、仮説を検証するための分析方法を述べる。

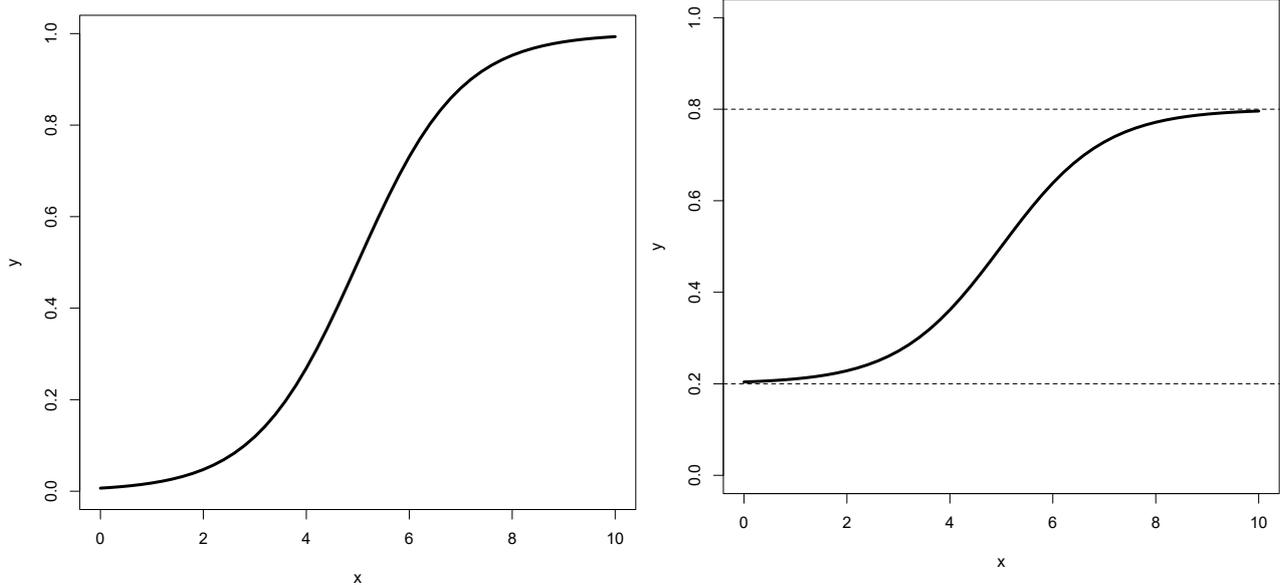
まず、実験結果より PAT を数値化する。本論では、 $\Delta PC_{A,R}$  を用いる。 $\Delta PC_{A,R}$  は、ある基準の刺激音  $R$  とそれ以外の刺激音  $A$  との PAT の差分のことである。 $\Delta PC_{A,R}$  を用いることで、3 音以上の音の PAT の関係の比較ができる。詳細は 2.2 節を参照してほしい。

次に、「立ち上がり時間の増加と PAT( $\Delta PC_{A,R}$ ) の移動との関係」を表す。本論では、この関係を図に示す。図は、ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、PAT (縦軸) が示される。すなわち、持続時間ごとの「立ち上がり時間の割合の増加と PAT の移動との関係」が示される。

続いて、持続時間ごとの「立ち上がり時間の増加と PAT( $\Delta PC_{A,R}$ ) の移動との関係」を最も近似する回帰モデルを求め、図に重ねる。このようにすると、実験における測定誤差を減らし、その関係をより明確に示すことができる。

では、この回帰モデルは何が適切であろうか。PAT の先行研究では、「立ち上がり時間」の増加と PAT の移動との関係は調査されていない。そのため、仮に、本論の実験にて「立ち上がり時間の割合」を 0% から 100% まで増加させたとして、その時の PAT( $\Delta PC_{A,R}$ ) が直線的に変化する確証はない。もしかしたら、「立ち上がり時間の割合」が小さい時から大きい時まで通して PAT( $\Delta PC_{A,R}$ ) が直線的に変化するかもしれないし、もしかしたら、「立ち上がり時間の割合」が小さい時あるいは大きい時には PAT( $\Delta PC_{A,R}$ ) が変化しないかもしれない。従って、これらの可能性に対応できる回帰モデルが望ましい。加えて、本論の仮説では、「立ち上がり時間の割合の増加と PAT( $\Delta PC_{A,R}$ ) の移動との関係」が持続時間によって異なると考えている。そのため、「立ち上がり時間の割合」を増加させた時の PAT がとる、最小値と最大値が持続時間によって異なる可能性も考えられる。従って、これにも対応できる回帰モデルが望ましい。これらの条件に合う回帰モデルは、4 パラメータロジスティック回帰モデルであるため、本論ではこれを用いる。

4 パラメータロジスティック回帰モデルとは、2 パラメータロジスティック回帰モデルを拡張したモデルである。2 パラメータロジスティック回帰モデルは、従属変数がカテゴリや比率、2 値変数である場合によく適用される。予測値が最小値 0 から最大値 1 までをとり、その間を S 字に変化していく (独立変数の変化に対する従属変数の変化量が、徐々に大きくなり、次第に従属変数の変化量がほぼ一定となり、さらに徐々に小さくなる形) モデルである。そのパラメータは、傾きと中央の位置の 2 つである。この 2 パラメータロジスティック回帰モデルを拡張したものが 4 パラメータロジスティック回帰モデルである。そのパラメータは、傾きと中央の位置、そして最小値と最大値である。このように拡張されたことで、予測値がデータに合わせた最小



(a) 2パラメータの場合

横軸が独立変数  $x$ 、縦軸が従属変数  $y$  である。パラメータは傾き  $b = -1$ 、中央の位置  $e = 5$  である。

(b) 4パラメータの場合

横軸が独立変数  $x$ 、縦軸が従属変数  $y$  である。パラメータは傾き  $b = -1$ 、中央の位置  $e = 5$ 、そして最小値  $c = 0.2$ 、最大値  $d = 0.8$  である。

図 2.6.1: ロジスティック回帰モデルの予測線の例

値と最大値をとることができ、予測値がデータに合わせた最小値から最大値までの間をS字に変化していくモデルになる。4パラメータロジスティック回帰モデルの式は以下である。従属変数を  $y$  とし、独立変数を  $x$ 、最小値を  $c$ 、最大値を  $d$ 、傾きを  $b$ 、中央の位置を  $e$  とする。

$$y = c + \frac{d - c}{1 + \exp(b(x - e))} \quad (2.10)$$

この回帰モデルの予測線の形は、S字曲線の形 (独立変数の変化に対する従属変数の変化量が、徐々に大きくなり、次第に変化量が一定になり、そして変化量が徐々に小さくなる、S字の全体) である。だが、4つのパラメータの値を変化させると、S字の真ん中の部分 (独立変数の変化に対する従属変数の変化量がほぼ一定であり直線に近い形となる部分) のみを当てはめたり、S字曲線の下の方の形 (独立変数の変化に対する従属変数の変化量が徐々に大きくなる部分) のみを当てはめたり、S字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) のみを当てはめたりすることも可能である。例として、図 2.6.1 に2パラメータロジスティック回帰モデルの予測線 (左図) と4パラメータロジスティック回帰モデルの予測線 (右図) を載せた。パラメータの値は、傾き  $b = -1$ 、最小値  $c = 0.2$ 、最大値  $d = 0.8$ 、中央の位置  $e = 5$  である。

最後に、「立ち上がり時間の割合の増加と  $\text{PAT}(\Delta PC_{A,R})$  の移動との関係」が持続時間によっ

て異なるのかを検討する。持続時間ごとの「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,R})$  の移動との関係」は、先述したように、4パラメータロジスティック回帰モデルのパラメータの値を、ひいてはロジスティック回帰モデルの予測線の形を変化させる。そのため、視覚的にはその関係が予測線の形に表われる。一方、特にその関係が表れる数値としては、「立ち上がり時間の割合が0の時の予測値 (PAT) と立ち上がり時間の割合が100の時の予測値 (PAT) との差分」と「パラメータの一つである傾き  $b$  の値」との2つの指標であると考えられる。前者は、立ち上がり時間の割合を0%から100%まで変化させたときに、 $PAT(\Delta PC_{A,R})$  がどのくらい移動するかを示し、後者は、立ち上がり時間の割合の変化がどのくらい PAT を変化させるかを示すからだ。従って、まず、ロジスティック回帰モデルの予測線の形の視認を行い、次に、2つの指標を用いてウォード法によるクラスター分析を行い、「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,R})$  の移動との関係」が持続時間によって異なるかを検証する。

## 実験構成

次章以降にて、仮説の検証のための実験をいくつか行った。ここではその実験の構成を述べる。必要に応じて、1章で掲載した図 1.4.1 の本論の研究フローを参照してほしい。

3章にて、三つの予備実験を行った。一つ目に、タッピング課題の妥当性を検証した。2.3.1章の中でタッピング課題の問題点として挙げたように、タッピング課題ではある A 音の PAT を測定するために必要な、A 音が繰り返される数 (タッピングの数) が明らかになっていない。従って、まずはそれを検討した。二つ目に、求められた適切なタッピングの数を用いたタッピング課題と、その他の PAT 測定課題である同時調整課題と等時調整課題とで、測定された PAT が有意に異ならないかを確認した。その際、3つの課題の PAT を、 $\Delta PC_{\text{test,base}}$  で比較した。合わせて、提示音圧レベルを変化させた場合に PAT が変化するかどうかを確認した。三つ目に、仮説検証のためには、本実験で用いる持続時間や「立ち上がり時間の割合」をいくつかの水準で変化させるのが適切かを検討した。合わせて、サンプルサイズも検討した。

5章と6章、7章にて、本実験として3つの実験を行った。

まず、5章にて、実験1を行った。音楽訓練を受けたことがある参加者 (音楽群) に対しタッピング課題を行った。刺激音はサイン波とし、その持続時間と立ち上がり時間の割合を変化させた。それらを変化させる水準は予備実験にて検討したものをを用いた。また、周波数は440 Hzとした。これは、現在の音楽における基準となる周波数だからであり、JIS規格でも楽器用標準周波数はト音記号の音名“イ”の音の周波数で440Hzであると定められている [13] からである。提示音圧レベルは予備実験で用いた水準では PAT は変化しないことが明らかとなったので、参加者間で固定した。この実験結果より、仮説を検証した。

次に、6章にて、実験2を行った。統合の時間窓の長さが異なる場合として音楽訓練の有無

を取り上げた。実験1の音楽群よりも統合の時間窓が短いと考えられる、音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行った。実験1と同様に、刺激音はサイン波とし、その持続時間と立ち上がり時間の割合を変化させた。持続時間は実験1より1水準減らしたが、立ち上がり時間の割合は実験1と同じであった。周波数も実験1と同様に440 Hzとし、提示音圧レベルも実験1と同様に参加者間で固定した。この実験結果より、仮説を検証した。

最後に、7章にて、実験3を行った。統合の時間窓の長さが異なる場合として周波数を取り上げた。実験1の440 Hzの刺激音よりも統合の時間窓が短いと考えられる、より高い周波数の刺激音を用いて、実験1と同様の音楽群に対しタッピング課題を行った。実験1と同様に、刺激音はサイン波とし、その持続時間と立ち上がり時間の割合を変化させた。それらを変化させる水準は、実験2と同じであった。周波数は3520 Hzとした。これは実験1で用いた周波数440 Hzの3オクターブ上である。先行研究を参考とし、また、ピアノの音域内であることより選択された。提示音圧レベルは実験1と同様に参加者間で固定した。この実験結果より、仮説を検証した。

## 2.7 まとめ

本論では、「時間の流れの中に配置された、知覚された音のアタックの瞬間」と定義される“Perceptual Attack Time; PAT”が、音の何の要素によってどのように移動するかを調査する。本論の目的は、PATの先行研究と聴覚情報の統合の時間窓の先行研究より導出した、「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」という仮説の検証である。

仮説検証のため、実験にて、持続時間と「立ち上がり時間の割合」を変化させたサイン波を用いてタッピング課題を行い、PATを測定する。その後、分析と考察を通して、「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,R}$ )の移動との関係」が、音の持続時間が統合の時間窓の範囲内か範囲外かによって異なるかどうかを検証する。

## 第3章 予備実験

仮説の検証のため、次章からの本実験にてPATを測定する。刺激音はサイン波とし、持続時間と「立ち上がり時間の割合(持続時間に占める立ち上がり時間の割合であり、数式では(立ち上がり時間)/(持続時間) $\times 100$ [%]である)」を変化させる。そして、この測定実験は音楽訓練を受けたことがある参加者(音楽群)と音楽訓練を受けたことがない参加者(一般群)とに対し行われる。

本章では、その前に予備実験を行う。

前章で述べた通り、PAT測定課題には、タッピング課題、同時調整課題、等時調整課題の3つがあるが、本論で用いるPAT測定はタッピング課題が望ましい。だが、タッピング課題を用いるには2つの問題がある。一つ目は、2.3.1項で述べたように、タッピング課題であるA音のPATを測定するために必要なA音が繰り返される数(タッピングの数)が明らかになっていないことだ。二つ目は、2.3.4項で述べたように、先行研究より3つの課題を同一の指標で比較すれば3つの課題間に有意差がない可能性が推察されたが、その確証は未だ得られていないことである。予備実験において、この2つの問題を検討する。これらが解消されたならば、本実験でタッピング課題を用いることができるだろう。

加えて、本実験で用いる刺激音の、持続時間と立ち上がり時間の割合をどのような水準で変化させるのが適切か、また、サンプルサイズはいくつが適切かも予備実験で検討する必要がある。

従って、本章では以下の3つの予備実験を行った。

予備実験1において、タッピング課題で、あるA音のPATを測定するために必要な、A音が繰り返される数(タッピングの数)を検討した。

予備実験2において、タッピング課題、同時調整課題、等時調整課題を同一の指標で比較した。このタッピング課題での各刺激音のタッピングの数は、予備実験1で求めた数とした。また、合わせて、提示音圧レベルを変化させた場合にPATが変化するかどうかを確認した。

予備実験3として、仮説の検証のためには、本実験で用いる刺激音の、持続時間と立ち上がり時間の割合をどのような水準で変化させるのが適切かを検討した。合わせて、適切なサンプルサイズも検討した。

## 3.1 予備実験1: タッピング課題におけるPAT測定に必要なタッピングの数の検討

タッピング課題とは、あるA音が、メトロノームのように等時間間隔に、数十回繰り返される音列を聴きながら、各音に合うように指でタッピングを行う課題である。本節にて、このタッピング課題でPATを測定するために必要なタッピングの数を検討した。

### 3.1.1 実験方法

#### 3.1.1.1 参加者

13人であった。内1人は実験者であった。彼らはみな、東京藝術大学大学院音楽音響創造研究分野に所属していた。また、参加者の殆どが音楽訓練を受けたことがあった。本論における音楽訓練を受けたことがある参加者の定義は「今までに定期的な演奏訓練の経験があった人」であった。

#### 3.1.1.2 刺激音

440 Hzのサイン波を用いた。持続時間と立ち上がり時間を変化させた<sup>1</sup>。持続時間は4水準(50, 75, 100, 125 ms)とし、立ち上がり時間は5水準(0, 10, 20, 30, 40 ms)とした。従って、計20個の刺激音が作成された。

立ち上がり部分と立ち下がり部分は、振幅エンベロープにおいて線形に変化させた。そのため、立ち上がり時間が0 msの刺激音は、振幅の急峻かつ大幅な変動によって、目的の周波数以外をも含むノイズが僅かに発生した。その影響を軽減するため、先述した立ち上がり時間の水準が0 msの刺激音のみ、実際には1 msの立ち上がり時間をもたせた。同様の理由で、立ち下がり時間も1 msで固定した。このように、実際には1 msの立ち上がり時間と立ち下がり時間であるが、以下では簡略的に0 msと記す。

以上の刺激音は、PC上でPythonを用いて生成され、標本周波数48000 Hz、量子化ビット数24 bitのモノラル音源として書き出された。

#### 3.1.1.3 実験手続き

「各音の鳴った瞬間とタップした瞬間が一致するよう」にタップするよう教示した。「各音の鳴った瞬間」は参加者の解釈にゆだねた。タップ箇所はマイク内蔵オーディオ・インターフェー

<sup>1</sup>この予備実験1では、「立ち上がり時間」を変化させた。

「立ち上がり時間の割合(持続時間に占める立ち上がり時間の割合であり、数式では(立ち上がり時間)/(持続時間)×100[%]である)」ではなかった。

ス (Line6 Sonic Port VX) のマイクグリル近傍のボディ部分を指定した。できる限り安定してタップできる体勢で、できる限り一定の強さでタップするよう教示した。

本番前にトレーニング課題を行ってもらった。刺激音は、持続時間 50 ms、立ち上がり時間 0 ms、立ち下がり時間 0 ms の 440 Hz のサイン波とした。この刺激音を 700 ms ごとに 20 回繰り返した音列を 2 度提示し、それぞれにタッピングさせた。

本番では、同一刺激音を 700 ms ごとに 110 回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため全 20 音列であった。これらの 20 音列は、ランダムに 10 音列を連結したものを 1 試行とし、全 2 試行が参加者に提示された。音列の開始合図音は、持続時間 1 秒、立ち上がり時間 0 ms、立ち下がり時間 0 ms の 880 Hz のサイン音であった。試行間に 5 分間の休憩が挟まれた。本番の所要時間は約 45 分であった。

音列の提示音圧レベルは、「自分のタップの音が聞こえる範囲」で、各参加者が好むレベルに調整させた。実験室で提示音圧レベルを測定したところ、参加者ごとの、全刺激音平均の提示音圧レベルは 76–79 dB であった<sup>2</sup>。

フィードバック音は提示しなかった。だが次項で述べるように、実験に用いたヘッドフォンはセミオープン型ヘッドフォンであったため、殆どの参加者は自身のタップ音が聞こえていたと考えられる。

#### 3.1.1.4 実験装置

刺激音は、PC、マイク内蔵オーディオ・インターフェース (Line6 Sonic Port VX)、セミオープン型ヘッドフォン (AKG K240 MK II) を通して参加者の両耳に提示された。各参加者のタッピングは、刺激音提示と同期して、マイク内蔵オーディオ・インターフェース (同上) を通し、PC 上のソフトウェアにて録音された。

刺激音の再生とタッピングの録音との遅延時間は、標本周波数 48000 Hz で 6 サンプルであった。

### 3.1.2 分析

まず、タッピング課題においてある  $A$  音の PAT を測定するために必要な  $A$  音が繰り返される数 (タッピングの数) とは、どのような数かを考える。2.3.1 項で述べたように、タッピング課題より PAT として算出されるのは、 $\Delta PC_{A,R}$  である。この  $\Delta PC_{A,R}$  は基準音である  $R$  音とその他の刺激音である  $A$  音の同調のズレ量の平均値より算出される。同調のズレ量の平均値と  $\Delta PC_{A,R}$  の算出では参加者のタッピングに含まれるバイアス  $b$  が一定であると仮定される。このバイアス  $b$  が一定であるという仮定は、ある参加者の中でのタッピングに含まれるバイアス (以下では「タッピングバイアス」と称する) が正規分布に従うため平均すると一定の値になる

<sup>2</sup>各刺激音の立ち上がり時間が異なることで、各刺激音の提示音圧レベルも異なる。そのため、各刺激音の提示音圧レベルを平均し、全刺激音平均の提示音圧レベルを報告した。

という仮定だと考えることができる。そうであれば、同調のズレ量の平均値と  $\Delta PC_{A,R}$  がタッピングバイアスに影響されないようにするには、一定量のタッピングの数が必要となる。だが、実験するうえでは各刺激音に対するタッピングの数はできる限り少ない方が望ましい。すなわち、タッピング課題において、ある  $A$  音の PAT を測定するために必要な  $A$  音が繰り返される数(タッピングの数)は「同調のズレ量の平均値と  $\Delta PC_{A,R}$  がタッピングバイアスにできる限り影響されない範囲で、かつ、できる限り少ない数」であると考えられる。

同調のズレ量の平均値がタッピングバイアスにできる限り影響されなければ、その差分である  $\Delta PC_{A,R}$  も影響されなくなる。従って、両者がタッピングバイアスにできる限り影響されないことを確認するには、すなわち、同調のズレ量の平均値のみを確認し、 $\Delta PC_{A,R}$  を確認しなくて良い。同調のズレ量の平均値とタッピングバイアスの関係を知りたいが、タッピングバイアス<sup>3</sup> は未知の値  $\tau_{\text{attack}}$  から算出されるものであるためタッピングバイアスそのものを分析に用いることはできない。そこで、同調のズレ量の平均値と同調のズレ量の平均値の標準誤差<sup>4</sup> を、その算出に用いるタッピングの数を変化させて算出することでタッピングバイアスの影響が小さいことを確認する。理論上、平均値は、算出に用いるデータ数が増えるにつれ値が一定の値に近づき、平均値の標準誤差は、算出に用いるデータ数が増えるにつれ値が小さくなる。従って、「同調のズレ量の平均値と  $\Delta PC_{A,R}$  がタッピングバイアスにできる限り影響されない範囲で、かつ、できる限り少ない数」は、「同調のズレ量の平均値ができる限り一定の値に近い値をとり、同調のズレ量の平均値の標準誤差ができる限り小さい値をとる、その範囲において最も少ないタッピングの数」と考えることができる。

これを求めるため、以下の流れで分析を行った。

まず、タッピングの録音データよりタップ時刻  $t_{\text{tap}}(A, i)$  を算出した。タップ時刻は同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  の算出に必要なだからだ。この方法の詳細は後述の 4.2.1 項を参照してほしい。

次に、算出に用いるタッピングの数を変化させて同調のズレ量の平均値とその標準誤差<sup>5</sup>を求めた。ある  $A$  音が繰り返される音列に対するタッピングのデータからいくつかのタッピング

<sup>3</sup>13 ページの、タッピング課題における  $\Delta PC_{A,B}$  の算出方法の項において、タップ時刻  $t_{\text{tap}}(A, i)$  は、タップしようと狙った PAT  $\tau_{\text{attack}}(A, i)$  にタッピングバイアス  $b$  が付与されたものであると述べた。

$$t_{\text{tap}}(A, i) = \tau_{\text{attack}}(A, i) - b$$

そうであれば、以下のように展開することができる。

$$\begin{aligned} t_{\text{tap}}(A, i) &= \tau_{\text{attack}}(A, i) - b \\ t_{\text{tap}}(A, i) - \tau_{\text{attack}}(A, i) &= -b \end{aligned}$$

すなわち、タッピングバイアスは PAT  $\tau_{\text{attack}}(A, i)$  とタップ時刻  $t_{\text{tap}}(A, i)$  との差分であること、ひいては未知の値  $\tau_{\text{attack}}$  から算出されることが明らかである。

<sup>4</sup>平均値の標準誤差とは平均値がサンプリングごとにどのくらいばらつくかを表すものである。

<sup>5</sup>標本の不偏分散の平方根を用いた、標準誤差の推定値を用いた。

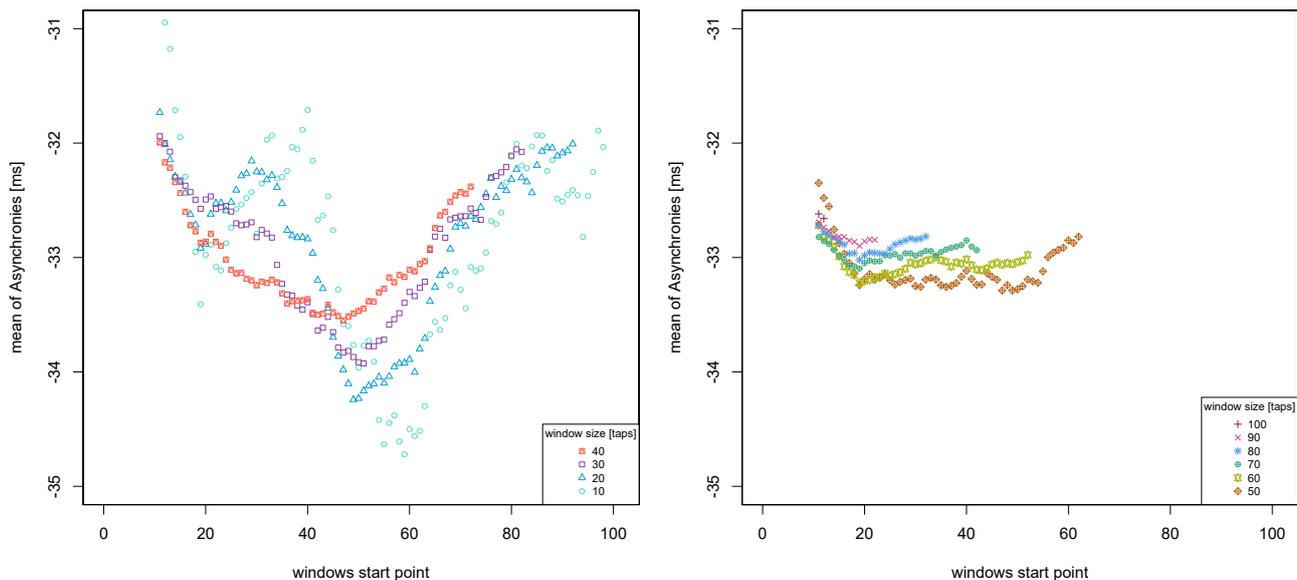
を抽出する場合、どこの位置から抽出するかも問題となる。そこで、算出に用いるタッピングの数と算出に用いるタッピングの開始位置を変化させた。算出に用いるタッピングの数は、10タップから100タップまでの10タップ刻み(10, 20, 30, 40, 50, 60, 70, 80, 90, 100タップ)とした。算出に用いるタッピングの開始位置は、11タップ目からそれぞれのタッピングの数が確保できる位置までの1タップ刻みとした。例えば、算出に用いるタッピングの数を10タップとした場合には、まず、11タップ目から20タップ目までを用いて同調のズレ量の平均値とその標準誤差を算出し、次に、算出に用いるタッピングの開始位置を1タップ後ろにずらし、12タップ目から21タップ目までを用いて算出し、さらに開始位置を1タップ後ろにずらす、と繰り返した。これらは、先に、各参加者の各刺激音 A において算出し、その後、刺激音で平均をとり、さらに参加者で平均をとった。

### 3.1.3 結果

算出に用いるタッピングの数と算出に用いるタッピングの開始位置を変化させた時の、同調のズレ量の平均値を図 3.1.1 に、同調のズレ量の平均値の標準誤差を図 3.1.2 に示した。これらの図では、算出に用いる、それぞれのタッピングの数(シンボル)が、ある開始位置にある時の(横軸)、同調のズレ量の平均値あるいは標準誤差(縦軸)が示された。視認性向上のため、便宜的に、左図にタッピングの数が10–40タップの場合を示し、右図に50–100タップの場合を示した。

図 3.1.1 に示した同調のズレ量の平均値について述べる。まず、算出に用いるタッピングの開始位置が変化するにつれて同調のズレ量の平均値は変動する。その変動量の最大値は、算出に用いるタッピングの数が10タップの場合の、4.141 msであった。次に、算出に用いるタッピングの数が40タップ以下(左図)よりも、窓の大きさが50タップ以上(右図)の方が、算出に用いるタッピングの開始位置が変化するにつれての同調のズレ量の平均値の変動量が小さかった。そして、算出に用いるタッピングの数が50タップ以上における算出に用いるタッピングの開始位置が21タップ目以降の場合では、開始位置が変化するにつれての同調のズレ量の平均値の変動量が僅かであった。なお、これは、タッピングの数が50タップの時の開始位置が61タップ目以降を除いた場合である。最後に、算出に用いるタッピングの数が50タップ以上では、タッピングの数が大きくなっても同調のズレ量の平均値が殆ど変化せず、ほぼ一定の値であった。

続いて、図 3.1.2 に示した、同調のズレ量の平均値の標準誤差について述べる。まず、算出に用いるタッピングの開始位置が変化するにつれて標準誤差は変動する。その変動量の最大値は、算出に用いるタッピングの数が10タップの場合の、0.977 msであった。また、平均値と同様に、タッピングの数が40タップ以下よりも50タップ以上の方が、タッピングの開始位置が変化するにつれての標準誤差の変動量が小さかった。一方、平均値と異なり、標準誤差ではタッピン



(a) 算出に用いるタッピングの数が 10 タップから 40 タップの場合 (b) 算出に用いるタッピングの数が 50 タップから 100 タップの場合

図 3.1.1: 算出に用いるタッピングの数とタッピングの開始位置を変化させた時の同調のズレ量の平均値

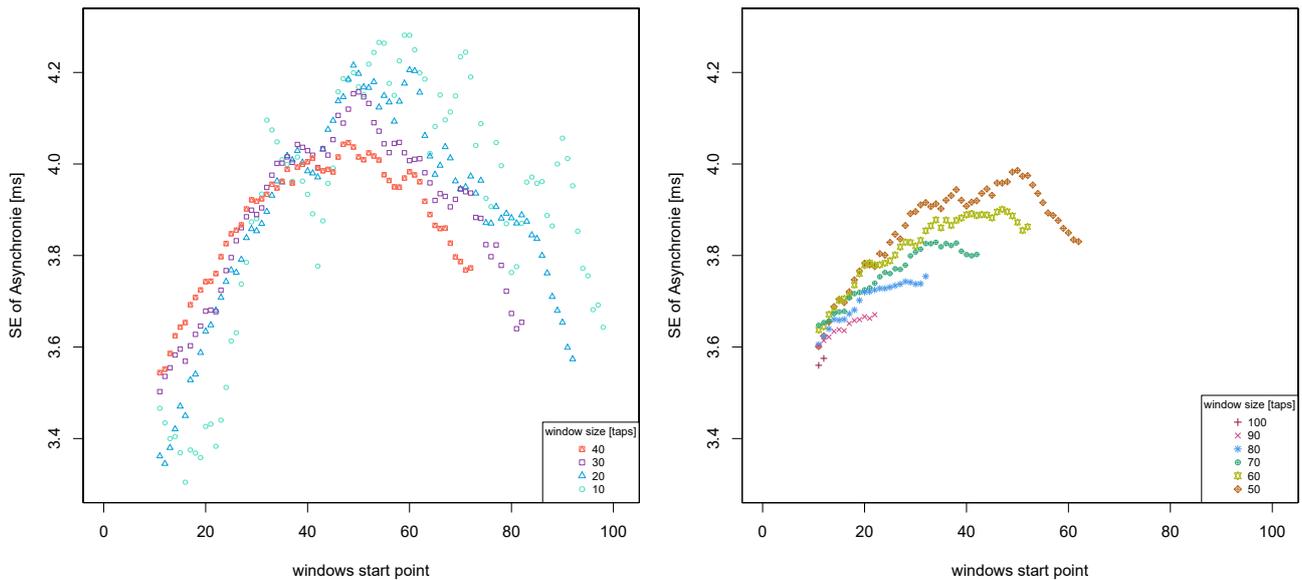
横軸が開始位置 (“windows start point”) であり、縦軸が同調のズレ量の平均値 (“mean of Asynchronies”) である。それらがタッピングの数 (“window size”) ごとに異なるシンボルで示されている。

グの数が 50 タップ以上でも、開始位置が変化するにつれて標準誤差が変動した。だが、タッピングの数が大きくなれば、開始位置が変化するにつれての変動量は小さくなった。最後に、平均値と異なり、標準誤差はタッピングの数が 50 タップ以上でも、タッピングの数が大きくなればなるほど標準誤差が、僅かであるが、小さくなった。

### 3.1.4 考察

以上の分析結果より、「同調のズレ量の平均値ができる限り一定の値に近い値をとり、かつ、同調のズレ量の平均値の標準誤差ができる限り小さい値をとる、その範囲において最も少ないタッピングの数」を検討する。

まず、「同調のズレ量の平均値ができる限り一定の値に近い値をとる」タッピングの数を考える。先述したように、平均値は、理論上、算出に用いるデータ数が増えるにつれ一定の値に近づいていく。確かに、分析結果より、算出に用いるタッピングの数が 50 タップより大きく、算出に用いるタッピングの開始位置が 21 タップ目以降であれば、算出される同調のズレ量の平均値がほぼ一定となることが示された。従って、同調のズレ量の平均値の観点からは、分析にて 21 タップ目から 50 タップ分を用いて平均値を算出できるように、実験で用いるタッピングの数は 70 が望ましい。



(a) 算出に用いるタッピングの数が 10 タップから 40 タップの場合 (b) 算出に用いるタッピングの数が 50 タップから 100 タップの場合

図 3.1.2: 算出に用いるタッピングの数とタッピングの開始位置を変化させた時の同調のズレ量の平均値の標準誤差

横軸が開始位置 (“windows start point”) であり、縦軸が同調のズレ量の平均値の標準誤差 (“SE of Asynchronies”) である。それらがタッピングの数 (“window size”) ごとに異なるシンボルで示されている。

次に、「同調のズレ量の平均値の標準誤差ができる限り小さい値をとる」タッピングの数を考える。先述したように、平均値の標準誤差は、算出に用いるデータ数が増えるにつれ値が小さくなる。確かに、分析結果より、算出に用いるタッピングの数が大きくなればなるほど、標準誤差が小さくなったし、加えて、タッピングの開始位置が変化するにつれての標準誤差の変動量も小さくなることが示された。従って、標準誤差の観点からは、算出に用いるタッピングの数ができるだけ大きい方が望ましいといえる。

最後に、平均値と標準誤差より導かれた望ましいタッピングの数の「範囲において最も少ないタッピングの数」を検討する。標準誤差の観点からは算出に用いるタッピングの数ができるだけ大きい方が望ましかった。そのため、平均値より導かれたタッピングの数の一段階大きいタッピングの数を採用することが望ましいと考える。すなわち、実験において 80 回のタッピングを行い、分析において 21 タップ目から 80 タップ目までの 60 タップを用いて同調のズレ量の平均値を算出することが望ましいと結論づける。

従って、以下のタッピング課題は、このように行った。

## 3.2 予備実験2: 実験課題の比較検討

2.3.4項で述べたように、PAT測定課題には、タッピング課題、同時調整課題、等時調整課題の3つがあるが、これらの3つの課題より得られるPATに有意差がないという確証は未だない。先行研究より明らかになっていることは、まず、2つの課題間より得られるPATの比較を行った研究より、タッピング課題と等時調整課題とで、また、同時調整課題と等時調整課題とで得られる値に有意差がないことである。このことより、そうであれば、タッピング課題と同時調整課題とで得られる値に有意差がない可能性が考えられる。次に、唯一3つの課題を比較した先行研究では、タッピング課題と2つの調整課題(同時調整課題と等時調整課題)の間で得られる値に有意差があったことである。だが、この研究では、タッピング課題と他の2つの調整課題とを異なる指標で比較していた。以上を鑑みると、3つの課題を同一の指標で比較すれば、3つの課題間に有意差がない可能性が考えられた。

そこで、この予備実験2にて、タッピング課題、同時調整課題、等時調整課題の3つを、同一の指標で比較した。このタッピング課題での音列内のタッピングの数は、予備実験1で求めた数とした。先行研究より提示音圧レベルがPATを変化させる可能性が示唆されているため、合わせて、提示音圧レベルを変化させた場合にPATが変化するかどうかを確認した。

### 3.2.1 実験方法

#### 3.2.1.1 参加者

3人であった。内1人は実験者であった。彼らはみな、東京藝術大学大学院音楽音響創造研究分野に所属しており、音楽訓練を受けたことがあった。

#### 3.2.1.2 刺激音

440 Hzのサイン波を用いた。

持続時間と立ち上がり時間<sup>6</sup>を変化させた。持続時間5水準(50, 75, 100, 125, 300 ms)×立ち上がり時間8水準(0, 10, 20, 30, 40, 80, 120, 240 ms)の計40の組み合わせより、持続時間よりも立ち上がり時間が短い13の組み合わせを刺激音とした。さらに、クリック音の刺激音(持続時間5 ms、立ち上がり時間5 ms)が追加された。従って、計14個の刺激音が作成された。

立ち上がり部分と立ち下がり部分は、振幅エンベロープにおいて線形に変化させた。そのため、立ち上がり時間が0 msの刺激音は、振幅の急峻かつ大幅な変動によって、目的の周波数以外をも含むノイズが僅かに発生した。その影響を軽減するため、先述した立ち上がり時間の水準が0 msの刺激音のみ、実際には1 msの立ち上がり時間をもたせた。同様の理由で、立ち下

<sup>6</sup>この予備実験2でも、「立ち上がり時間」を変化させた。

「立ち上がり時間の割合(持続時間に占める立ち上がり時間の割合であり、数式では(立ち上がり時間)/(持続時間)×100[%]である)」ではなかった。

がり時間も 1 ms で固定した。このように、実際には 1 ms の立ち上がり時間と立ち下がり時間であるが、以下では簡略的に 0 ms と記す。

以上の刺激音は、PC 上で Python を用いて生成され、標本周波数 48000 Hz、量子化ビット数 24 bit のモノラル音源として書き出された。

### 3.2.1.3 実験手続き

実験課題は、タッピング課題、等時調整課題、同時調整課題の 3 つであった。また、各課題で刺激音の提示音圧レベルを 2 水準 (76 dB、84 dB) で変化させた。この提示音圧レベルは、聴感上明らかに差があるが、長時間聴いていても耳に害がないレベルとした。

それぞれの課題の実験手続きは以下の通りである。

#### タッピング課題

タッピング課題の実験手続きは、予備実験 1 のタッピング課題の実験手続き (3.1.1.3 項) とほぼ同様であった。重複するため、詳細は割愛する。

予備実験 1 と異なる点は、提示音圧レベルとタッピングの数、そして本番の試行の構成であった。まず、提示音圧レベルは先述した通りである。次に、タッピングの数は水準 (80 回、110 回) とした。すなわち、同一刺激音を 700 ms ごとに 110 回あるいは 80 回繰り返した音列を用いた。本節の予備実験においてタッピングの数を 2 水準用いたのは、前節でタッピング課題では実験にて同一刺激が 80 回繰り返される音列を用いることが望ましいと結論づけたことより、80 回のタッピングと 110 回のタッピングとから算出したに統計的有意差がないかを確認するためであった。従って、全部で 56 音列 (14 個の刺激音 × 2 水準の提示音圧レベル × 2 つの音列内のタッピングの数) が用意された。そして、これらの全 56 音列を用いた実験全体は、14 つの音列を 3 試行に分割したものを 1 ブロックとし、計 4 ブロックで構成された。各ブロックは、タッピングの数 (2 水準) と提示音圧レベル (2 水準) のうち一つの水準のみをまとめられたものであった。ブロックの提示順は参加者間でカウンターバランスがとられた。1 ブロック内の 14 音列は 3 試行 (5 音列 + 4 音列 + 4 音列) に分けられた。クリック音以外の 13 音列がランダムに 5 音列あるいは 4 音列を連結されたものを 1 試行とした。3 試行目の最後にクリック音が繰り返される音列が提示された。各試行間に 5 分間の休憩が挟まれた。本番の所要時間は 1 時間半であった。

#### 同時調整課題と等時調整課題

同時調整課題と等時調整課題の実験手続きは殆ど同じである。そのため、まとめて述べる。

両課題のベース音は、クリック音 (持続時間 5 ms、立ち上がり時間 5 ms) と、クリック音を除き持続時間と立ち上がり時間が最も短い刺激音 (持続時間 50 ms、立ち上がり時間 0 ms) の 2 つ

であった。テスト音は、クリック音を除く、13個の刺激音であった。従って、両課題で調整させたペアの数は、計52ペア(13個のテスト音×2個のベース音×2水準の提示音圧レベル)であった。

これらの52ペアを用いた実験全体は、13つのペア音列を2試行に分割したものを1ブロックとし、計4ブロックで構成された。各ブロックは、タッピングの数(2水準)と提示音圧レベル(2水準)のうち一つの水準のみをまとめられたものであった。ブロックの提示順は参加者間でカウンターバランスがとられた。1ブロック内の13ペアはランダムに2試行(7ペア+6ペア)に分けられ、提示された。各試行間に5分の休憩が挟まれた。本番の所要時間は2時間であった。

参加者が調整する前の各ペアのベース音とテスト音は、聴感上明らかに同時あるいは等時ではない位置に配置されるよう、それらの物理的オンセットの差分が±0–100msの範囲内に配置された。その際、その正負はカウンターバランスがとられ、その値はランダム化された。

両課題のベース音やテスト音の時間間隔について述べる。同時調整課題では、ベース音とベース音、及び、テスト音とテスト音の物理的オンセットの時間間隔が700ms(約86BPM)であった。等時調整課題では、ベース音とベース音、及び、テスト音とテスト音の物理的オンセットの時間間隔が1400msであった。すなわち、調整後のベース音とテスト音の時間間隔が理論的には700ms(約86BPM)となる時間間隔であった。

ベース音とテスト音はループ提示された。参加者の好きな時に一時停止や頭出し再生が可能であった。等時調整課題ではダイオティックで、同時調整課題ではダイコティックで両耳に提示された。視覚的情報は与えられなかった。

参加者が各ペアを調整し終えたら、実験者によって手動で値が記録された。記録された値は、同時調整課題では、この課題における $\Delta PC_{\text{test,base}}$ の算出方法より、調整後のベース音とテスト音の物理的オンセットの差分であった。簡略化した式2.8を再掲する。

$$\Delta PC_{\text{test,base}} = t_{\text{onset}}(\text{base}) - t_{\text{onset}}(\text{test})$$

一方、等時調整課題では、調整後のベース音とテスト音の物理的オンセットの差分に、理論的なベース音とテスト音の時間間隔(本実験の場合700ms)を加えた値であった。簡略化した式2.9を再掲する。

$$\Delta PC_{\text{test,base}} = t_{\text{onset}}(\text{base}) - t_{\text{onset}}(\text{test}) + \text{idealIOI}$$

#### 3.2.1.4 実験装置

それぞれの課題の実験装置は以下の通りであった。

#### タッピング課題

予備実験1の実験装置(3.1.1.4項)と同様であった。重複するため、詳細は割愛する。

## 同時調整課題と等時調整課題

刺激音の提示方法は、タッピング課題と同様であった。

参加者のテスト音の位置調整には、MIDI コントローラ (KORG nanoPAD2) のボタンを用いた。MIDI コントローラは PC に繋がられ、そのボタンのうち、2つのボタンがテスト音を 10 ms 分時間的に前後に動かすように設定され、他の 2つのボタンがテスト音を 1 ms 分時間的に前後に動かすように設定された<sup>7</sup>。ボタンを押したときに音が何 ms 動くかは参加者には伝えなかった。

### 3.2.2 分析と結果

分析にて、3つの課題より得られる PAT を同一指標で比較し、有意に異なるかどうかを確認した。また、タッピング課題より算出される同調のズレ量の平均値が 2水準のタッピングの数によって有意に異なるかどうかと、3つの課題より得られる PAT が 2水準の提示音圧レベルによって有意に異なるかをも確認した。

#### 前処理

分析の前に、まず、「立ち上がり時間」を「立ち上がり時間の割合」に変換した。すなわち、(立ち上がり時間)/(持続時間)×100 [%] を算出した。全 14 個の刺激音の持続時間と「立ち上がり時間」と「立ち上がり時間の割合」を表 3.2.1 に示した。

3つの課題を比較する指標として、 $\Delta PC_{A,R}$  を算出した。この予備実験 2 では、同時調整課題と等時調整課題で用いたベース音の 2 音を基準の刺激音  $R$  とした。以後、クリック音 (持続時間 5 ms、立ち上がり時間 5 ms) を “D5R5 音” と称し、クリック音を除いた最も持続時間と立ち上がり時間が短い音 (持続時間 50 ms、立ち上がり時間 0 ms) を “D50R0 音” と称す。従って、3つの課題より、 $\Delta PC_{A,D5R5}$  と、 $\Delta PC_{A,D50R0}$  とが算出された。また、分析の都合上、基準の刺激音と基準の刺激音との  $\Delta PC_{R,R} = 0$  もある方が望ましいため、これも算出した。

より具体的に、課題ごとに算出したものを述べる。タッピング課題においては、まず、タッピングの数 2水準と提示音圧レベル 2水準とにおける全 56 刺激音に対する同調のズレ量の平均値が算出された。タッピングの数が 80 回の音列では 21 タップ目から 80 タップ目までを、110 回の音列では 11 タップ目から 110 タップ目までを用いて算出された。その後、基準の刺激音以外の 13 個の刺激音との  $\Delta PC_{A,D5R5}$  および  $\Delta PC_{A,D50R0}$  が算出され、加えて、基準の刺激音同士の  $\Delta PC_{D5R5,D5R5} = 0$  および  $\Delta PC_{D50R0,D50R0} = 0$  も算出された。一方、同時調整課題と等時調整課題では、D5R5 音が基準の刺激音の場合には、提示音圧レベル 2水準における、全 13 個のテスト音との  $\Delta PC_{A,D5R5}$  が算出され、加えて、基準の刺激音同士の  $\Delta PC_{D5R5,D5R5} = 0$  も算

<sup>7</sup>REAPER の “Nudge right/left” を割り振った。

表 3.2.1: 刺激音の持続時間と立ち上がり時間 (立ち上がり時間の割合)

	持続時間 [ms]			
	5	50	100	300
立ち上がり時間 [ms] (立ち上がり時間の割合 [%])		0 (0%)	0 (0%)	0 (0%)
	5 (100%)			
		20 (40%)	20 (20%)	20 (7%)
		40 (80%)	40 (40%)	40 (13%)
			80 (80%)	80 (27%)
				120 (40%)
				240 (80%)

出された。そして、D50R0 音が基準の刺激音の場合には、提示音圧レベル 2 水準における、基準の刺激音以外の 12 個のテスト音との  $\Delta PC_{A,D50R0}$  が算出され、加えて、基準の刺激音同士の  $\Delta PC_{D50R0,D50R0} = 0$  も算出された。

### 3.2.2.1 タッピング課題におけるタッピングの数が異なる時の同調のズレ量の平均値

タッピング課題において、算出される同調のズレ量の平均値がタッピングの数の 2 水準 (80 回、110 回) によって有意に異なるかどうかを確認した。

まず、算出された同調のズレ量の平均値をタッピングの数の 2 水準 (80 回、110 回) ごとに箱ひげ図に示した<sup>8</sup>。それが図 3.2.1 である。この箱ひげ図の見方は図のキャプションを参照してほしい。この箱ひげ図より、タッピングの数に関わらず、中央値も四分位範囲も、外れ値の数もほぼ同じであることが明らかである。

次に、同調のズレ量の平均値を従属変数とし、タッピングの数を独立変数として、一要因参加者内計画の分散分析 (Greenhouse-Geisser の  $p$  値補正付き) を行った。その結果、タッピングの数の主効果は 5% 水準で有意でなかった ( $F(1, 2) = 1.001, p = .423, \eta_{partial}^2 = 0.334$ )。

従って、タッピング課題において、タッピングの数を 110 回から 80 回に削減しても同調のズレ量の平均値に有意差がなかったことが確認された。

<sup>8</sup>それぞれの箱にあるデータは、提示音圧レベル 2 水準 × 基準の刺激音を含めた全 14 個の刺激音の、計 28 個の同調のズレ量の平均値である。

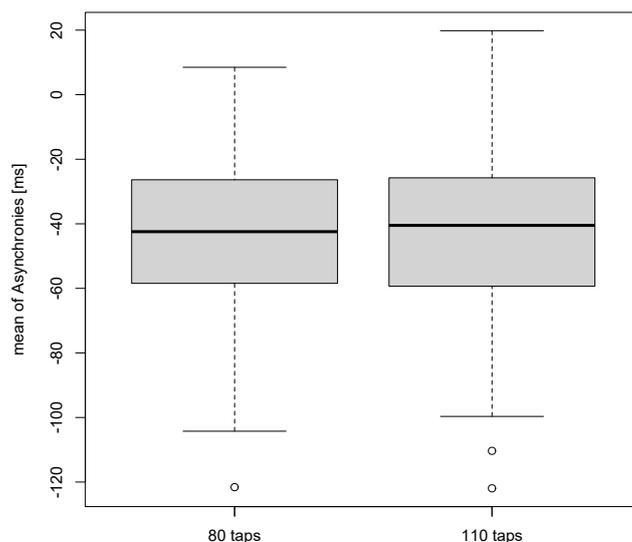


図 3.2.1: タッピングの数ごとの同調のズレ量の平均値の箱ひげ図

横軸がタッピングの数で 80 回 (“80 taps”) の場合と 110 回 (“110 taps”) の場合であり、縦軸が同調のズレ量の平均値 (“mean of Asynchronies”) である。

この箱ひげ図では、中央値が箱の中の太線で表され、四分位範囲 (第 1 四分位から第 3 四分位の差分) すなわちデータのばらつきが箱の大きさで表される。また、外れ値は「第 1 四分位・第 3 四分位から四分位範囲の 1.5 倍以上離れた値」と定義され、白い丸のシンボルで表される。そして、外れ値を除いた最小値と最大値がひげの端で表される。

以下ではタッピングの数 2 水準の結果をまとめて分析を行った。

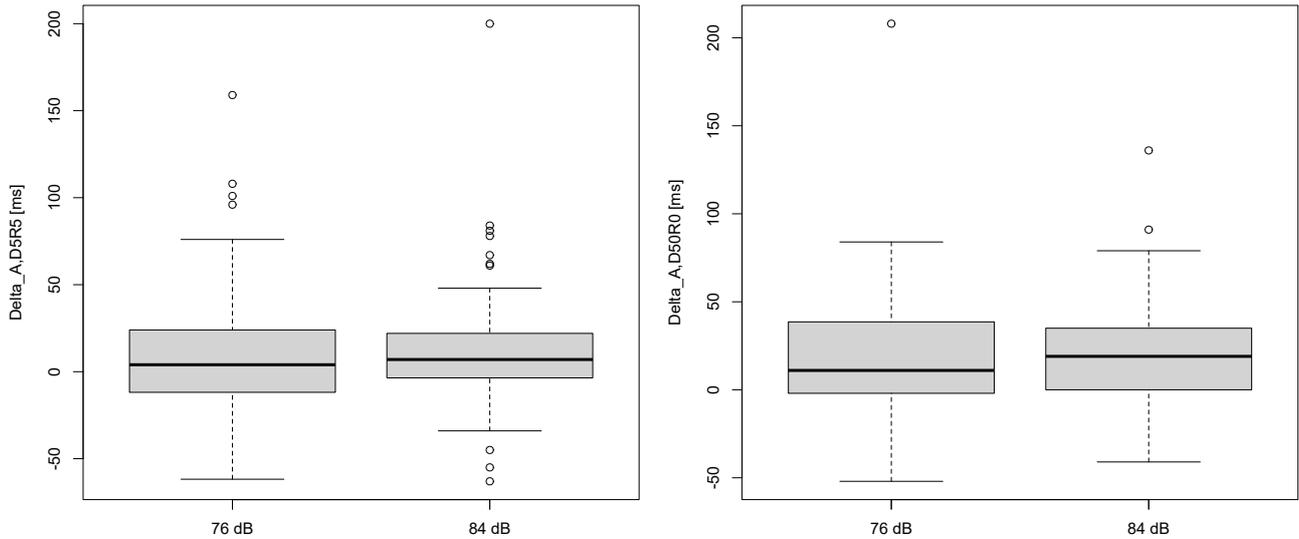
### 3.2.2.2 3つの課題における提示音圧レベルが異なる時の $\Delta PC_{A,R}$

3つの課題において、得られる  $\Delta PC_{A,R}$  が提示音圧レベル 2 水準 (76 dB、84 dB) によって有意に異なるかどうかを確認した。

まず、3つの課題より算出された  $\Delta PC_{A,D5R5}$  と  $\Delta PC_{A,D50R0}$  とを、提示音圧レベルごとに箱ひげ図で示した<sup>9</sup>。  $\Delta PC_{A,D5R5}$  の場合が図 3.2.2a であり、  $\Delta PC_{A,D50R0}$  の場合が図 3.2.2b である。これらの図より、  $\Delta PC_{A,D5R5}$  でも、  $\Delta PC_{A,D50R0}$  でも、提示音圧レベルが異なると、外れ値の数とひげの長さが異なるが、中央値は大きく異ならなかった。

次に、  $\Delta PC_{A,D5R5}$  と  $\Delta PC_{A,D50R0}$  を従属変数とし、提示音圧レベルを独立変数として、一要因参加者内計画の分散分析 (Greenhouse-Geisser の  $p$  値補正付き) を行った。その結果、  $\Delta PC_{A,D5R5}$  を従属変数とした分散分析でも、  $\Delta PC_{A,D50R0}$  を従属変数とした分散分析でも、提示音圧レベルの主効果は 5% 水準で有意でなかった (順に  $F(1, 2) = 0.182, p = .711, \eta_{partial}^2 = 0.083, F(1, 2) = 0.406, p = .589, \eta_{partial}^2 = 0.169$ )。

<sup>9</sup>それぞれの箱にあるデータは、  $\Delta PC_{A,D5R5}$  の場合は 56 個 (タッピング課題で 28 個+同時調整課題で 14 個+等時調整課題で 14 個) であり、  $\Delta PC_{A,D50R0}$  の場合は 54 個 (28+13+13) である。



(a)  $\Delta PC_{A,D5R5}$  (基準の刺激音が D5R5 音) の場合

(b)  $\Delta PC_{A,D50R0}$  (基準の刺激音が D50R0 音) の場合

図 3.2.2: 提示音圧レベルごとの  $\Delta PC_{A,R}$  の箱ひげ図

横軸が提示音圧レベル 2 水準 (76 dB, 84 dB) であり、縦軸が  $\Delta PC_{A,R}$  である。

この箱ひげ図の見方は図 3.2.1 と同様である。

最後に、先行研究との比較のため、持続時間、立ち上がり時間の割合、提示音圧レベルを独立変数とした、三要因の分散分析を行いたい。だが、二要因以上の分散分析では、独立変数のある要因における水準数と異なる要因における水準数との全部の組み合わせの刺激音に対する従属変数のデータが必要である。例えば、持続時間 (要因) が 2 水準 (50, 300 ms) で、立ち上がり時間の割合 (要因) が 2 水準 (0, 100%) あったならば、2 水準  $\times$  2 水準の計 4 個の組み合わせの刺激音に対する従属変数のデータを用いる必要がある。この予備実験 2 では、持続時間における水準と立ち上がり時間の割合における水準との全部の組み合わせの刺激音を用いていない (刺激音の持続時間と立ち上がり時間の割合をまとめた表 3.2.1 参照)。そこで、データを絞り、基準の刺激音が D5R5 音である  $\Delta PC_{A,D5R5}$  のみを従属変数とし、持続時間 3 水準 (50, 100, 300 ms)、立ち上がり時間の割合 3 水準 (0, 40, 80%)、提示音圧レベル 2 水準 (76 dB, 84 dB) を独立変数として、三要因参加者内計画の分散分析 (Greenhouse-Geisser の  $p$  値補正付き) を行った。その結果、立ち上がり時間の割合の主効果 ( $F(2, 4) = 13.969, p = .016, \eta_{partial}^2 = 0.875$ ) のみ、5%水準で有意であった。一方、提示音圧レベルの主効果と提示音圧レベル、その他の要因との交互作用は、5%水準で有意でなかった。

従って、この予備実験 2 で用いた提示音圧レベル 2 水準 (76 dB, 84 dB) では、提示音圧レベルが異なっても PAT は変化しないことが明らかになった。

以下では提示音圧レベル 2 水準もまとめて分析を行った。

### 3.2.2.3 3つの課題より算出される $\Delta PC_{A,R}$

3つの課題より算出される  $\Delta PC_{A,R}$  が有意に異なるのか、また有意ではなくとも傾向として差異があるのかを確認した。

この際、まず、持続時間と立ち上がり時間の割合を考慮せずに分析を行い、次に、持続時間と立ち上がり時間の割合を考慮して分析を行った。

まず、持続時間と立ち上がり時間の割合を考慮せず、全刺激音の  $\Delta PC_{A,D5R5}$  と  $\Delta PC_{A,D50R0}$  が、3つの課題によって異なるかを確認した<sup>10</sup>。

先に、 $\Delta PC_{A,D5R5}$  と  $\Delta PC_{A,D50R0}$  を従属変数とし、PAT 測定課題の種類を独立変数として、一要因参加者内計画の分散分析 (Greenhouse-Geisser の  $p$  値補正付き) を行った。その結果、 $\Delta PC_{A,D5R5}$  を従属変数とした分散分析でも、 $\Delta PC_{A,D50R0}$  を従属変数とした分散分析でも、PAT 測定課題の種類の主効果が 5%水準で有意であった (順に  $F(2, 4) = 15.861, p = .013, \eta_{\text{partial}}^2 = 0.888$ ,  $F(2, 4) = 8.681, p = .035, \eta_{\text{partial}}^2 = 0.813$ )。だが、下位検定として Shaffer の多重比較を行ったところ、5%水準で調整後  $p$  値が有意な組み合わせはなかった<sup>11</sup>。

続いて、課題ごとに箱ひげ図に示し、傾向として差異があるかを確認した。 $\Delta PC_{A,D5R5}$  の場合が図 3.2.3a であり、 $\Delta PC_{A,D50R0}$  の場合が図 3.2.3b である。これらの図より、以下の五つが明らかであった。一つ目に、 $\Delta PC_{A,D5R5}$  でも  $\Delta PC_{A,D50R0}$  でも、課題ごとに中央値が異なった。中央値が大きい順に課題を並べると、等時調整課題、同時調整課題、タッピング課題であった。二つ目に、 $\Delta PC_{A,D5R5}$  でも  $\Delta PC_{A,D50R0}$  でも、課題ごとに四分位範囲の大きさが異なった。四分位範囲が大きい順に並べると、等時調整課題、同時調整課題、タッピング課題であった。三つ目に、 $\Delta PC_{A,D5R5}$  よりも、 $\Delta PC_{A,D50R0}$  の方が四分位範囲が大きかった。これは同時調整課題と等時調整課題において顕著であった。四つ目に、 $\Delta PC_{A,D5R5}$  と  $\Delta PC_{A,D50R0}$  とを合わせて考えると、課題ごとに外れ値の数が異なった。同時調整課題と等時調整課題とでは外れ値の数が多く、タッピング課題では外れ値が少なかった。五つ目に、 $\Delta PC_{A,D50R0}$  よりも、 $\Delta PC_{A,D5R5}$  の方が外れ値の数が多かった。

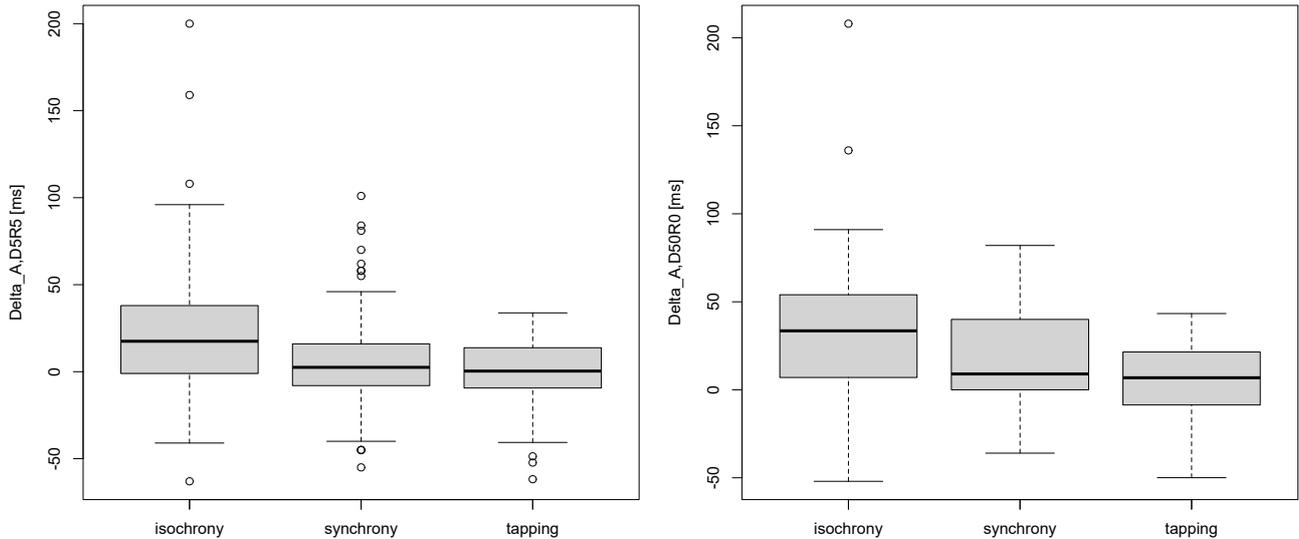
次に、持続時間と立ち上がり時間の割合を考慮し、持続時間ごとの立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D5R5}$  や  $\Delta PC_{A,D50R0}$  の変化が3つの課題で傾向としてどのように異なるかを確認した。

<sup>10</sup>分析に用いたデータは、 $\Delta PC_{A,D5R5}$  の場合、タッピング課題で 56 個、同時調整課題と等時調整課題で 28 個であった。一方、 $\Delta PC_{A,D50R0}$  場合、タッピング課題で 56 個、同時調整課題と等時調整課題で 26 個であった。

<sup>11</sup>多重比較の結果を述べる。

$\Delta PC_{A,D5R5}$  を従属変数とした場合、タッピング課題と同時調整課題の組み合わせが  $t(2) = 1.660, \text{adj.}p = 0.239$ 、タッピング課題と等時調整課題の組み合わせが  $t(2) = 3.906, \text{adj.}p = 0.179$ 、同時調整課題と等時調整課題の組み合わせが  $t(2) = 2.434, \text{adj.}p = 0.179$  であった。

$\Delta PC_{A,D50R0}$  を従属変数とした場合、タッピング課題と同時調整課題の組み合わせが  $t(2) = 3.813, \text{adj.}p = 0.115$ 、タッピング課題と等時調整課題の組み合わせが  $t(2) = 4.965, \text{adj.}p = 0.115$ 、同時調整課題と等時調整課題の組み合わせが  $t(2) = 2.657, \text{adj.}p = 0.117$  であった。



(a)  $\Delta PC_{A,D5R5}$  (基準の刺激音が D5R5 音) の場合

(b)  $\Delta PC_{A,D50R0}$  (基準の刺激音が D50R0 音) の場合

図 3.2.3: 3つの課題ごとの  $\Delta PC_{A,R}$  の箱ひげ図

横軸が3つの課題で、左から、等時調整課題 (“isochrony”)、同時調整課題 (“synchrony”)、タッピング課題 (“tapping”) である。縦軸が  $\Delta PC_{A,R}$  である。

この箱ひげ図の見方は図 3.2.1 と同様である。

先に、基準の刺激音ごとに、3つの課題ごと (図ごと) の、持続時間ごと (シンボル) の立ち上がり時間の割合の増加 (横軸) に対する  $\Delta PC_{A,D5R5}$  (縦軸) を図示した。  $\Delta PC_{A,D5R5}$  の場合が図 3.2.4 であり、  $\Delta PC_{A,D50R0}$  の場合が図 3.2.5 である。この図より、以下の二つのことが明らかである。一つ目に、  $\Delta PC_{A,D5R5}$  でも、  $\Delta PC_{A,D50R0}$  でも、殆どの持続時間で、タッピング課題よりも同時調整課題と等時調整課題の方が、立ち上がり時間の割合の増加に対する  $\Delta PC_{A,R}$  の変化量が大きかった。二つ目に、同時調整課題と等時調整課題では、  $\Delta PC_{A,D5R5}$  の場合は持続時間によらず、立ち上がり時間の割合が増加すると  $\Delta PC_{A,D5R5}$  も増加したが、  $\Delta PC_{A,D50R0}$  の場合は持続時間によって、立ち上がり時間の割合の増加に対して  $\Delta PC_{A,D50R0}$  が一定だったり増加したりした。これはタッピング課題では見られなかった。

以上のことより、課題ごとに得られる  $\Delta PC_{A,R}$  は、詳細にみれば異なる傾向があるが、統計的には有意差がないことが明らかになった。

### 3.2.3 考察

以上の分析結果をもとに、考察を行う。

まず、タッピング課題において、同一刺激音へのタッピングの数を 110 回から 80 回に削減しても、同調のズレ量の平均値に有意差がなかったことが確認された。このことより、前節の予

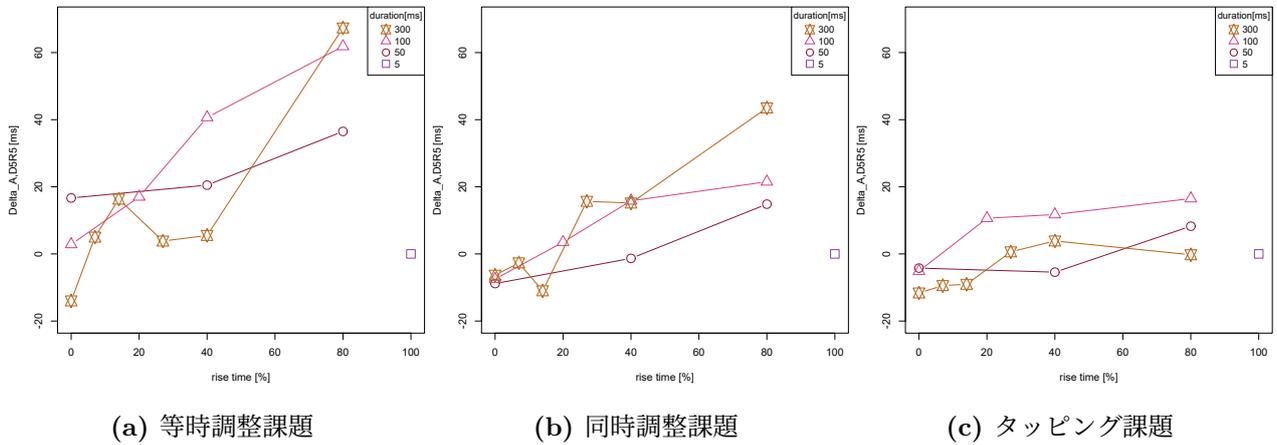


図 3.2.4: 課題ごとの、持続時間ごとの立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D5R5}$  (基準の刺激音が D5R5 音) の変化  
 3つの課題ごと (図ごと) の、持続時間ごと (シンボル) の立ち上がり時間の割合の増加 (横軸) に対する  $\Delta PC_{A,D5R5}$  (縦軸) が示された。

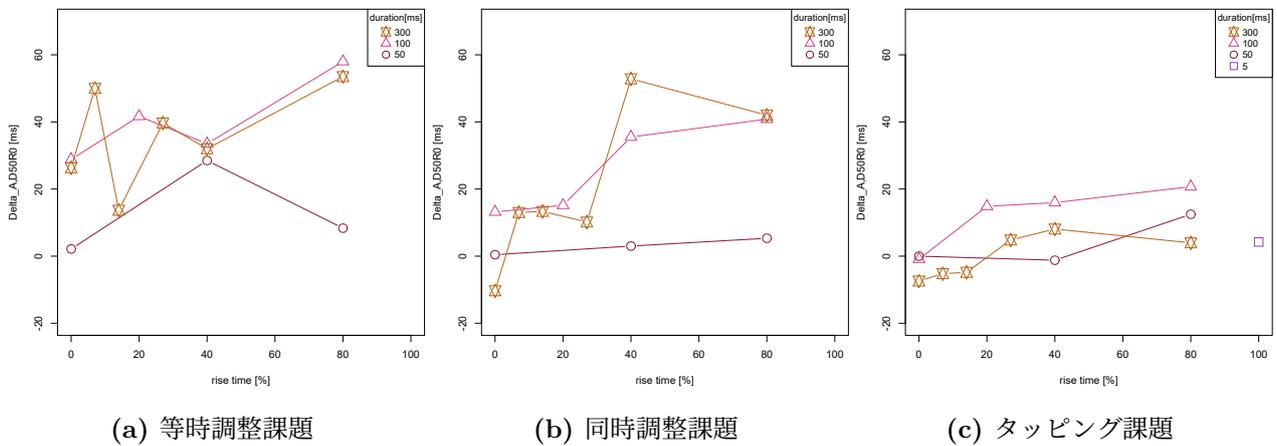


図 3.2.5: 課題ごとの、持続時間ごとの立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D50R0}$  (基準の刺激音が D50R0 音) の変化  
 3つの課題ごと (図ごと) の、持続時間ごと (シンボル) の立ち上がり時間の割合の増加 (横軸) に対する  $\Delta PC_{A,D50R0}$  (縦軸) が示された。

備実験1において、タッピング課題におけるA音のタッピングの数と分析に用いるデータ範囲として、「同一刺激音を一定間隔ごとに80回繰り返す音列を提示し、タッピングを行ってもらおう。そして分析にて、その音列の21タップ目から80タップ目までの60タップを用いて同調のズレ量の平均値を算出する」と結論付けたことが妥当であると確かめられた。

次に、3つの課題において、この予備実験2で用いた提示音圧レベル(76 dB、84 dB)では、提示音圧レベルが異なってもPATは変化しないことが明らかになった。また、三要因の分散分析を行った結果、提示音圧レベルと、持続時間や立ち上がり時間の割合との交互作用も有意ではなかったことが明らかになった。

2.4.4節にて述べたように、PATの先行研究であるJ.Vosら[34]とBechtoldら[1]は、提示音圧レベルとその他の音の要素との交互作用があると報告した。従って、この予備実験2は、先行研究と矛盾した結果であった。この矛盾を引き起こした原因として、実験に用いた提示音圧レベルとその差分が考えられる。J.Vosらが用いた提示音圧レベルは37, 57, 77 dBであり、その差分が20 dBと40 dBであった。一方、Bechtoldらが用いたのはRMSした音圧レベルが刺激音間で-17.1- -36.9 dBの範囲であり、その差分が19.8 dBであった。一方、本論の予備実験2では76 dBと84 dBであり、その差分が8 dBであった。従って、提示音圧レベルがその他の音の要素と合わさってPATを変化させるには、ある程度の音圧レベルの差分が必要である可能性が考えられた。だが、2.6章で述べたように、提示音圧レベルはPATを直接的に移動させるのではなく、統合の時間窓の長さを変化させ、その影響でPATの知覚が間接的に変化した可能性も考えられる。本論では提示音圧レベルについてはこれ以上は扱わないが、提示音圧レベルがPATの知覚をどのように変化させるのかについては、今後の研究が必要である。

最後に、タッピング課題、同時調整課題、等時調整課題より算出される $\Delta PC_{A,R}$ は、詳細にみれば異なる傾向があるが、統計的には有意差がないことが明らかになった。

この予備実験2の冒頭にて、先行研究の報告より推察した結果、3つの課題を同一の指標で比較すれば、3つの課題間に有意差がない可能性を述べた。分析結果より、この可能性が支持された。

また、箱ひげ図(図3.2.3)や、 $\Delta PC_{A,R}$ の変化の図(図3.2.4と図3.2.5)より、詳細にみれば傾向として差異があることが示されたが、これらの内容は2.3章で述べた先行研究の報告と一致していた。一つ目に、先行研究では、課題間で得られる標準偏差が異なる傾向があると報告された。本論では標準偏差を算出していないが、箱ひげ図より、データのばらつきを表す四分位範囲の大きさと外れ値の数が課題によって異なることが確認された。二つ目に、先行研究では、立ち上がり時間が長い音のPATが課題間で異なる傾向があると報告され、また、持続時間が増加した時のPATの変化量と、立ち上がり時間が増加した時のPATの変化量とのどちらの方が大きいかが課題間で異なる傾向があると報告された。本論でも、 $\Delta PC_{A,R}$ の変化の図よ

り、基準の刺激音によらず、課題ごとに、立ち上がり時間の割合の増加に対する  $\Delta PC_{A,R}$  の変化量が異なることが確認された。三つ目に、先行研究によって、同時調整課題の問題点として、ベース音によって調整の難易度と得られる  $\Delta PC_{\text{test,base}}$  の標準偏差が異なることが報告された。本論では標準偏差を算出していないが、箱ひげ図より、基準の刺激音によって、四分位範囲の大きさと外れ値の数が異なることも確認された。さらに、 $\Delta PC_{A,R}$  の変化の図より、基準の刺激音によって、同時調整課題と等時調整課題では、持続時間ごとの「立ち上がり時間の増加と  $\Delta PC_{A,R}$  の変化との関係」も異なることも確認された。すなわち、本論では、同時調整課題に加え、等時調整課題でも、ベース音による差異が見られた。だが、これらの差異はタッピング課題では見られなかった。

### 3.2.4 本論で用いる PAT 測定課題の検討

本論では、多くの刺激音を用いることになるため、刺激音の数に対する実験時間が短い方が望ましく、また、音楽訓練を受けたことがない参加者に対しても実験を行うため、実験課題はより単純でより労力が少ない方が望ましい。そのため、タッピング課題の2つの問題が解消されたならば、本実験の PAT 測定課題にタッピング課題を用いるのが望ましいと考えた。実際、本章の予備実験1と予備実験2において問題が解決されたため、タッピング課題を用いることに問題がないことが確認された。

また、予備実験2より、タッピング課題では刺激音の数に対する実験時間が短いことが確認された。予備実験2の実験所要時間は、どの課題でも、1時間半–2時間であった。だが、予備実験2では同時調整課題と等時調整課題において各ペアを1度ずつ調整させたが、本来、これらの課題では複数回ずつ調整させ、複数回の調整後の値を平均して PAT を算出することが望ましい。もし、同時調整課題と等時調整課題で各ペアを4度ずつ調整を行うならば、タッピング課題と比べ、4倍の時間がかかることになる。

さらに、同時調整課題と等時調整課題では、ベース音によって難易度が変化したり、調整中の意識の向け方や調整中のベース音とテスト音の時間間隔によって調整の遂行が難しいことが想定された。実際、この予備実験2の参加者より、「タッピング課題よりも同時調整課題と等時調整課題は難しい」とのコメントがあった。

加えて、同時調整課題と等時調整課題では、ベース音によって ( $\Delta PC_{A,R}$  における基準の刺激音  $R$  によって)、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,R}$  の変化との関係」が異なってしまうことが明らかになった。これは、本論において、「立ち上がり時間の割合の増加と PAT の変化との関係」が持続時間によって異なるかを調べ、ひいては本論の仮説の検証を行うことに、大きな問題となることが想定される。

以上のことより、本論では、やはりタッピング課題を用いることが望ましく、また妥当であ

ると結論付けた。

### 3.3 予備実験3: 仮説検証方法の確認並びに本実験条件の検討

本節では、2.6節で述べた仮説検証の実験方法と分析方法によって仮説の検証が可能であるかを確認した。その後、本実験において、刺激音の持続時間と立ち上がり時間の割合をどのような水準で変化させるのが適切かを検討した。合わせて、適切なサンプルサイズも検討した。

#### 3.3.1 実験方法

##### 3.3.1.1 参加者

3人であった。内1人は実験者であった。彼らははみな、東京藝術大学音楽学部音楽環境創造科、並びに同大学院音楽音響創造研究分野に所属していた。彼らはみな、音楽訓練を受けたことがあった。

##### 3.3.1.2 刺激音

440 Hz のサイン波を用いた。

持続時間と立ち上がり時間の割合を変化させた。持続時間は8水準(50, 100, 150, 200, 250, 300, 350, 400 ms)とし、立ち上がり時間の割合は10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)とした<sup>12</sup>。持続時間50 msで立ち上がり時間の割合90%の刺激音では、立ち下がり時間10 msの確保が不可能となるため、これは削除された。従って、持続時間8水準×立ち上がり時間の割合10水準より1個の刺激音を除いた、計79個の刺激音が作られた。

立ち上がり部分は振幅エンベロープにおいて線形に変化させたが、立ち下がり部分はコサインカーブ状に変化させた。立ち下がり時間は10 msで固定した。

以上の刺激音は、PC上でPythonを用いて生成され、標本周波数48000 Hz、量子化ビット数24 bitのモノラル音源として書き出された。

##### 3.3.1.3 実験手続き

タッピング課題の実験手続きは、予備実験1のタッピング課題の実験手続き(3.1.1.3項)とほぼ同様であった。重複するため、詳細は割愛する。

予備実験1と異なる点は、提示音圧レベル、教示内容、本番の試行の構成であった。まず、提示音圧レベルは76 dBとし、参加者間で固定とした。次に、教示内容は「各音に合うようにタッピングする」であった。また、その「各音に合うようにタッピングするとは、音に合わせて手拍

<sup>12</sup>予備実験1や予備実験2と異なり、「立ち上がり時間の割合」が0%の刺激音は、「立ち上がり時間」が0 msであった(1 msではなかった)。

子する時のぴったり合う感覚が得られる位置、あるいは、アンサンブルで演奏するときぴったり合う位置を目指してタッピングを行う感覚」であるとも教示した。最後に、本番では、同一刺激音を 700 ms(約 86 BPM) ごとに 80 回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため全 79 音列であった。これらの全 79 音列より、ランダムに 7 音列あるいは 8 音列を連結したものを 1 試行とし、全 10 試行が参加者に提示された。試行間に休憩が挟まれた。本番の所要時間は約 1 時間 30 分であった。

#### 3.3.1.4 実験装置

予備実験 1 の実験装置 (3.1.1.4 項) と同様であった。重複するため、詳細は割愛する。

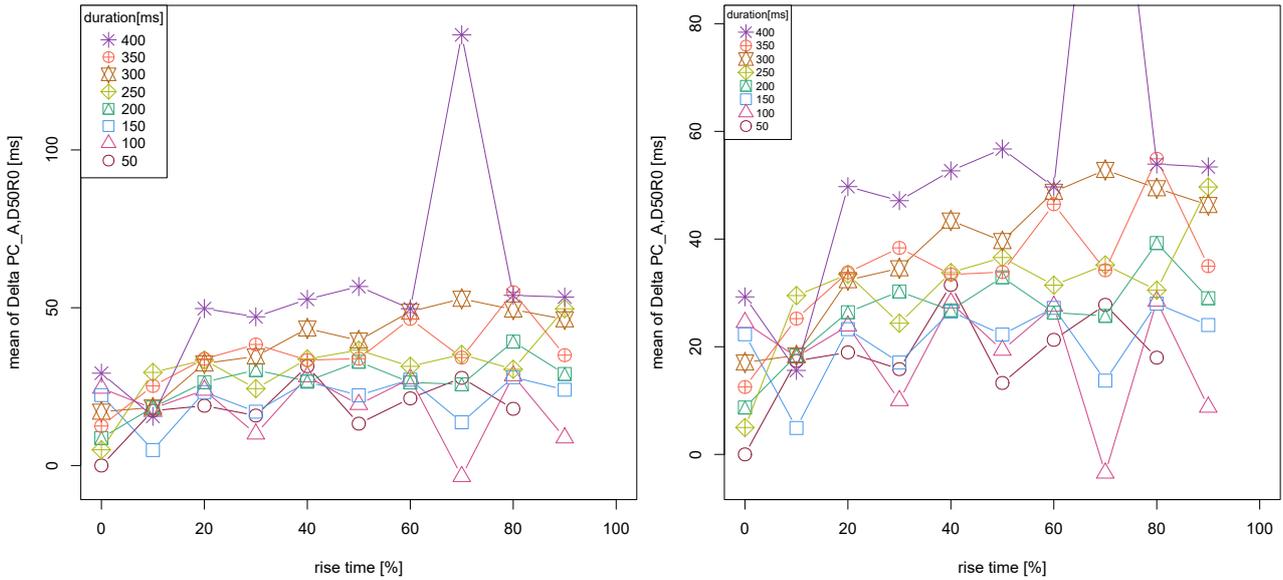
### 3.3.2 分析と結果、考察

2.6 節で述べた仮説検証の実験方法と分析方法によって仮説の検証が可能であるかを確認するため、分析を行った。

先に、分析の前処理を行った。まず、タップ時刻  $t_{\text{tap}}(A, i)$  を、タッピングの録音データより算出した。この方法の詳細は後述の 4.2.1 項を参照してほしい。次に、タップ時刻を用いて、各刺激音  $A$  の音列における、21 タップ目から 80 タップ目までを用いて、全 79 個の刺激音の同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  を算出した。この際、より厳格に参加者の反応を得るために、刺激音の再生とタッピングの録音との遅延時間 (標準化周波数が 48 kHz の時の 6 サンプル) を差し引いた。そして、 $\Delta\text{PC}_{A,R}$  を算出した。この予備実験 3 では、最も持続時間と立ち上がり時間の割合が小さい音を基準の刺激音  $R$  とした。すなわち、持続時間 50 ms、立ち上がり時間 0 ms の音であった。以後 “D50R0 音” と称す。従って、基準の刺激音と、基準の刺激音を含めた全 79 個の刺激音との  $\Delta\text{PC}_{A,D50R0}$  が算出された。

次に、全 79 個の刺激音の  $\Delta\text{PC}_{A,D50R0}$  を用いて、2.6 節で述べた方法で分析を行った。この分析の目的は「立ち上がり時間の割合の増加と PAT の移動との関係」が持続時間によって異なるかどうかを明らかにすることであった。

まず、「立ち上がり時間の割合の増加と  $\Delta\text{PC}_{A,D50R0}$  の変化との関係」を図示するため、ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、 $\Delta\text{PC}_{A,D50R0}$  (縦軸) を図 3.3.1 に示した。この予備実験 3 では、基準音  $R$  と持続時間 400 ms で立ち上がり時間の割合 70 % の刺激音 (D400R70 音) との  $\Delta\text{PC}_{A,D50R0}$ 、すなわち  $\Delta\text{PC}_{D400R70,D50R0}$  が、その他の刺激音との  $\Delta\text{PC}_{A,D50R0}$  よりもかなり大きい値であった。 $\Delta\text{PC}_{D400R70,D50R0}$  を図の枠内に含めると、その他の  $\Delta\text{PC}_{A,D50R0}$  の変化が見にくくなるため、図 3.3.1 の左には  $\Delta\text{PC}_{D400R70,D50R0}$  を図の枠内に含めた図を、右には  $\Delta\text{PC}_{D400R70,D50R0}$  を含めない図を載せた。これらの図より、持



(a)  $\Delta PC_{D400R70,D50R0}$  を図の枠内に含めた図

(b)  $\Delta PC_{D400R70,D50R0}$  を図の枠内に含めない図

図 3.3.1: 持続時間ごとの立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係

ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、 $\Delta PC_{A,D50R0}$  (縦軸) が示された。

なお、 $\Delta PC_{D400R70,D50R0}$  とは、持続時間 400 ms で立ち上がり時間の割合 70% の刺激音 (D400R70 音) と基準音 (D50R0 音) との  $\Delta PC_{A,D50R0}$  である。

持続時間が長い時には、立ち上がり時間の割合が増加すると  $\Delta PC_{A,D50R0}$  が大きくなる傾向、すなわち持続時間と立ち上がり時間の割合との交互作用が明らかであった。

次に、各持続時間における「立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係」を最も近似する 4 パラメータロジスティック回帰モデルを求めた。求められた回帰モデルの予測線を先ほどの図に重ねたものが図 3.3.2 である。ここでは、予測線の変化が見やすくなるよう、 $\Delta PC_{D400R70,D50R0}$  を図の枠内に含めない図のみ載せた。この図でも、持続時間が長い時には、立ち上がり時間の割合が増加すると  $\Delta PC_{A,D50R0}$  が大きくなる傾向、すなわち持続時間と立ち上がり時間の割合と交互作用が明らかであった。具体的には、持続時間が短い場合は、立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D50R0}$  の変化量が小さく、持続時間が長い場合は、 $\Delta PC_{A,D50R0}$  の変化量が大きいことが示された。

さらに、4 パラメータロジスティック回帰モデルにおける、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係」を強く反映する 2 つの指標、すなわち「立ち上がり時間の割合が 0% の時の予測値 (PAT) と立ち上がり時間の割合が 100% の時の予測値 (PAT) との差分」と「パラメータの一つである傾き  $b$  の値」を、求められた回帰モデルより算出し、これらの値が持続時間によって異なるかをクラスター分析を行って確認した。クラスター分析の結果を、図 3.3.3 に示した。この図より、持続時間が (1) 50, 100, 150, 200 ms、(2) 250, 300, 350, 400 ms の 2 つの

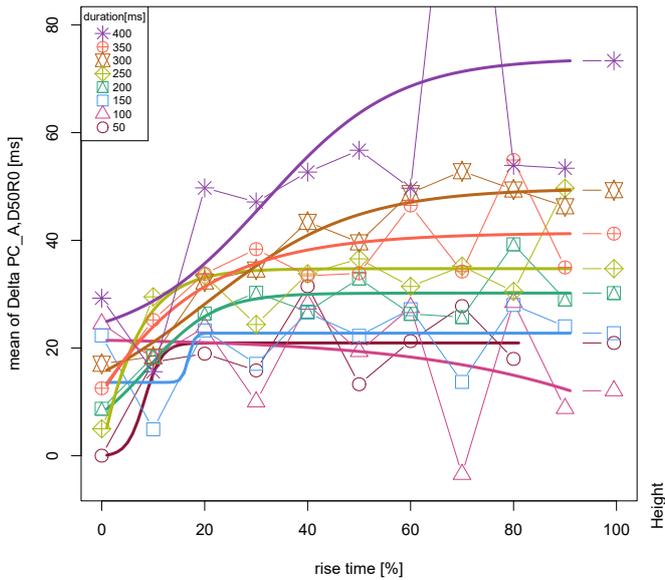


図 3.3.2: 持続時間ごとの立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係とそれを近似したロジスティック回帰モデルの予測線 ( $\Delta PC_{D400R70,D50R0}$  を図の枠内に含めない図)

ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  の変化 (縦軸) が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

なお、 $\Delta PC_{D400R70,D50R0}$  とは、持続時間 400 ms で立ち上がり時間の割合 70% の刺激音 (D400R70 音) と基準音 (D50R0 音) との  $\Delta PC_{A,D50R0}$  である。

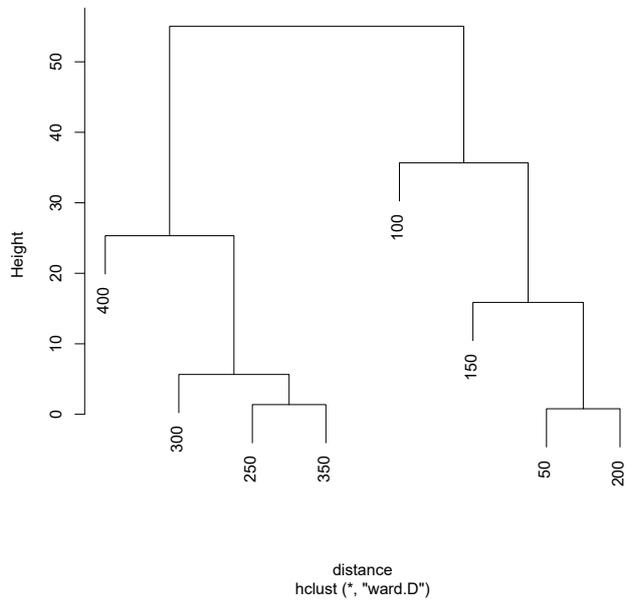


図 3.3.3: 2つの指標を用いたクラスター分析

グループに分けられた。すなわち、200 ms 以下と 250 ms 以上とのグループに分けられることが明らかになった。

以上の予備実験の結果より、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係」が持続時間が 200 ms 以下と 250 ms 以上とで異なる可能性が示唆された。2.5 節で述べた通り、統合の時間窓の長さは 150–300 ms の範囲のどこかにあると考えられた。すなわち、この 200–250 ms は、統合の時間窓が存在するとされる時間範囲に重なる。実験に用いた刺激音の持続時間が、統合の時間窓の範囲内から範囲外までに及ぶ以上、統合の時間窓の影響を受けないとは考えにくい。そこで、この予備実験 3 の条件下における参加者たちの統合の時間窓の長さが 200–250 ms であると考えれば、予備実験の結果、仮説が支持される可能性が示唆されたこととなる。もちろん、予備実験はサンプルサイズが小さいため、正確な結果とは言えないが、少なくとも、持続時間と立ち上がり時間の割合を変化させた刺激音を用いることで、仮説の検証が可能であることが示された。

### 3.3.3 本実験における刺激音の水準とサンプルサイズの検討

この予備実験 3 では全 79 個の刺激音を用いたが、本実験では仮説の検証に不可欠な刺激音のみに絞る必要がある。同時に、本実験ではサンプルサイズを増やす必要もある。そこで、本項にて、本実験で用いる刺激音の水準、並びに、サンプルサイズを検討した。

検討に入る前に、予備実験 3 で用いられた刺激音を再度述べる。持続時間は 8 水準 (50, 100, 150, 200, 250, 300, 350, 400 ms) であり、立ち上がり時間の割合は 10 水準 (0, 10, 20, 30, 40, 50, 60, 70, 80, 90 %) であった。これらの刺激音より、持続時間 50 ms で立ち上がり時間の割合 90 % の刺激音を除いた、計 79 個の刺激音が用いられた。

#### 刺激音の水準の検討

分析より、持続時間が 200 ms 以下と 250 ms 以上とで「立ち上がり時間の割合の増加と  $\Delta PC_{A,D50R0}$  の変化との関係」が異なる可能性が示唆された。このことより、仮説検証のためには、少なくとも 200 ms と 250 ms の持続時間を刺激音として用いる必要がある。だが、予備実験はサンプルサイズが小さかったため、結果が不安定である可能性が高い。そこで、200 ms の持続時間の一つ前の水準である 150 ms と、250 ms の一つ後の水準である 300 ms の持続時間も刺激音として用いる方が望ましい。加えて、本論では、仮説の検証のため、音楽群への実験 1 を基準として、これよりも統合の時間窓がより短いと想定される場合として、実験 2 にて一般群を対象に PAT 測定を行い、実験 3 にてより高い周波数の刺激音への PAT 測定を行う (2.5 章参照)。そのため、150 ms の持続時間よりさらに一つ短い持続時間である 100 ms も追加する方が望ましい。

一方、立ち上がり時間の割合の水準数は、ロジスティック回帰モデルの安定性を左右する。立ち上がり時間の割合を1水準減らすだけでもロジスティック回帰モデルが不安定になってしまう可能性が高いため、本実験の立ち上がり時間の割合の水準は予備実験のままとする。

従って、本実験の刺激音は、持続時間を5水準(100, 150, 200, 250, 300 ms)で変化させ、立ち上がり時間の割合を10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)で変化させる。

## サンプルサイズの検討

4パラメータロジスティック回帰モデルではサンプルサイズを決める方法が確立されていない。そこで、本論では、次の二つの指標を用いて、サンプルサイズの検討を行った。一つ目は赤池情報量基準(Akaike's Information Criterion; AIC)<sup>13</sup>であり、二つ目は予測値と実測値との相関係数は、4パラメータロジスティック回帰モデルの予測値と実際に測定された $\Delta PC_{A,D50R0}$ とのピアソンの積率相関係数であった。

この2つの指標の値を、予備実験にて得られた全3人のデータよりサンプルサイズが1-3人のそれぞれの場合の可能な組み合わせ( ${}_nC_r$ )において算出し、その後、可能な組み合わせの数で平均した<sup>14</sup>。続いて、サンプルサイズが1-3人の場合におけるAICと相関係数の値に最も近似する予測線を当てはめた。これにより、サンプルサイズが4人以上の場合の2つの指標の予測を行った。AICは上限や下限がある指標ではなく、見た目としても直線回帰の当てはまりが良いように感じられたため、直線回帰で近似した。相関係数は上限が1であり下限が0である指標であるため、0-1を限界とする2パラメータロジスティック回帰モデルで近似した。以上のように算出された、サンプルサイズが1-3人の場合の2つの指標の値とその予測線を図3.3.4に示した。

サンプルサイズを決定するためには閾値が必要である。恣意的ではあるが、AICの閾値はサンプルサイズが1の時のAICの値の半分である32とし、予測値と実測値との相関係数の閾値は、実際の実験結果より算出した場合にその上限である1となることが考えにくいため、1の手前である0.99とした。閾値を初めて超えるサンプルサイズを求めたところ、相関係数ではサンプルサイズが7人であり、AICではサンプルサイズが8人であることがわかった。これらにより、本実験のサンプルサイズを8人と決めた。

<sup>13</sup>赤池情報量基準(Akaike's Information Criterion; AIC)とは、より良いモデルを選択するとき用いられる基準の一つである。AICはモデルの予測の良さを重視した基準であり、値が小さい方が予測が良いとされる。

<sup>14</sup>例えば、3人のデータをA, B, Cとし、サンプルサイズ2人の場合として算出することを想定する。サンプルサイズが2人の場合、可能な組み合わせ( ${}_3C_2$ )は、AB, AC, BCの計3つの組み合わせである。このそれぞれの組み合わせの時の、 $\Delta PC_{A,D50R0}$ を求め、ロジスティック回帰モデルで近似をし、AICと相関係数を算出した。その後、組み合わせの数である3で平均をとった。このような計算をサンプルサイズが1-3人のそれぞれの場合において行った。

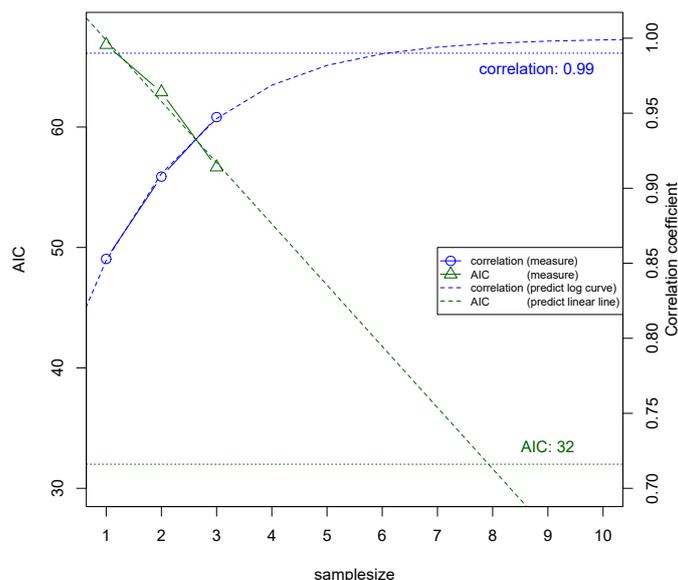


図 3.3.4: サンプルサイズを変化させたときの AIC 並びに予測値と実測値の相関係数  
 サンプルサイズが 1-3 の時の AIC と相関係数の値を実線とシンボルにて示した。また、それらを近似する回帰モデルの予測値を破線にて示した。そして、閾値を点線にて示し、閾値の数値も右側に記した。

### 3.4 結論

本章にて 3 つの予備実験を行った。

予備実験 1 において、タッピング課題で、ある A 音の PAT を測定するために必要な、A 音が繰り返される数 (タッピングの数) を検討した。実験と分析より、タッピング課題においてある A 音の PAT を測定するためには、実験において 80 回のタッピングを行い、分析において 21 タップ目から 80 タップ目までの 60 タップを用いて同調のズレ量の平均値を算出することが望ましいと判断された。

予備実験 2 において、タッピング課題、同時調整課題、等時調整課題の 3 つを、同一の指標 ( $\Delta PC_{A,R}$ ) で比較した。また、合わせて、提示音圧レベルを変化させた場合に PAT が変化するかどうかを確認した。実験と分析より、3 つの課題より得られる PAT が有意に異なることが明らかになった。また、詳細にみれば異なる傾向があるが、これらは全て、先行研究で報告されたことと一致した。さらに、タッピング課題において、音列内のタッピングの数を 110 回から 80 回に削減しても、同調のズレ量の平均値に有意差がなかったことが確認された。加えて、提示音圧レベルを 76 dB と 84 dB とで変化させても PAT が変化しないことも明らかになった。そのため、本論では提示音圧レベルを統制して実験を行い、提示音圧レベルについてはこれ以上は扱わないこととした。

以上の予備実験 1 と予備実験 2 より、本実験においてタッピング課題を用いることが妥当で

あると示された。

予備実験3として、2.6節で述べた仮説検証の実験方法と分析方法を用いて実験と分析を行った。考察の結果、確かに、仮説の検証が可能であることが明らかになった。また、本実験での刺激音は、持続時間が5水準(100, 150, 200, 250, 300 ms)、立ち上がり時間の割合が10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)が望ましいと判断された。また、この予備実験3と同様の音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行うならば、サンプルサイズが8人であることが望ましいと判断された。

# 第4章 本実験に共通する実験方法と分析方法

本章では、本実験として行う3つの実験、すなわち、実験1(5章)、実験2(6章)、実験3(7章)に共通する実験方法ならびに分析方法を述べる。

## 4.1 実験方法

### 4.1.1 刺激音

サイン波を用いた。実験1と実験2では周波数が440 Hzであった。これは、現在の音楽における基準となる周波数だからであり、JIS規格でも楽器用標準周波数はト音記号の音名“イ”の音の周波数で440 Hzであると定められている [13] からである。一方、実験3では、統合の時間窓の長さが異なる場合として周波数を高くし、3520 Hzとした。これは、先行研究を参考とし、また、440 Hzの3オクターブ上でありピアノの音域内であることより決定した。

持続時間と「立ち上がり時間の割合(持続時間に占める立ち上がり時間の割合であり、数式では(立ち上がり時間)/(持続時間)×100[%]である)」を変化させた。持続時間は、実験ごとに異なった。実験1では5水準(100, 150, 200, 250, 300 ms)であり、実験2と実験3では4水準(100, 150, 200, 250 ms)であった。実験2と実験3は実験1よりも統合の時間窓が短いと想定されたので、実験を簡略化するため持続時間が最も長い1水準を除いたからだ。立ち上がり時間の割合は、共通して10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)であった<sup>1</sup>。従って、実験1では、持続時間が5水準×立ち上がり時間の割合が10水準で、計50個の刺激音が作成された。実験2と実験3では、持続時間4水準×立ち上がり時間の割合10水準で、計40個の刺激音が作成された。

立ち上がり部分あるいは立ち下がり部分を振幅エンベロープにおいて線形に変化させると、振幅の急峻かつ大幅な変動によって目的の周波数以外をも含むノイズが僅かに発生する。そこで、立ち下がり部分は、このノイズを軽減させるため、10 ms分コサインカーブ状に変化させ、一方、立ち上がり部分は僅かなノイズの発生を許容し、直線に変化させた。これは、立ち上がり時間の割合を10水準(10%刻み)で変化させることを優先したためであった。

---

<sup>1</sup>予備実験1や予備実験2と異なり、立ち上がり時間の割合が0%の刺激音は、「立ち上がり時間」が0 msであった(1 msではなかった)。

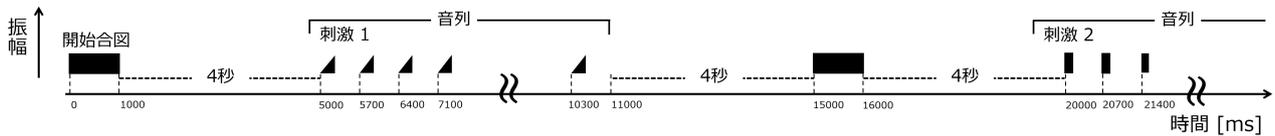


図 4.1.1: タッピング課題における提示した 1 試行の冒頭の模式図 (振幅エンベロープ)

横軸が時間を、縦軸が振幅を表す。1 試行は、開始合図音と刺激音の音列より構成される。開始合図音は持続時間 1 秒で 880 Hz のサイン波である。この開始合図音が鳴り終わってから 4 秒後に音列が開始する。音列では同一刺激音が 700 ms ごとに 80 回繰り返された。1 試行内では、5-8 音列が提示された。音列と音列の間は、作成の都合上、前の音列の最後の刺激音の物理的オンセットから 700 ms 後をその音列の末尾とし、そこから 4 秒後に次の音列の開始合図音が鳴り始めるように設定された。

刺激音は、PC 上で Python を用いて生成され、標本周波数 48000 Hz、量子化ビット数 24 bit のモノラル音源として書き出された。

#### 4.1.2 実験手続き

タッピング課題を用いた。タッピング課題とは、ある A 音が、メトロノームのように等時間間隔に、数十回繰り返される音列を聴きながら、各音に合うように指でタッピングを行う課題である。

実験では、「各音に合うようにタッピングする」よう教示した。また、その「各音に合うようにタッピングするとは、音に合わせて手拍子する時のぴったり合う感覚が得られる位置、あるいは、アンサンブルで演奏するときぴったり合う位置を目指してタッピングを行う感覚」であると教示した。

タップ箇所はマイク内蔵オーディオ・インターフェース (Line6 Sonic Port VX) のマイクグリル近傍のボディ部分を指定した。できる限り安定してタップできる体勢で、できる限り一定の強さでタップするよう教示した。

本番前にトレーニング課題を設けた。刺激音は、持続時間が 50 ms、立ち上がり時間が 0 ms、立ち下がり時間が 0 ms の 440 Hz のサイン波とした。この刺激音を 700 ms (約 86 BPM) ごとに 20 回繰り返した音列を 2 度提示し、それぞれにタッピングさせた。

本番では、同一刺激音を 700 ms (約 86 BPM) ごとに 80 回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため、実験 1 では 50 音列、実験 2 と実験 3 では 40 音列が用意された。これらの音列よりランダムに 5 音列から 8 音列を連結したものを 1 試行とし、複数の試行が参加者に提示された。各音列の開始合図音は、持続時間 1 秒、立ち上がり時間 0 ms、立ち下がり時間 0 ms の 880 Hz のサイン音であった。1 試行内は休憩を挟まず続けさせ、試行間で必要に応じて休憩を挟んだ。この提示した 1 試行の冒頭を模式図として示したのが図 4.1.1 である。

3.2 節の予備実験 2 より、提示音圧レベルが変化しても PAT は有意に変化しなかった。その

ため、提示音圧レベルは参加者間で固定とした。各刺激音の立ち上がり時間が異なることで提示音圧レベルも異なるが、全刺激音平均の提示音圧レベルは 76 dB であった。この提示音圧レベルの測定は、比較的反響の少ないスタジオにて、ダミーヘッド (Brüel&Kjær Type 4128C)、アンプ (NEXUS Conditioning Amplifier)、実験に使用したセミオープン型ヘッドフォン (AKG K240 MK II)、音圧校正済み騒音計の  $z$  特性 (周波数の重み付けが無いもの) で行われた。

フィードバック音は提示しなかった。だが、実験に用いたヘッドフォンはセミオープン型ヘッドフォンであったため、殆どの参加者は自身のタップ音が聞こえていたと考えられる。ただし、実験 2(6 章) の一般群の参加者の内、特にタップの弱かった 2-3 人は、自身のタップ音は聞こえていなかった可能性も考えられた。

### 4.1.3 実験装置

刺激音は、PC、マイク内蔵オーディオ・インターフェース (Line6 Sonic Port VX)、セミオープン型ヘッドフォン (AKG K240 MK II) を通して参加者の両耳に提示した。各参加者のタッピングは、刺激音提示と同期して、マイク内蔵オーディオ・インターフェース (同上) を通し、PC 上のソフトウェアにて録音された。

刺激音の再生とタッピングの録音との遅延時間は、ヘッドフォン (同上) をマイク (同上) の上に置き、刺激音を録音し、目視で測定したところ、標本周波数 48000 Hz で 6 サンプルであった。

## 4.2 分析方法

### 4.2.1 分析の前処理

まず、参加者ごとに、実験より得られたタッピングの録音データから、ピーク検出を行い、各刺激音  $A$  の音列における全てのタップ時刻  $t_{\text{tap}}(A, i)$  を算出した。その方法は次の通りである。

タッピングの録音データより、振幅エンベロープを算出し、時間遅延が生じないように、そのエンベロープにローパスフィルタ (第二種チェビシェフフィルタ、2 次、通過域端周波数 60 Hz) をかけた。その後、その振幅エンベロープを dB に変換し、以下のピーク位置の条件を満たすサンプルを検出した。その条件は三つあり、一つ目はそのサンプルの振幅が検出閾値 (音列全体の最大振幅値より 18 dB 下の値<sup>2</sup>) よりも大きいこと、二つ目はそのサンプルの振幅値が一つ前のサンプルの振幅の値と一つ後のサンプルの振幅の値よりも大きいこと、三つ目はそのサンプルから 10 ms 前のサンプルまでの間に 8 dB 以上のレベル差があること、であった。この三つの条件が全て満たされたサンプルをピーク位置とした。そして、このピーク位置より 6 dB 下がったところをタップ時刻  $t_{\text{tap}}(A, i)$  とした。

<sup>2</sup>実験 2(6 章) の一般群の参加者の内の 2-3 人は、特にタップの弱く、この検出閾値ではタップの数が本来のタップの数の半分以下しか検出されなかった。そのため、この参加者たちのみ、検出閾値を音列全体の最大振幅値より 26 dB 下の値に変更した。

まれに、 $i$  回目の  $A$  音に対して複数のピーク位置が検出されることがあった。その場合、 $i$  回目の  $A$  音の物理的オンセットに最も近いピーク位置を採用し、タップ時刻を算出した。また、 $i$  回目の  $A$  音の物理的オンセットから 400 ms ( $A$  音を繰り返した時間間隔が 700 ms であるため、その半分強の時間間隔) の間にピーク位置が検出されなかった場合、タップし損ねたと判断し、その音に対するタップ時刻の算出を行わなかった。

次に、参加者ごとの、各刺激音  $A$  に対する同調のズレ量の平均値  $\overline{\text{Asy}}(A)$  を算出した。同調のズレ量の平均値の定義式を再掲する (11 ページの式 2.4)。

$$\overline{\text{Asy}}(A) = \frac{1}{n} \sum_{i=1}^n (t_{\text{tap}}(A, i) - t_{\text{onset}}(A, i))$$

すなわち、参加者ごとに、各刺激音  $A$  の音列における、21 音目から 80 音目までの  $A$  音の物理的オンセット  $t_{\text{onset}}(A, i)$  と、それらに対するタップ時刻  $t_{\text{tap}}(A, i)$  の差分としての同調のズレ量を算出し、その後その平均値を算出した。

この際、より厳格に参加者の反応を得るために、刺激音の再生とタッピングの録音との遅延時間 (標本化周波数が 48 kHz の時の 6 サンプル) を差し引いた。これにより、2.3.1 節でタッピング課題の問題点として挙げた、タップを記録する装置の遅延が含まれてしまう問題を解消した。

その後、 $\Delta\text{PC}_{A,R}$  が算出された。基準の刺激音  $R$  は、刺激音の中で、持続時間と立ち上がり時間の割合が最も小さい音とした。すなわち、共通して、持続時間が 100 ms、立ち上がり時間の割合が 0% の刺激音であった。以後 “D100R0 音” と略記する。従って、基準の刺激音と基準の刺激音を含めた全ての刺激音との  $\Delta\text{PC}_{A,D100R0}$  を算出した。これはまず参加者ごとに算出され、その後、参加者の数で平均された。

同調のズレ量の平均値から  $\Delta\text{PC}_{A,D100R0}$  の算出過程は 2.3.1 ページにて述べた。それを  $\Delta\text{PC}_{A,D100R0}$  に合わせて書き換えた式を以下に、その模式図を図 4.2.1 に示す。合わせて参照してほしい。

$$\begin{aligned} \Delta\text{PC}_{A,D100R0} &= \Delta\text{AO}_A - \Delta\text{AO}_{D100R0} \\ &= (\Delta\text{AO}_A - b) - (\Delta\text{AO}_{D100R0} - b) \\ &= \overline{\text{Asy}}(A) - \overline{\text{Asy}}(D100R0) \end{aligned}$$

## 4.2.2 分析方法

分析の目的は、「立ち上がり時間の割合の増加と  $\Delta\text{PC}_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを明らかにし、仮説を検証することであった。

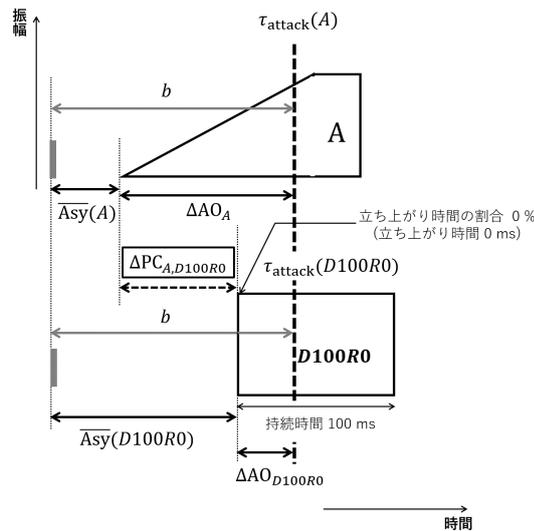


図 4.2.1:  $\Delta PC_{A,D100R0}$  の模式図、及び、基準の刺激音 D100R0 音の同調のズレ量の平均値  $\overline{Asy}(D100R0)$  と A 音の同調のズレ量の平均値  $\overline{Asy}(A)$  と  $\Delta PC_{A,D100R0}$  の関係図 (振幅エンベロープ)  
横軸が時間を、縦軸が振幅を表す。灰色の長方形のシンボルがタップを表す。この図では、A 音は例として適当な刺激音とした。 $\Delta PC_{A,D100R0}$  を算出する際には、この図における A 音の位置に、それぞれの刺激音を当てはめて計算する。

分析に先立ち、生データより、実験の意図に沿わないものや他と大きく異なるデータがないかを確認した。まず、各刺激音 A に対する 60 個の同調のズレ量  $Asy(A, i)$  を確認した。次に、同調のズレ量の平均値  $\overline{Asy}(A)$  を確認した。さらに、同調のズレ量  $Asy(A, i)$  の標準偏差も確認した。その後、分析として、次の 4 工程を行った。

まず、ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、PAT (縦軸) を図にし、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を示した。

次に、各持続時間におけるこの関係を最も近似する 4 パラメータロジスティック回帰モデルを求めた。そして、求められた回帰モデルの予測線を先ほどの図に重ね、その予測線の形が持続時間によって異なるかを視覚的に確認した。

さらに、4 パラメータロジスティック回帰モデルにおける、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を強く反映する 2 つの指標、すなわち「立ち上がり時間の割合が 0% の時の予測値 (PAT) と立ち上がり時間の割合が 100% の時の予測値 (PAT) との差分」と「パラメータの一つである傾き  $b$  の値」を求められた回帰モデルより算出した。そして、これらの値が持続時間によって異なるかを、ワード法によるクラスター分析を行って確認した。

最後に、 $\Delta PC_{A,D100R0}$  と従属変数とし、持続時間、立ち上がり時間の割合を独立変数とした、二要因参加者内計画の分散分析である。参加者内計画の分散分析の前提である球面性の仮定が成り立たない場合には、Greenhouse-Geisser の自由度調整を行い、 $p$  値補正を行った。本論では非線形であるロジスティック回帰モデルを前提とするため、線形モデルを用いる分散分析は本論の仮説の検証には適切ではないと考えられる。だが殆どの先行研究では分散分析を行って

いるため、それらとの比較を目的に本論でも分散分析を行った。

# 第5章 本実験1: 持続時間や立ち上がり時間の割合の増加とPATの変化との関係と、統合の時間窓

本章より、タッピング課題を用いて、本論の仮説である「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことを検証する。

本章では、タッピング課題を、音楽訓練を受けたことがある参加者(音楽群)に対して行った。

## 5.1 実験方法

### 5.1.1 参加者

3.3節の予備実験3での検討に従い、参加者は8人であった。参加者はみな、東京藝術大学音楽学部音楽環境創造科、並びに同大学院音楽音響創造研究分野に所属しており、音楽訓練を受けたことがあった。本論における音楽訓練を受けたことがある参加者の定義は「今までに定期的な演奏訓練の経験があった人」であった。

平均年齢は26.3歳であった。音響経験(録音やPAなど)の平均年数が5.25年であった。演奏経験の平均年数が11.9年、そのうち師についていた平均年数が9.6年であった。演奏本番経験のある人(今までに1回でもコンクールや演奏発表会などに出たことがある)が8人中4人であった。演奏楽器はピアノ、箏、ヴァイオリン、サクソ、ギター、トロンボーンであった。音楽聴取は多くの人が一日に1時間程度であり、主にBGMとしてであった。自発的にはあまり聴かない人もいたが、その人たちも仕事や練習では聴いたり演奏したりしていた。

### 5.1.2 刺激音

440 Hzのサイン波を用いた。持続時間は5水準(100, 150, 200, 250, 300 ms)とし、立ち上がり時間の割合は10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)とした。従って、計50個の刺激音が作られた。

立ち上がり部分と立ち下がり部分の形状、立ち下がり時間、そして生成方法は本論共通のもの(4章参照)を使用した。

### 5.1.3 実験手続き

実験手続きは、本論共通のもの(4章)を使用した。以下では、本番の試行の構成のみ述べる。

本番では、同一刺激音を700ms(約86BPM)ごとに80回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため、全50音列であった。これらの全50音列より、ランダムに7音列あるいは8音列を連結したものを1試行とし、全7試行が参加者に提示された。1試行内は休憩を挟まず続けさせ、必要に応じて試行間に休憩を挟んだ。実験所要時間は約1時間15分であった。

### 5.1.4 実験装置

実験装置は、本論共通のもの(4章)を使用した。

### 5.1.5 実験環境と実験日程

実験は、常時も突発的にも大きな音がしない静かな部屋で、2020年8月3日から8月5日までの間に一人ずつ行われた。

## 5.2 分析と結果

仮説の検証のため、「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」が持続時間によって異なるかを分析した。

主な分析の流れは本論共通の手順(4章)にて記載した通りである。

### 5.2.1 生データの確認およびデータ処理

生データの確認を行い、実験の意図に沿わないものや、他と大きく異なるデータがないかを確認した。

#### 同調のズレ量のバイオリンプロット

参加者ごとに各刺激音 $A$ に対する60個の同調のズレ量 $A_{sy}(A, i)$ を算出し、これらをバイオリンプロットで示した(付録A)。バイオリンプロットとは、ヒストグラムの平滑化の方法の一つであり、横向きのヒストグラムをカーネル密度推定を用いて平滑化したものである。そのため、バイオリンプロットより、どの範囲に値が多いかを知ることができ、また、値が多い範囲から離れたところにも値があるかどうかを知ることができる。さらに、23ページで述べたように、「PATが何らかの確率密度分布に従う」という考え方に立脚する先行研究では、PATの分

布が実験結果よりカーネル密度推定で算出された。バイオリンプロットを示せば、それらの先行研究と比較することも可能となる。

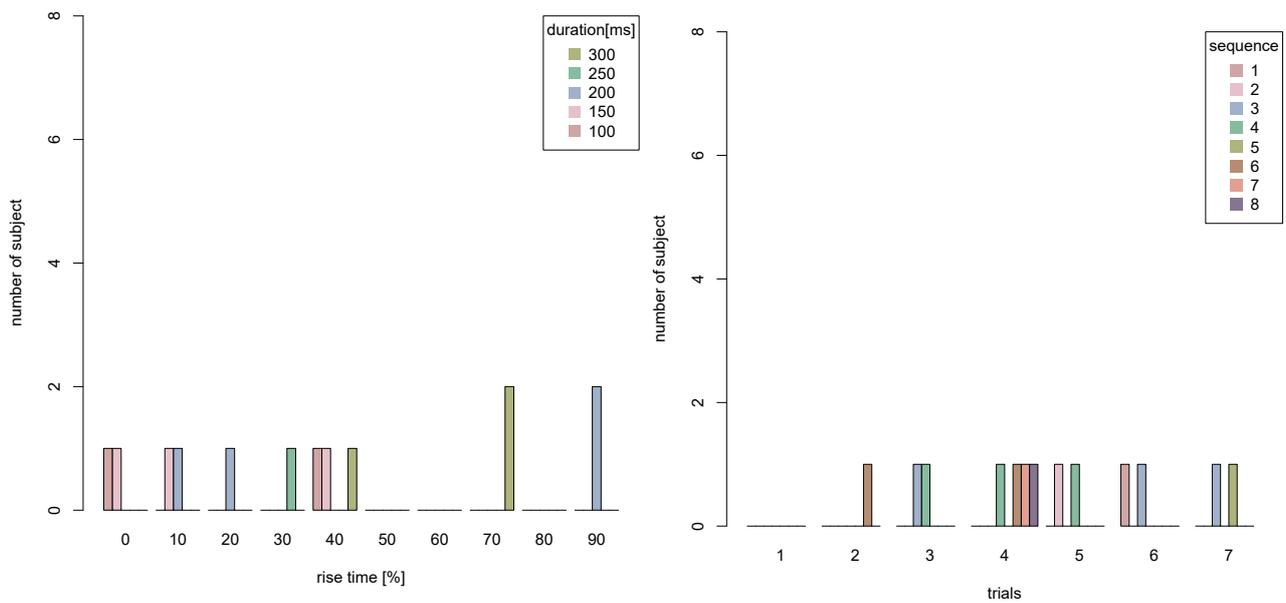
これらのバイオリンプロットより、全ての参加者の全ての刺激音の同調のズレ量の分布が単峰分布であることが確認された。多くの刺激音の分布は、両端が細くなる(正規分布のような)形を示した。だが、たまに、長方形に近い(尖度が負のような)形や、左右非対称の(歪度の絶対値が大きいような)形もあった。例えば、長方形に近い形であったのは、参加者 D の刺激音 49 などであった。また、左右非対称の形であったのは、参加者 A の刺激音 1 や、参加者 B の刺激音 1 や刺激音 3 などであった。

さらに、ごくまれに、ある刺激音の同調のズレ量の分布の端が非常に長かった。これは、その刺激音の同調のズレ量の中に、1-2 個程度、大きく外れた値があったことを表している。この 1-2 個程度の大きく外れた同調のズレ量はタッピングのミスより偶発的に得られた可能性が高いと考えられた。それらの大きく外れた同調のズレ量の値は  $\pm 200$  ms 以上であった。そこで、 $\pm 200$  ms 以上の同調のズレ量となったタッピングが多くの参加者に共通して、特定の刺激音や特定の提示順にあるかどうかを確認した。 $\pm 200$  ms 以上の同調のズレ量であったタッピングが属する刺激音と提示順を調べ、このタッピングが一つでもあった参加者の数を刺激音ごとにあるいは提示順ごとに数えた。これを棒グラフで表したものが図 5.2.1 である。この図の左図に刺激音ごとの棒グラフを、右図に提示順ごとの棒グラフを掲載した。図の詳しい見方はキャプションを参照してほしい。これらの図より、刺激音でも提示順でも、 $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数は最大で 2 人であった(この実験 1 の参加者は全 8 人であった)。以上より、極めてまれに  $\pm 200$  ms 以上の同調のズレ量のタッピングが確認されたが、これらは刺激音や提示順に依存しないことが確認され、偶発的に得られたタッピングであると考えられた。そうであれば、これらのタッピングは集中力の低下やタッピングの際の手指の動作ミスに起因すると推察される。このようなデータは実験意図に沿わないものであるため、本論では、 $\pm 200$  ms 以上の同調のズレ量のタッピングは削除した。

### 参加者ごとの各刺激音に対する同調のズレ量の平均値の変化量

参加者ごとに、各刺激音 A に対する同調のズレ量の平均値を算出し、縦に並べて図 5.2.2 に示した。この図では、各参加者において、白丸のシンボルがどのくらい縦に広がっているかによって、各刺激音に対する同調のズレ量の平均値がどのくらい変化し、どの範囲で変化したかを知ることができる。すなわち、その変化量と変化範囲を知ることができる。前者の変化量は数値として表 5.2.1 にも示した。

まず、表 5.2.1 より、各刺激音に対する同調のズレ量の平均値の変化量が参加者によって異なることが明らかである。参加者 A の変化量が最も小さく、56 ms であった。一方、参加者 C の



(a) 刺激音ごと

横軸の“rise time[%]”は刺激音の立ち上がり時間の割合を表し、縦軸は参加者の数を表し、凡例の“duration[ms]”は刺激音の持続時間を表す。各棒グラフは±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。横軸は刺激音の立ち上がり時間の割合であり、それぞれの立ち上がり時間の割合の箇所には、左から右へ、持続時間が短い刺激音から長い刺激音に対する棒グラフが並んでいる。

(b) 提示順ごと

横軸の“trials”は試行を表し、凡例の“sequence”は音列を表す。各棒グラフは±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。実験で提示された試行が横軸に示され、各試行の箇所には、左から右へ、各試行内で提示された順の各音列に対する棒グラフが並んでいる。すなわち、左から右へと、提示された時系列で並んでいる。

図 5.2.1: ±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数

変化量が最も大きく、100 ms であった。だが、図 5.2.2 を見ると、参加者 A と参加者 C を含めた全参加者の、大半の白丸のシンボルの縦方向への拡がり具合は大きく異ならなかった。加えて、同図より、全参加者の、白丸のシンボルが拡がる範囲も大きく異ならなかった。強いて言うならば、参加者 A と参加者 C の範囲がその他の参加者よりも負の方向にある程度であった。

以上のことより、各刺激音に対する同調のズレ量の平均値の変化量や変化範囲は、参加者によって僅かに異なるが、それらが他の参加者と全く異なる参加者はいなかったと判断した。

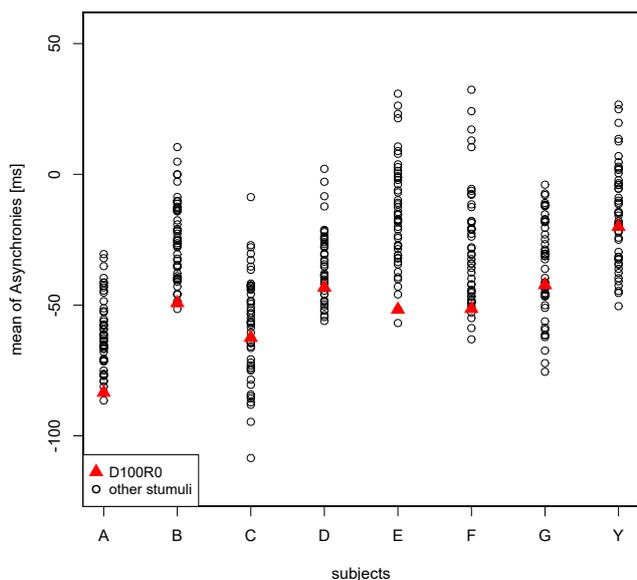


図 5.2.2: 参加者ごとの、各刺激音に対する同調のズレ量の平均値  
 図の横軸の “subjects” は参加者番号を表し、縦軸の “mean of asynchronies[ms]” は同調のズレ量の平均値を表す。白丸のシンボルが、各刺激音 A に対する同調のズレ量の平均値を、赤色塗りつぶしの三角のシンボルが、 $\Delta PC_{A,D100R0}$  の基準の刺激音である D100R0 音に対する同調のズレ量の平均値を示す。

表 5.2.1: 参加者ごとの、各刺激音に対する同調のズレ量の平均値の変化量 (最大値 – 最小値)

参加者	変化量 [ms]
A	56
B	62
C	100
D	58
E	88
F	95
G	72
Y	77

### 同調のズレ量の標準偏差の確認

まず、参加者ごとに、全刺激音で平均した、同調のズレ量の標準偏差を算出し、箱ひげ図として図 5.2.3 に示した。この箱ひげ図の見方はキャプションを参照してほしい。この箱ひげ図より外れ値がないことが明らかである。すなわち、全刺激音で平均した標準偏差が他の参加者と大きく異なる参加者がいなかったことが確認された。

次に、刺激音ごとに、全参加者で平均した、同調のズレ量の標準偏差を算出し、棒グラフとして図 5.2.4 に示した。この図の見方はキャプションを参照してほしい。

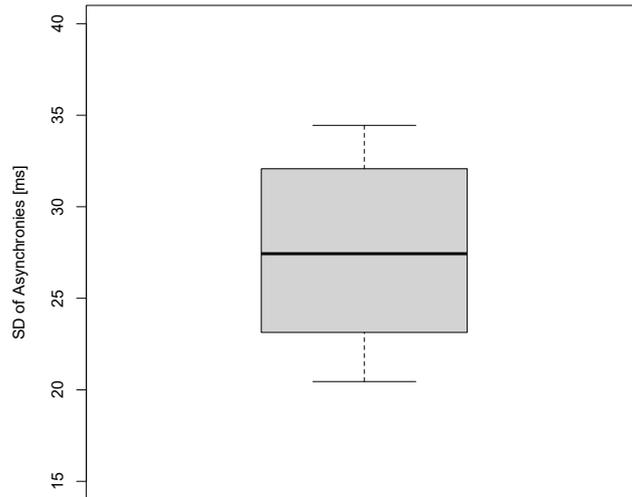


図 5.2.3: 参加者ごとの、全刺激音で平均した同調のズレ量の標準偏差の箱ひげ図

縦軸の“SD of Asynchronies [ms]”が同調のズレ量の標準偏差を表す。この箱ひげ図では、中央値が箱の中の太線で表され、四分位範囲(第1四分位から第3四分位の差分)すなわちデータのばらつきが箱の大きさで表される。また、外れ値は「第1四分位・第3四分位から四分位範囲の1.5倍以上離れた値」と定義され、白い丸のシンボルで表される。そして、外れ値を除いた最小値と最大値がひげの端で表される。

この図より、全参加者で平均した同調のズレ量の標準偏差は、刺激音によってばらつくが、持続時間や立ち上がり時間の割合が増加するにつれて標準偏差も増加するという傾向は見られなかった。

また、2.3.1節で述べた通り、音楽訓練を受けた人のSMS課題における標準偏差はIOIの2-4%であると言われる。本実験の、刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差は、値にして24-33 msであり<sup>1</sup>、割合にしてIOI(700 ms)の3.4-4.7%であった。本実験の参加者は専門的な音楽訓練を受けた人と趣味の人が混在したが、先行研究と近い値が得られたといえる。

以上のように生データの確認を行った。同調のズレ量の中に、偶発的でかつ他と大きく異なるデータが含まれていたため、それらのみ削除した。その他に大きく異なるデータは見られず、データの削除が必要なものはなかった。

<sup>1</sup>この値は、全刺激音においても、持続時間が250 msまでの刺激音に限っても、同一であった。このことは、次章以降にて実験間の標準偏差を比較する際に想起してほしい。

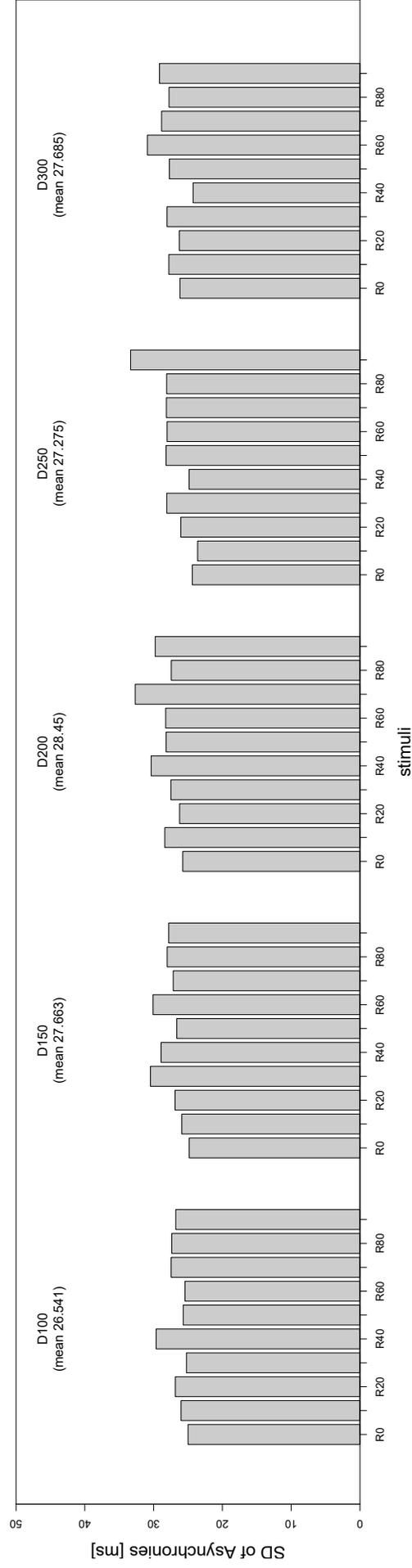


図 5.2.4: 刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差の棒グラフ

全 50 個の刺激音 (持続時間 5 水準 × 立ち上がり時間の割合 10 水準) に対する、全参加者で平均した同調のズレ量の標準偏差が棒グラフで表されている。それは、10 本ごとに一つのまとまりになっており、これらは同じ持続時間を持つ刺激音である。10 本の棒グラフの上方に、その持続時間の値が “D + 値” で略記された。合わせて、その下に、その 10 個の刺激音で平均をとった、全参加者で平均した同調のズレ量の平均値の標準偏差の値も示された。その 10 本の棒グラフは、それぞれ、左から順に、立ち上がり時間の割合が小さい刺激音から大きい刺激音に対する値を表す棒グラフが並んでいる。横軸に立ち上がり時間の割合の値が “R + 値” と略記された。

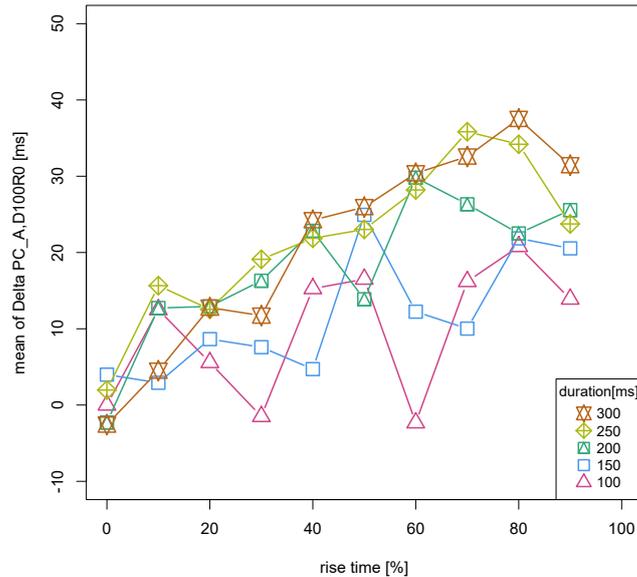


図 5.2.5: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」  
 横軸の“rise time [%]”は刺激音の立ち上がり時間の割合を表し、縦軸の“mean of Delta PC\_A,D100R0”  
 は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の立ち上がり時間  
 の割合を増加させた時の (横軸)、PAT (縦軸) が示されている。

### 5.2.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

仮説の検証のため、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを分析した。

まず、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を図に示した。それが図 5.2.5 である。この図より、持続時間が長くなると、立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D100R0}$  の変化量が大きくなる傾向が明らかであった。

次に、各持続時間におけるこの関係を最も近似する 4 パラメータロジスティック回帰モデルを求め、求められた回帰モデルの予測線を先ほどの図に重ねた。それが図 5.2.6 である。この図より、持続時間が 150 ms 以下の時の予測線と、持続時間が 200 ms 以上の時の予測線とで、その予測線の形が異なることが示唆された。

そして、4 パラメータロジスティック回帰モデルにおける、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を強く反映する 2 つの指標を算出した。その値を表 5.2.2 に示した。これらの 2 つの指標の値を用いてウォード法によるクラスター分析を行い、持続時間を分類した。その結果が、図 5.2.7 である。この図より、持続時間が、(1)100, 150 ms、(2-1)200, 250 ms、(2-2)300 ms のグループに分けられると示された。すなわち、「立ち上がり時間の割合の増加と PAT の移動との関係」が、持続時間が 150 ms 以下と、持続時間が 200 ms 以上とで異

表 5.2.2: 近似した4パラメータロジスティック回帰モデルにおける2つの指標の値

持続時間 [ms]	パラメータの一つである傾き $b$ の値	立ち上がり時間の割合が0%の時の予測値 (PAT) と立ち上がり時間の割合が100%の時の予測値 (PAT) との差分
100	-0.018	12.716
150	-0.018	17.322
200	-0.043	25.449
250	-0.037	26.781
300	-0.052	36.479

なることが示された。

以上より、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」は、持続時間が150 ms 以下と持続時間が200 ms 以上とで次のように異なるといえるだろう。持続時間が150 ms 以下では、立ち上がり時間の割合が増加するにつれ、 $\Delta PC_{A,D100R0}$  がゆるやかに増加した。立ち上がり時間の割合が0-90%まで増加した時の  $\Delta PC_{A,D100R0}$  の増加量は、約13-17 ms であり、持続時間が200 ms 以上の場合のその増加量よりも小さかった。ロジスティック回帰モデルの予測線の形としては、S字曲線の下の方の形(独立変数の変化に対する従属変数の変化量が徐々に大きくなる)か、あるいはS字の真ん中の部分(独立変数の変化に対する従属変数の変化量がほぼ一定であり直線に近い形となる部分)であると示唆された。一方、持続時間が200 ms 以上では、立ち上がり時間の割合が増加するにつれ、 $\Delta PC_{A,D100R0}$  が急激に増加した。立ち上がり時間の割合が0-90%まで増加した時の  $\Delta PC_{A,D100R0}$  の増加量は、約25-36 ms であり、持続時間が150 ms 以下の場合のその増加量よりも大きかった。加えて、立ち上がり時間の割合が60-70%以上になると、立ち上がり時間の割合が増加するにつれての  $\Delta PC_{A,D100R0}$  の増加量は小さくなった。ロジスティック回帰モデルの予測線の形としては、S字の上の方(独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分)であると示唆された。

### 5.2.3 分散分析

先行研究との比較のため、 $\Delta PC_{A,D100R0}$  を従属変数とし、持続時間と立ち上がり時間の割合を独立変数とした、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内計画の分散分析を行った(表5.2.3)。

その結果、持続時間と立ち上がり時間の割合との交互作用 ( $F(36, 252) = 1.524, p = 0.034, \eta_p^2 = 0.179$ ) が5%水準で有意であった。交互作用が有意であったため、主効果はあまり意味をなさないが、先行研究との比較のために主効果も報告する。立ち上がり時間の割合の主効果 ( $F(9, 63) = 9.681, p < 0.000, \eta_p^2 = 0.580$ ) と持続時間の主効果 ( $F(4, 28) = 5.678, p = 0.002, \eta_p^2 = 0.448$ ) が5%水準で有意であった。

交互作用の下位検定の結果、有意な組み合わせがあった。そのため、さらに下位検定としてShafferの多重比較を行った。その結果、調整後  $p$  値が5%水準で有意であった組み合わせは、持

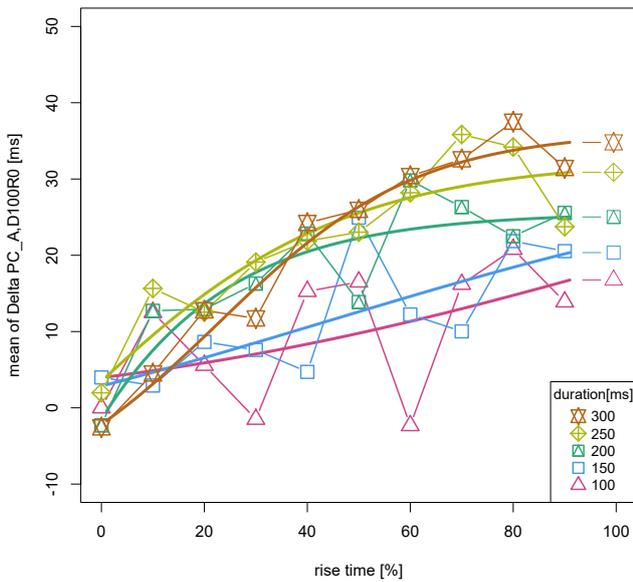


図 5.2.6: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」と、それを最も近似する4パラメータロジスティック回帰モデルの予測線  
横軸の“rise time [%]”は立ち上がり時間の割合を表し、縦軸の“mean of Delta  $PC_{A,D100R0}$ ”は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音(シンボル)における、その音の立ち上がり時間の割合を増加させた時の(横軸)、 $\Delta PC_{A,D100R0}$ (縦軸)が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

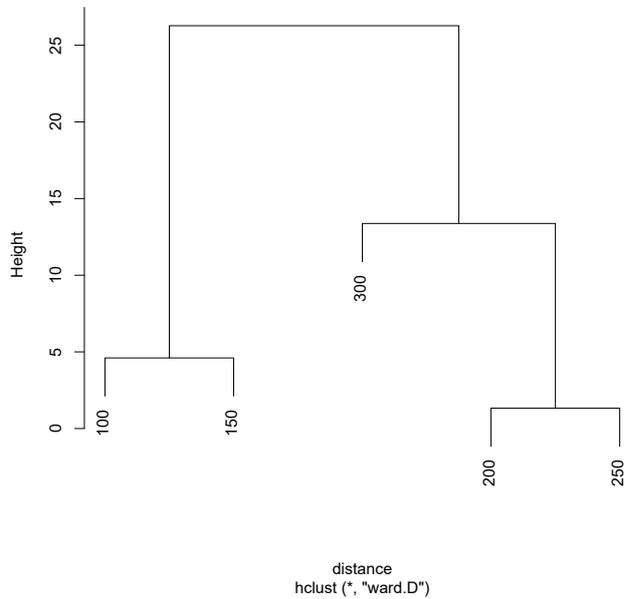


図 5.2.7: 近似した4パラメータロジスティック回帰モデルにおける2つの指標を用いたクラスター分析の結果

表 5.2.3: 実験1の分散分析表

	SS	df	MS	F-ratio	p-value	$\eta_p^2$
参加者	49 849.283	7	7 121.326			
持続時間	9 316.121	4	2 329.030	5.677	0.002	0.448
参加者 × 持続時間	11 486.572	28	410.235			
立ち上がり時間の割合	25 307.299	9	2 811.922	9.681	0.000	0.580
参加者 × 立ち上がり時間の割合	18 298.155	63	290.447			
持続時間 × 立ち上がり時間の割合	11 133.874	36	309.274	1.524	0.034	0.179
参加者 × 持続時間 × 立ち上がり時間の割合	51 131.544	252	202.903			
Total	176 522.848	399	442.413			

続時間と立ち上がり時間の割合の各水準との間では、以下の4つであり、

- 立ち上がり時間の割合 60%の時の、持続時間 100 ms と 300 ms の間 ( $t(7) = 6.093, adj.p = 0.005$ )
- 立ち上がり時間の割合 60%の時の、持続時間 100 ms と 200 ms の間 ( $t(7) = 4.798, adj.p = 0.012$ )
- 立ち上がり時間の割合 70%の時の、持続時間 150 ms と 250 ms の間 ( $t(7) = 4.123, adj.p = 0.044$ )
- 立ち上がり時間の割合 70%の時の、持続時間 150 ms と 200 ms の間 ( $t(7) = 3.662, adj.p = 0.048$ )

持続時間と立ち上がり時間の割合の各水準との間では、以下の3つが有意であった。

- 持続時間 200 ms の時の、立ち上がり時間の割合 0%と 70%の間 ( $t(7) = 6.739, adj.p = 0.012$ )
- 持続時間 200 ms の時の、立ち上がり時間の割合 0%と 60%の間 ( $t(7) = 5.343, adj.p = 0.039$ )
- 持続時間 250 ms の時の、立ち上がり時間の割合 0%と 30%の間 ( $t(7) = 7.903, adj.p = 0.004$ )

すなわち、立ち上がり時間の割合が60%と70%の時のみ、持続時間が150 ms以下の条件と200 ms以上の条件との間に有意な組み合わせがあった。また、持続時間が200 msと250 msの時のみ、立ち上がり時間の割合が0%との組み合わせが有意であったといえる。

## 5.3 考察

以下では、分析の流れに沿って、考察を行い、仮説の検証を行う。

### 5.3.1 生データの確認

まず、各参加者の各刺激音に対する同調のズレ量のバイオリンプロットより、全ての参加者の全ての刺激音の同調のズレ量の分布が単峰分布であることが明らかとなった。23 ページで述べたように、先行研究ではタッピング課題で測定された PAT は単峰分布となることが報告された。従って、本論の結果は先行研究と一致した。

だが、本論のバイオリンプロットにおいて、多くの刺激音の分布は両端が細くなる形を示したが、たまに、長方形に近い形や左右非対称の形もあった。先行研究では、同時調整課題や等時調整課題では刺激音によって PAT が双峰分布になることもあると報告された。本論では分布の算出以外を行わないが、PAT の分布の変化が偶然得られたものなのか、刺激音や参加者、測定課題に起因するのかなどは今後の研究が必要である。

そして、同調のズレ量の平均値と同調のズレ量の標準偏差からは、個人差が大きいことが確認された。SMS 課題の先行研究より同調のズレ量の平均値は個人差が大きいことが報告され、PAT の先行研究からも得られた結果は個人差が大きいことが報告された。従って、本論の結果は、先行研究と一致した。

### 5.3.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

分析の結果、持続時間 150 ms 以下と 200 ms 以上とで、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が異なることが明らかになった。2.5 節で述べた通り、統合の時間窓の長さは 150–300 ms の範囲のどこかにあると考えられた。すなわち、この 150–200 ms は、統合の時間窓が存在するとされる時間範囲に重なる。実験で用いた刺激音の持続時間が統合の時間窓の範囲内から範囲外までに及ぶ以上、統合の時間窓の影響を受けないとは考えにくい。そこで、本実験の条件下における参加者たちの統合の時間窓の長さが 150–200 ms の範囲にあると考えれば、本論の仮説である「PAT の知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持されたといえる。

加えて、クラスター分析においてグループ (2) は、(2-1) 持続時間 200, 250 ms と (2-2) 300 ms の下位グループに分けられた。回帰モデルの予測線と合わせて考えると、持続時間が統合の時間窓以上の場合、持続時間が長くなるにつれ、立ち上がり時間の割合の増加に対する  $\Delta PC_{A,D100R0}$  の変化量が、より大きくなる可能性が示唆された。

そうとはいえ、本論では、統合の時間窓を直接測定していない。そのため、以上の結論は、この実験 1 の参加者の統合の時間窓の長さが 150–200 ms の範囲にあるという仮定の下でのものである。本当に仮説が支持されるかどうかを結論づけるには、実験 2 と実験 3 にて統合の時間窓の長さが異なる時の刺激音の PAT と比較をした後に検討する必要がある。

### 5.3.3 分散分析

$\Delta PC_{A,D100R0}$  を従属変数とし、持続時間、立ち上がり時間の割合を独立変数とした、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内計画の分散分析を行った。その結果、持続時間と立ち上

がり時間の割合との交互作用が5%水準で有意であった。そして、多重比較の結果、立ち上がり時間の割合が60%と70%の時のみ、持続時間が150ms以下の条件と200ms以上の条件との間に有意な組み合わせがあった。また、持続時間が200msと250msの時のみ、立ち上がり時間の割合が0%との組み合わせが有意であった。

以上のことより、立ち上がり時間の割合が大きい場合のみ、持続時間による差が有意となったことが明らかである。また、持続時間が長い場合のみ、立ち上がり時間の割合が0%の時とそれ以上の時との間で有意となったことも明らかである。どちらの場合も、持続時間と立ち上がり時間の割合との交互作用を反映している。言い換えると、分析にて明らかになった『持続時間150ms以下と200ms以上とで、「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」が異なる』ことを反映していると考えられる。一方で、それを正確に反映した場合に有意になるはずの組み合わせが有意にならなかったことも明らかである。例えば、持続時間が300msの時には有意な組み合わせがなく、立ち上がり時間の割合が80%と90%の時にも有意な組み合わせがなかった。これは、刺激音による標準偏差の違いや個人差が分散分析結果に影響した可能性も、 $\Delta PC_{A,D100R0}$ の実測値が予測値から上下にばらついたことが影響した可能性も考えられる。

先行研究の分散分析の結果と本論の結果との比較は考察の章(8.1節)にて行う。

## 5.4 結論

本論の仮説である、「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことを検証するため、持続時間と立ち上がり時間の割合を変化させた刺激音へのタッピング実験を、音楽訓練を受けたことがある参加者(音楽群)に対して行った。その結果、以下のことが明らかとなった。

- 持続時間150ms以下と200ms以上とで、「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」が異なることが明らかとなった。これにより、本論の仮説が支持される可能性が示された。

だが、統合の時間窓を直接測定していないため、本当に仮説が支持されるかどうかは、実験2と実験3にて統合の時間窓の長さが異なる時の刺激音のPATとの比較をした後に検討する必要がある。

- より具体的には、「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の変化との関係」は、音の持続時間が統合の時間窓の範囲内か範囲外かで、以下のように異なる可能性が示された。

- 音の持続時間が統合の時間窓以下 (150 ms 以下) では、立ち上がり時間の割合が増加するにつれ、 $PAT(\Delta PC_{A,D100R0})$  はゆるやかに時間的後方へ移動する。その移動量は音の持続時間が統合の時間窓以上 (200 ms 以上) の場合のその移動量よりも小さい。ロジスティック回帰モデルの予測線の形としては、S 字曲線の下の方の形 (独立変数の変化に対する従属変数の変化量が徐々に大きくなる) か、あるいは S 字の真ん中の部分 (独立変数の変化に対する従属変数の変化量がほぼ一定であり直線に近い形となる部分) である。
- 音の持続時間が統合の時間窓以上 (200 ms 以上) では、立ち上がり時間の割合が増加するにつれ、 $PAT(\Delta PC_{A,D100R0})$  が急激に時間的後方へ移動する。その移動量は音の持続時間が統合の時間窓以下 (150 ms 以下) の場合よりも大きい。だが、統合の時間窓以下と異なり、統合の時間窓以上では、立ち上がり時間の割合が 60–70 % 以上になると、立ち上がり時間の割合が増加するにつれての  $PAT$  の移動量が徐々に小さくなる。ロジスティック回帰モデルの予測線の形としては、S 字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) である。
- 統合の時間窓の範囲外 (持続時間 200 ms 以上) において、持続時間が長くなるにつれ、立ち上がり時間の割合の増加に対する  $PAT(\Delta PC_{A,D100R0})$  の変化量が、より大きくなる可能性も示唆された。

以上のように、本実験にて仮説が支持される可能性が示された。次章以降にて、統合の時間窓が異なる場合として、音楽訓練の有無 (6 章) と周波数 (7 章) を取り上げ、同じ刺激音の  $PAT$  を測定する。その後、統合の時間窓の長さが異なる時の  $PAT$  との比較を通して、更なる仮説の検証を行う。

## 第6章 本実験2: 統合の時間窓の長さが異なる場合(音楽訓練の有無)における、持続時間や立ち上がり時間の割合の増加とPATの変化との関係と、統合の時間窓

前章の実験1では、音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行い、本論の仮説の検証を行った。その結果、本論の仮説である「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持される可能性が示された。だが、本論では統合の時間窓を直接測定していないため、PATの知覚に統合の時間窓が関係しているのかを確認するには、統合の時間窓の長さが異なる時のPATと比較する必要があった。

そこで、実験2では、統合の時間窓の長さが異なる場合として、音楽訓練を受けたことがあるかどうか(音楽訓練の有無)を取り上げた。実験1の音楽群の参加者よりも統合の時間窓が短いと想定される、音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行った。

### 6.1 実験方法

#### 6.1.1 参加者

この実験2では音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行った。一般群はタッピングのばらつきが大きいことが考えられるため、音楽群に行った実験1よりも参加者の数を増やした。参加者は10人であった。

音楽訓練を受けたことがない参加者の定義は「小学生までの習い事を除き、音楽の授業以外で演奏経験がない人」とした。参加者の平均年齢は25歳であった。

10人の参加者の内、2人が小学校6年生までピアノを習っており、1人が小学生の時に3か月間のみピアノを習っていた。そして、7人が小学生の時を含めて音楽の授業以外で演奏経験が全くなかった。音楽聴取時間は平均して一日に2.65時間程度であった。音楽聴取ジャンルはJ-popを中心に様々なものを聴いていた。

実験後のインタビューで、リズムゲーム(通称「音ゲー」)の経験とダンスの経験を質問したが、多くの参加者はどちらも殆ど経験がなかった。これらは、リズムに合わせて指や身体を動かすことに慣れているか、また、リズム感が養われているかを調査するために質問された。これらの経験が豊富であれば、タッピングがより安定すると考えられたからである。

### 6.1.2 刺激音

440 Hz のサイン波を用いた。

一般群の参加者は、タッピングに不慣れだと想定された。そのため、集中して課題を遂行できるように、実験1よりも持続時間の水準を一つ割愛した。実験1よりも、実験2の統合の時間窓の方が短いと想定されたため、実験1の持続時間条件の中で最も長い300 msを除いた。すなわち、持続時間は4水準(100, 150, 200, 250 ms)、立ち上がり時間の割合は10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)とした。従って、持続時間4水準×立ち上がり時間の割合10水準の計40個の刺激音が作られた。

立ち上がり部分と立ち下がり部分の形状、立ち下がり時間、そして生成方法は本論共通のもの(4章参照)を使用した。

### 6.1.3 実験手続き

実験手続きは、本論共通のもの(4章)を使用した。以下では、本番の試行の構成のみ述べる。

本番では、同一刺激音を700 ms(約86 BPM)ごとに80回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため、全40音列であった。これらの全40音列より、ランダムに5音列あるいは7音列を連結したものを1試行とし、全6試行が参加者に提示された。1試行内は休憩を挟まず続けさせ、必要に応じて試行間に休憩を挟んだ。実験所要時間は約1時間であった。

### 6.1.4 実験装置

実験装置は、本論共通のもの(4章)を使用した。

### 6.1.5 実験環境と日程

実験は、常時も突発的にも大きな音がしない、静かな部屋で、2020年9月16日から10月23日までの間に一人ずつ行われた。

## 6.2 分析と結果

仮説の検証のため、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを分析した。

主な分析の流れは本論共通の手順 (4章) にて記載した通りである。

### 6.2.1 生データの確認およびデータ処理

生データの確認を行い、実験の意図に沿わないものや、他と大きく異なるデータがないかを確認した。

#### 同調のズレ量のバイオリンプロット

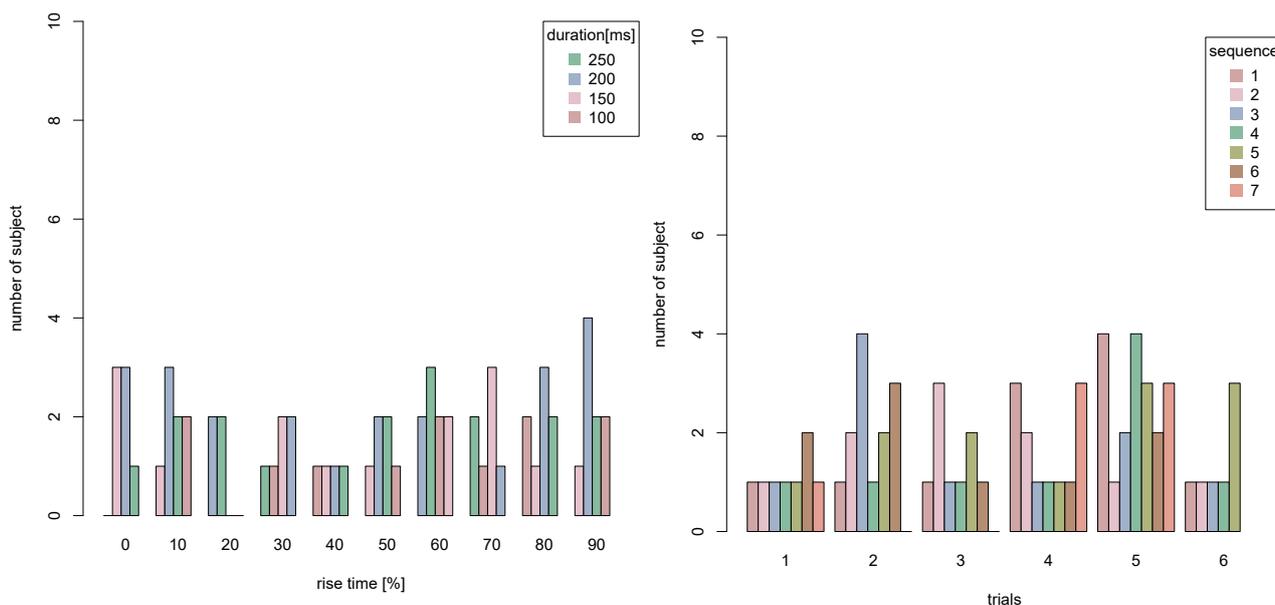
参加者ごとに各刺激音  $A$  に対する 60 個の同調のズレ量  $Asy(A, i)$  を算出し、これらをバイオリンプロットで示した (付録 B)。バイオリンプロットが何かについては前章の 5.2 節を参照してほしい。

これらのバイオリンプロットより、実験 1 と同様に、殆どの参加者の殆どの刺激音の同調のズレ量の分布は単峰分布であることが確認された。多くの刺激音の分布は、両端が細くなる (正規分布のような) 形を示した。だが、たまに、長方形に近い (尖度が負のような) 形や、左右非対称の (歪度の絶対値が大きいような) 形もあった。例えば、長方形に近い形であったのは、参加者 I の刺激音 39 や参加者 J の刺激音 17 などであった。また、左右非対称の形であったのは、参加者 H の刺激音 37 などであった。

さらに、実験 1 と同様に、ごくまれに、ある刺激音の同調のズレ量の分布の端が非常に長かった。それらの端の同調のズレ量の値は、 $\pm 200$  ms 以上であり、タッピングのミスより偶発的に得られた可能性が高いと考えられた。

だが、実験 1 と異なるのは、この実験 2 では、参加者 L に限って、ある刺激音のバイオリンプロットにおいて同調のズレ量が  $+200$ – $+300$  ms の間でも太くなることがあった (参加者 L の刺激音 11 と刺激音 34)。これは、実験後のインタビューにおける「思わず裏拍にタッピングしてしまったことがあった」とのコメントより、裏拍にタッピングしたことに起因すると考えられた。実際、参加者 L の、多くの刺激音の同調のズレ量は中央値が  $-100$ – $-50$  ms であるため、バイオリンプロットで同調のズレ量が  $+200$ – $+300$  ms 付近で太くなった場合のタッピングは、多くの刺激音よりも約 300–350 ms 遅く、すなわち同一刺激音が繰り返される IOI 700 ms の半分程度遅くタッピングしたことになる。すなわち、裏拍にタッピングしたことが明らかである。

タッピングのミスであれ、裏拍にタッピングしたのであれ、この  $\pm 200$  ms 以上の同調のズレ量のタッピングが、多くの参加者に共通して、特定の刺激音や特定の提示順にあるかどうかを



(a) 刺激音ごと

横軸の“rise time[%]”は刺激音の立ち上がり時間の割合を表し、縦軸は参加者の数を表し、凡例の“duration[ms]”は刺激音の持続時間を表す。各棒グラフは  $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。横軸は刺激音の立ち上がり時間の割合であり、それぞれの立ち上がり時間の割合の箇所には、左から右へ、持続時間が短い刺激音から長い刺激音に対する棒グラフが並んでいる。

(b) 提示順ごと

横軸の“trials”は試行を表し、凡例の“sequence”は音列を表す。各棒グラフは  $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。実験で提示された試行が横軸に示され、各試行の箇所には、左から右へ、各試行内で提示された順の各音列に対する棒グラフが並んでいる。すなわち、左から右へと、提示された時系列で並んでいる。

図 6.2.1:  $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数

確認した。実験 1 と同様に、 $\pm 200$  ms 以上の同調のズレ量であったタッピングが属する刺激音と提示順を調べ、このタッピングが一つでもあった参加者の数を刺激音ごとにあるいは提示順ごとに数え、棒グラフで表したものが図 6.2.1 である。図の詳しい見方はキャプションを参照してほしい。これらの図より、刺激音でも提示順でも、 $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数は最大で 4 人であった (この実験 2 の参加者は全 10 人であった)。また、実験 1 よりも実験 2 の方が、刺激音ごとでも提示順ごとでも、 $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数が多かった。以上のことより、この実験 2 でも、まれに  $\pm 200$  ms 以上の同調のズレ量のタッピングが確認されたが、これらは刺激音や提示順に依存しないことが確認され、偶発的に得られたタッピングであると考えられた。これらのタッピングは、実験 1 と同様に、集中力の低下やタッピングの際の手指の動作ミスに起因するものもあるが、この実験 2 では、誤って裏拍にタッピングしてしまったことに起因するものもあると推察された。いずれにせよ、このようなデータは実験意図に沿わないものであるため、本論では、 $\pm 200$  ms 以上の同調のズレ量のタッピングは削除した。

## 参加者ごとの刺激音に対する同調のズレ量の平均値の変化量

参加者ごとに、各刺激音 A に対する同調のズレ量の平均値を算出し、縦に並べて図 6.2.2 に示した。この図では、各刺激音に対する同調のズレ量の平均値の変化量と変化範囲を知ることができる。図の詳しい見方はキャプションを参照してほしい。前者の変化量は数値として表 6.2.1 にも示した。

まず、表 5.2.1 より、各刺激音に対する同調のズレ量の平均値の変化量が参加者によって異なることが明らかである。参加者 M の変化量が最も小さく、56 ms であった。一方、参加者 L の変化量が最も大きく、151 ms であった。確かに、図 6.2.2 における縦方向への拡がり具合を見ると、参加者 M はあまり拡がらずにまとまっており、参加者 L は拡がっている。だが、縦方向への拡がり具合は参加者によって大きくばらついているため、参加者 L や参加者 M のみが他の参加者と全く異なるとはいえなかった。加えて、同図より、全参加者の、白丸のシンボルが拡がる範囲を見ても、参加者 Q は他の参加者よりも負の方向にあるとはいえ、他の参加者と全く異なるとは言えなかった。

以上のことより、各刺激音に対する同調のズレ量の平均値の変化量や変化範囲は、参加者によってばらつくが、それらが他の参加者と全く異なる参加者がいるとは言えなかった。音楽訓練を受けたことがない(一般群の)参加者であることも考慮すると、ばらつきが大きくなることは想定されるため、実験意図に全く沿わない反応をした参加者はいなかったと判断した。

## 同調のズレ量の標準偏差の確認

まず、参加者ごとに、全刺激音で平均した、同調のズレ量の標準偏差を算出し、箱ひげ図として図 6.2.3 に示した。この箱ひげ図の見方はキャプションを参照してほしい。この箱ひげ図より、全刺激音で平均した標準偏差が外れ値となった参加者が二人いることが確認された。第 1 四分位からの外れ値(箱よりも下の白い丸のシンボル)であったのは参加者 K であり、第 3 四分位からの外れ値(箱よりも上の白い丸のシンボル)であったのは参加者 L であった。どちらも音楽経験がなかった。第 1 四分位からの外れ値は標準偏差が小さいことを示すため問題にはならないが、第 3 四分位からの外れ値は標準偏差が大きいことを示すため、問題になる。だが、第 3 四分位からの外れ値であった参加者 L は、同調のズレ量のバイオリンプロットの項で述べたように、裏拍にタッピングしたため、 $\pm 200$  ms 以上の同調のズレ量であったタッピングが多くあった参加者であった。これらのタッピングが削除されたため、参加者 L の同調のズレ量のデータ数は、他の参加者よりも少なかった。これらのことが影響し、参加者 L の同調のズレ量の標準偏差が外れ値になったと考えられた。このようなタッピングも一般群の参加者の反応の一つであると考え、データを削除しなかった。

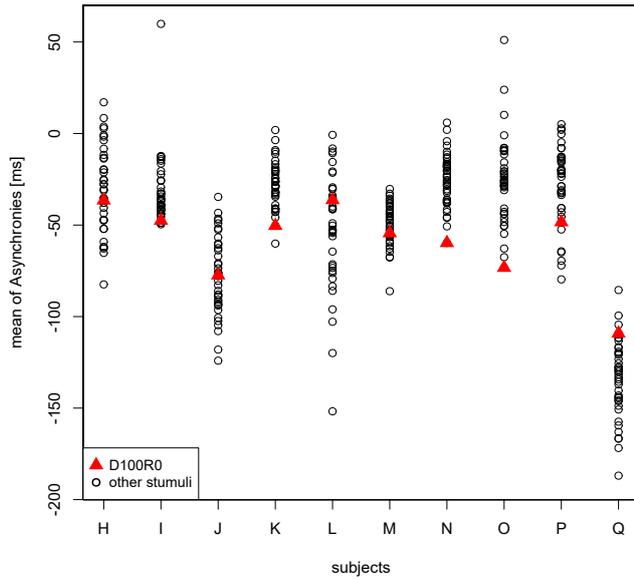


図 6.2.2: 参加者ごとの、各刺激音に対する同調のズレ量の平均値  
 図の横軸の“subjects”は参加者番号を表し、縦軸の“mean of asynchronies[ms]”は同調のズレ量の平均値を表す。白丸のシンボルが、各刺激音 A に対する同調のズレ量の平均値を、赤色塗りつぶしの三角のシンボルが、 $\Delta PC_{A,D100R0}$  の基準の刺激音である D100R0 音に対する同調のズレ量の平均値を示す。

表 6.2.1: 参加者ごとの、各刺激音に対する同調のズレ量の平均値の変化量 (最大値 – 最小値)

参加者	変動量 [ms]
H	99
I	109
J	89
K	62
L	151
M	56
N	66
O	124
P	85
Q	101

次に、刺激音ごとに、全参加者で平均した、同調のズレ量の標準偏差を算出し、棒グラフとして図 6.2.4 に示した。この図の見方はキャプションを参照してほしい。

この図より、実験 1 と異なり、全参加者で平均した同調のズレ量の標準偏差は、持続時間が長くなるにつれ、僅かに大きくなる傾向が見られた。各持続時間の標準偏差を平均した値 (図中の“D+ 値”) を参照すると、持続時間が 100 ms の時は標準偏差が約 30 ms であるが、持続時間が 150 ms と 200 ms では約 32 ms であり、持続時間が 250 ms では約 33 ms であった。ごく僅かであるが増加傾向があることが示された。

また、2.3.1 節で述べた通り、音楽訓練を受けた人の SMS 課題における標準偏差は IOI の 2–4% であると言われる。実際、実験 1 の音楽群における、刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差は、値にして 24–33 ms であり、割合にして IOI(700 ms) の 3.4–4.7% であった。そして、この実験 2 の一般群における、刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差は、値にして 27–39 ms であり、割合にして、IOI の 3.8–5.6% であった。従って、実験 1 の音楽群の参加者よりも、この実験 2 の一般群の参加者の標準偏差の方がやや大きいことが確認された。

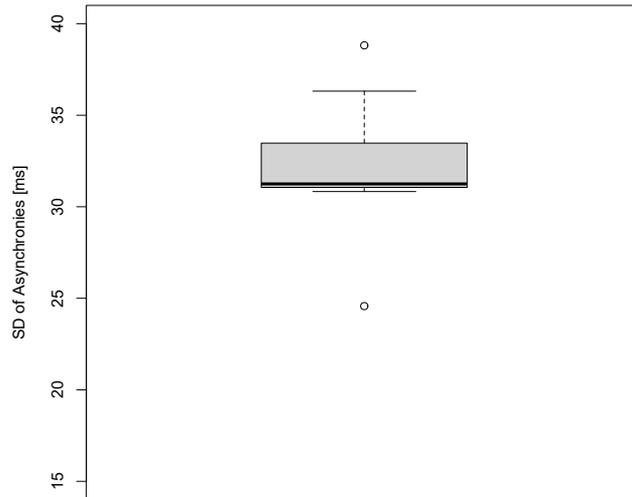


図 6.2.3: 参加者ごとの、全刺激音で平均した同調のズレ量の標準偏差の箱ひげ図

縦軸の“SD of Asynchronies[ms]”が同調のズレ量の標準偏差を表す。この箱ひげ図では、中央値が箱の中の太線で表され、四分位範囲(第1四分位から第3四分位の差分)すなわちデータのばらつきが箱の大きさで表される。また、外れ値は「第1四分位・第3四分位から四分位範囲の1.5倍以上離れた値」と定義され、白い丸のシンボルで表される。そして、外れ値を除いた最小値と最大値がひげの端で表される。

以上のように生データの確認を行った。同調のズレ量の中に、偶発的であつ他と大きく異なるデータや実験意図に沿わないデータが含まれていたため、それらのみ削除した。全体的に個人差やばらつきが大きかったが、音楽訓練を受けたことがない参加者の反応であると判断し、データの削除は行わなかった。

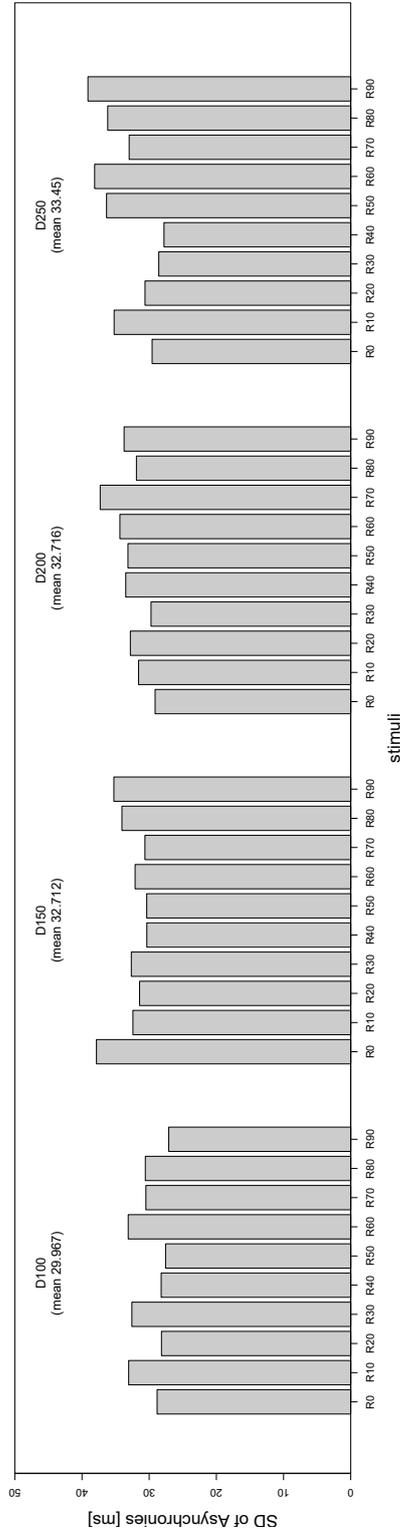


図 6.2.4: 刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差の棒グラフ

全 40 個の刺激音 (持続時間 4 水準 × 立ち上がり時間の割合 10 水準) に対する、全参加者で平均した同調のズレ量の標準偏差が棒グラフで表されている。それらは、10 本ごとに一つのまとまりになっており、一つのまとまりの 10 本は同じ持続時間の刺激音である。10 本の棒グラフの上方に、その持続時間の値が “D + 値” で略記された。合わせて、その下に、その 10 個の刺激音で平均をとった、全参加者で平均した同調のズレ量の平均値の標準偏差の値も示された。その 10 本の棒グラフは、それぞれ、左から順に、立ち上がり時間の割合が小さい刺激音から大きい刺激音に対する値を表す棒グラフが並んでいる。横軸に立ち上がり時間の割合の値が “R + 値” と略記された。

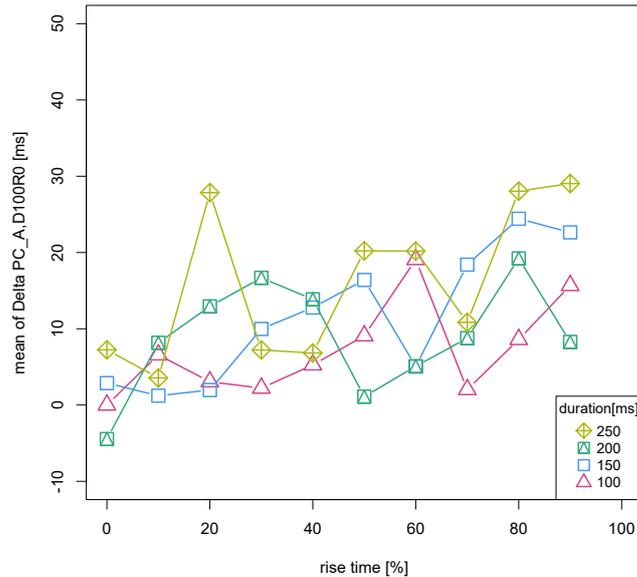


図 6.2.5: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」  
 横軸の“rise time [%]”は刺激音の立ち上がり時間の割合を表し、縦軸の“mean of Delta PC\_A,D100R0”は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、PAT (縦軸) が示されている。

## 6.2.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

仮説の検証のため、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを分析した。

まず、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を図に示した。それが図 6.2.5 である。

次に、各持続時間におけるこの関係を最も近似する 4 パラメータロジスティック回帰モデルを求め、求められた回帰モデルの予測線を先ほどの図に重ねた。それが図 6.2.6 である。この図より、持続時間が 100 ms の予測線の形と 200 ms の予測線の形とが類似しており、また、持続時間が 150 ms の時の予測線と 250 ms の予測線とが類似していることが示唆された。

そして、4 パラメータロジスティック回帰モデルにおける、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を強く反映する 2 つの指標を算出した。その値を 6.2.2 に示した。これらの 2 つの指標の値を用いてウォード法によるクラスター分析を行い、持続時間を分類した。その結果が、図 6.2.7 である。この図より、持続時間が、(1)100 ms、(2-1)200 ms、(2-2)150、250 ms のグループに分けられると示された。

これらの分析より、持続時間が 100 ms 以下と 150 ms 以上において、「立ち上がり時間の割合

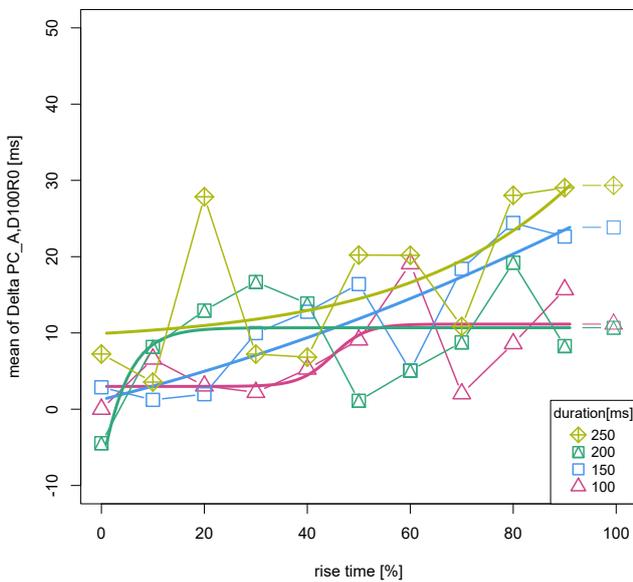


図 6.2.6: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」と、それを最も近似する4パラメータロジスティック回帰モデルの予測線  
 横軸の“rise time [%]”は立ち上がり時間の割合を表し、縦軸の“mean of Delta  $PC_{A,D100R0}$ ”は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音(シンボル)における、その音の立ち上がり時間の割合を増加させた時の(横軸)、 $\Delta PC_{A,D100R0}$  の変化(縦軸)が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

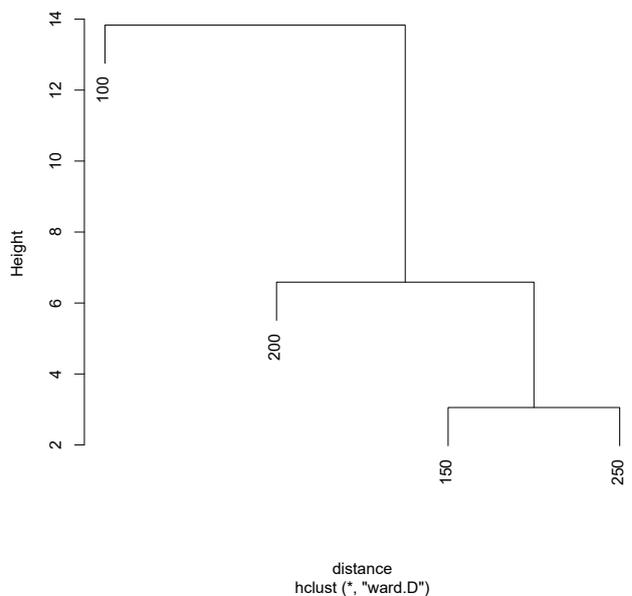


図 6.2.7: 近似した4パラメータロジスティック回帰モデルにおける2つの指標を用いたクラスター分析の結果

表 6.2.2: 近似した4パラメータロジスティック回帰モデルにおける2つの指標の値

持続時間 [ms]	パラメータの一つである傾き $b$ の値	立ち上がり時間の割合が0%の時の予測値 (PAT) と立ち上がり時間の割合が100%の時の予測値 (PAT) との差分
100	-0.306	8.169
150	-0.012	22.423
200	-0.225	15.195
250	-0.032	19.369

の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が異なることが示されたともいえるかもしれない。だが、そうとは言い切れない。なぜなら、一つ目に、クラスター分析の結果より、全ての持続時間が一つのグループであると判断することもできるからである。また、二つ目に、持続時間グループ (2) に属する 3 つの持続時間のロジスティック回帰モデルの予測線の形がそれぞれ異なるからである。具体的には、持続時間が 150 ms の時は、回帰モデルの予測線の形が S 字の真ん中の部分 (独立変数の変化に対する従属変数の変化量がほぼ一定であり直線に近い形となる部分) であり、持続時間が 250 ms の時は S 字曲線の下の方の形 (独立変数の変化に対する従属変数の変化量が徐々に大きくなる) であった。持続時間が 200 ms の時は S 字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) で、立ち上がり時間の割合 20% 以降では予測線が水平であった。これらのことを踏まえて敢えて述べるならば、全ての持続時間が一つのグループであり、どの持続時間でも、立ち上がり時間の割合が 70% 程度以降でのみ、立ち上がり時間の割合が増加すると  $\Delta PC_{A,D100R0}$  が僅かに増加するが、それ以前 (立ち上がり時間の割合が 0-70%) では立ち上がり時間の割合が増加しても  $\Delta PC_{A,D100R0}$  が変化しないといえるかもしれない。

このような結果となった原因として、参加者ごとに、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が異なる可能性が考えられた。そこで、これを確認した。まず、参加者ごとの、各刺激に対する  $\Delta PC_{A,D100R0}$  を用いて、参加者ごとの、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を最も近似する 4 パラメータロジスティック回帰モデルを求めた。次に、参加者ごとの、回帰モデルにおける 2 つの指標を算出した。そして、その参加者ごとの 2 つの指標を用いてクラスター分析を行い、参加者のグループ分けを行った。その結果が図 6.2.8 である。この図より、参加者が 3 人のグループと 7 人のグループに分けられることが明らかとなった。

続いて、それぞれのグループごとに、参加者平均した各刺激音に対する  $\Delta PC_{A,D100R0}$  を算出し、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」の図を示した。そして、各持続時間のこの関係を最も近似する 4 パラメータロジスティック回帰モデルを求め、求められた回帰モデルの予測線を先ほどの図に重ねた。それが、図 6.2.9 である。左図に参加者が 7 人のグループの結果を、右図に 3 人のグループの結果を示した。これらの図より、7 人のグループは、立ち上がり時間の割合が 70% 以降のみ立ち上がり時間の割合が増加すると  $\Delta PC_{A,D100R0}$  が僅かに増加するが、それ以前 (立ち上がり時間の割合が 0-70%) では立ち上がり時間の割合が増加しても  $\Delta PC_{A,D100R0}$  が変化しないことが示された。一方、3 人のグループでは、どの持続時間でも、立ち上がり時間の割合が増加するにつれて  $\Delta PC_{A,D100R0}$  が急激に増加した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれての  $\Delta PC_{A,D100R0}$  の移動量が徐々に小さくなるかあるいは変化しなくなった。回帰モデルの予測線の形としては、S 字の上の方と S 字曲線全体の形であった。

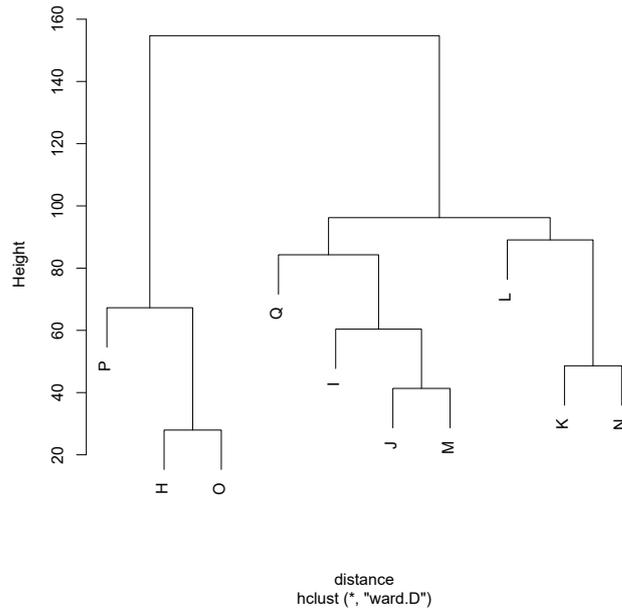


図 6.2.8: 参加者ごとに近似した4パラメータロジスティック回帰モデルにおける2つの指標を用いた、参加者を分類したクラスター分析の結果

以上のように、参加者は、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が異なる、2つのグループに分けられた。3人のグループと7人のグループとの参加者において、演奏経験や、音楽聴取時間や音楽聴取ジャンル、そしてリズムゲーム(通称「音ゲー」)やダンスの経験に差異はなかった。すなわち、本論では、グループ内の参加者のみに共通する特徴は明らかにならなかった。

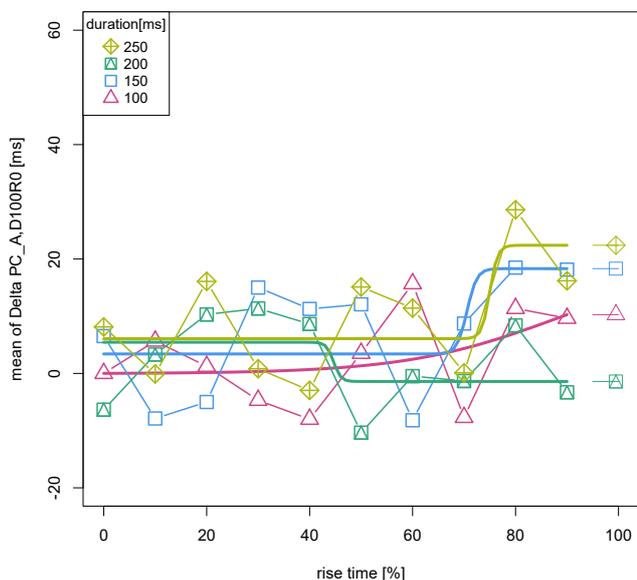
### 6.2.3 分散分析

先行研究との比較のため、 $\Delta PC_{A,D100R0}$  を従属変数とし、持続時間、立ち上がり時間の割合を独立変数として、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内分散分析を行った(表 6.2.3)。

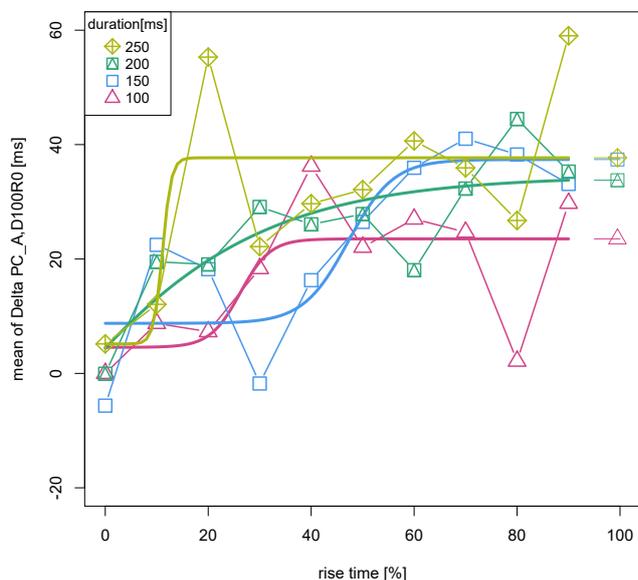
その結果、立ち上がり時間の割合の主効果 ( $F(4.06, 36.52) = 3.727, p = 0.012, \eta_p^2 = 0.293$ ) のみが5%水準で有意であった。持続時間的主効果、持続時間と立ち上がり時間の割合との交互作用は5%水準で有意でなかった。

立ち上がり時間の割合の主効果の下位検定として Shaffer の多重比較を行った。その結果、調整後  $p$  値が5%水準で有意であった組み合わせは、以下の一つのみであった。

- 立ち上がり時間の割合0%と立ち上がり時間の割合80% ( $t(9) = 4.781, adj.p = 0.045$ )



(a) 参加者 7 人のグループ



(b) 参加者 3 人のグループ

図 6.2.9: 参加者グループごとの、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」と、それを最も近似する 4 パラメータロジスティック回帰モデルの予測線  
 横軸の“rise time [%]”は立ち上がり時間の割合を表し、縦軸の“mean of Delta PC\_A,D100R0”は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  の変化 (縦軸) が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

表 6.2.3: 実験 2 の分散分析表

	SS	df	MS	F-ratio	p-value	$\eta_p^2$
参加者	169 720.327	9	18 857.814			
持続時間	4526.920	3	1508.973	1.779	0.175	0.165
参加者 × 持続時間	22 908.744	27	848.472			
立ち上がり時間の割合	11 328.613	4.060	2791.679	3.727	0.012	0.293
参加者 × 立ち上がり時間の割合	27 354.538	36.520	748.989			
持続時間 × 立ち上がり時間の割合	12 620.707	27	467.434	1.276	0.170	0.124
参加者 × 持続時間 × 立ち上がり時間の割合	88 986.811	243	366.201			
Total	337 446.659	399	845.731			

## 6.3 考察

以下では、分析の流れに沿って、考察を行い、仮説の検証を行う。加えて、実験1の結果と比較することで、音楽訓練の有無がPATにどのように影響するかも検討する。

### 6.3.1 生データの確認

まず、実験1と同様に、各参加者の各刺激音に対する同調のズレ量のバイオリンプロットより、殆どの参加者の殆どの刺激音の同調のズレ量の分布が単峰分布となった。この結果は先行研究と一致するものであった。

また、実験1と同様に、多くの刺激音の分布は両端が細くなる形を示したが、たまに、長方形に近い形や左右非対称の形もあった。このようなPATの分布の変化が何に起因するかは今後の研究が俟たれる。

そして、同調のズレ量の平均値と、同調のズレ量の標準偏差からは、実験1以上に個人差が大きいことが確認された。これも先行研究と一致する結果であった。

### 6.3.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

全参加者平均した $\Delta PC_{A,D100R0}$ を用いた分析より、持続時間が100msと150ms以上とで「立ち上がり時間の割合の増加と $PAT(\Delta PC_{A,D100R0})$ の移動との関係」が異なる可能性が示唆された。だが、そう言い切ることはできず、むしろ、全ての持続時間で「立ち上がり時間の割合の増加と $PAT(\Delta PC_{A,D100R0})$ の移動との関係」は同じであるとも考えられた。

そこで、参加者ごとに、持続時間ごとの「立ち上がり時間の割合の増加と $PAT(\Delta PC_{A,D100R0})$ の移動との関係」が同じであることを確認するため、各参加者の、各刺激音における $\Delta PC_{A,D100R0}$ を用いて、参加者をグループ分けをした。その結果、参加者は2つのグループに分けられた。

一つ目のグループは7人の参加者が属した。このグループでは、立ち上がり時間の割合が70%以降でのみ、立ち上がり時間の割合が増加するにつれ $PAT(\Delta PC_{A,D100R0})$ が僅かに時間的後方へ移動した。だが、立ち上がり時間の割合が0-70%では、立ち上がり時間の割合が増加しても $PAT(\Delta PC_{A,D100R0})$ は移動しなかった。この結果は実験1の音楽群の結果と全く異なった。

一方、もう一つのグループは3人の参加者が属した。このグループではどの持続時間でも、すなわち、持続時間が100ms以上において、立ち上がり時間の割合が増加するにつれ、 $PAT(\Delta PC_{A,D100R0})$ が急激に時間的後方へ移動した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれてPATの移動量が徐々に小さくなるか、あるいは移動しなくなることも明らかであった。回帰モデルの予測線の形としては、S字の上の方と、S字曲線全体の形であった。この傾向は、まさに、実験1で示された、統合の時間窓の範囲外と仮定さ

れた持続時間が200 ms以上での傾向と一致した。強いて違いを言うならば、この実験2のみで、ロジスティック回帰モデルの予測線の形に、S字曲線全体の形が見られた。

本論では、統合の時間窓を直接測定していない。実験1では統合の時間窓の長さが150–200 msの間のどこかにある仮定された。そして、実験1における統合の時間窓の範囲外の傾向が、実験1よりも統合の時間窓が短い場合として行われた実験2では、より短い持続時間から見られた。以上のことより、PAT知覚に統合の時間窓が影響していると考えることができ、本論の仮説が支持されたといえる。だがこれは3人の参加者を平均した結果であるため、この結果が不安定である可能性も考えられる。従って、実験3の結果より、同様のことが示されるかを確認する必要がある。

### 6.3.3 分散分析

$\Delta PC_{A,D100R0}$  を従属変数とし、持続時間、立ち上がり時間の割合を独立変数として、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内分散分析を行った。その結果、立ち上がり時間の割合の主効果のみ5%水準で有意であった。また、立ち上がり時間の割合の主効果の下位検定において、立ち上がり時間の割合0%と立ち上がり時間の割合80%の間のみ有意であることが明らかとなった。

先述した通り、全参加者を平均した  $\Delta PC_{A,D100R0}$  を用いて分析した結果、どの持続時間でも、立ち上がり時間の割合が70%程度以降でのみ立ち上がり時間の割合が増加すると  $\Delta PC_{A,D100R0}$  が僅かに増加するが、それ以前では  $\Delta PC_{A,D100R0}$  が変化しない傾向が見られた。すなわち、この分散分析の結果は、全参加者平均をした場合には、持続時間と立ち上がり時間の割合との交互作用が見られず、立ち上がり時間の割合が大きいときに  $\Delta PC_{A,D100R0}$  が僅かに増加したことを反映したと考えられた。

先行研究の分散分析の結果と本論の結果との比較は、考察(8.1節)にて行う。

## 6.4 結論

本論の仮説である、「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加とPATの移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことを検証するため、持続時間と立ち上がり時間の割合を変化させた刺激音へのタッピング実験を、音楽訓練を受けたことがない参加者(一般群の参加者)に対して行った。その結果、以下のことが明らかとなった。

- 一般群の参加者の大半が示した、持続時間ごとの「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」は、実験1の音楽群と全く異なるものであった。

- 具体的には、立ち上がり時間の割合が 70% 以降のみ立ち上がり時間の割合が増加すると  $PAT(\Delta PC_{A,D100R0})$  が僅かに時間的後方へ移動するが、それ以前 (立ち上がり時間の割合が 0-70%) では立ち上がり時間の割合が増加しても  $PAT(\Delta PC_{A,D100R0})$  が移動しなかった。
- 参加者 10 人中 7 人がこの傾向を示した。
- 一般群の参加者の一部が示した、持続時間が 100 ms 以上での「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,D100R0})$  との関係」は、実験 1 の音楽群の持続時間 200 ms 以上の場合の関係と同様であった。

実験 1 における統合の時間窓の範囲外の傾向が、実験 1 よりも統合の時間窓が短い場合として行われた実験 2 では、実験 1 より短い持続時間以上で見られた。このことより、本論の仮説が支持されたといえる。

- 具体的には、持続時間 100 ms 以上において、立ち上がり時間の割合が増加するにつれ、 $PAT(\Delta PC_{A,D100R0})$  が急激に時間的後方へ移動した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれての  $PAT$  の移動量が徐々に小さくなるか、あるいは移動しなくなった。回帰モデルの予測線の形としては、S 字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) と、S 字曲線全体の形 (独立変数の変化に対する従属変数の変化量が、徐々に大きくなり、次第に変化量が一定になり、そして変化量が徐々に小さくなる、S 字の全体) であった。
- 参加者 10 人中 3 人がこの傾向を示した

以上より、本論の仮説が確からしいことが確認された。だが、この実験 2 で仮説が確からしいと導かれたのは、3 人の参加者を平均した結果である。そのため、結果が不安定である可能性も考えられる。従って、次章にて、統合の時間窓が異なる場合として、周波数を取り上げる。統合の時間窓が短いとされる、周波数が高い刺激音への  $PAT$  を測定する。そして、実験 1 やこの実験 2 と比較を行い、同様のことが示されるかを確認し、更なる仮説の検証を行う。

# 第7章 本実験3: 統合の時間窓の長さが異なる場合(異なる周波数)における、持続時間や立ち上がり時間の割合の増加とPATの変化との関係と、統合の時間窓

本論では統合の時間窓を直接測定せず、統合の時間窓の長さが異なる時のPATと比較することでPATの知覚に統合の時間窓が関係しているのかを確認する。

本章の実験3では、実験1と統合の時間窓の長さが異なる場合として、周波数を取り上げた。実験1よりも統合の時間窓が短いと想定される、周波数がより高い刺激音を用いて、音楽群の参加者に対しタッピング課題を行った。

## 7.1 実験方法

### 7.1.1 参加者

3.3節の予備実験3での検討に従い、参加者は8人であった。このうちの7人は、実験1にも参加していた。参加者はみな、東京藝術大学音楽学部音楽環境創造科、並びに同大学院音楽音響創造研究分野に所属しており、音楽訓練を受けたことがあった。音楽訓練を受けたことがある参加者の定義は実験1と同様であった。

平均年齢は23.9歳であった。音響経験(録音やPAなど)の平均年数が2.8年であった。演奏経験の平均年数が11.4年、そのうち師についていた平均年数が9.0年であった。演奏本番経験のある人(今までに1回でもコンクールや演奏発表会などに出たことがある)が8人中4人であった。演奏楽器はピアノ、箏、ヴァイオリン、コントラバス、サックス、ギター、トロンボーンであり、合唱や声楽を行う参加者もいた。音楽聴取は多くの人が一日に1時間程度BGMとして聴いていた。自発的にはあまり聴かない人もいたが、その人たちも仕事や練習では聴いたり演奏したりしていた。

### 7.1.2 刺激音

サイン波を用いた。周波数は 3520 Hz とした。これは実験 1 で用いた 440 Hz の 3 オクターブ上であり、次の三つを考慮し決定された。一つ目がピアノの音域内であること、二つ目が PAT の先行研究の一つであり本論と条件が近い Danielsen ら [5] と London ら [15] が周波数の一つの水準に 3000 Hz を用いたこと、そして三つ目が統合の時間窓の先行研究において、周波数が異なる場合の臨界継続時間を検討した Plomp ら [21] と萩原 [43] が周波数の一つの水準に 4000 Hz を用いたことであった。

持続時間は 4 水準 (100, 150, 200, 250 ms) とし、立ち上がり時間の割合は 10 水準 (0, 10, 20, 30, 40, 50, 60, 70, 80, 90 %) とした。従って、計 40 個の刺激音が作られた。これは実験 2 と同様であった。

立ち上がり部分と立ち下がり部分の形状、立ち下がり時間、そして生成方法は本論共通のもの (4 章参照) を使用した。

### 7.1.3 実験手続き

実験手続きは、本論共通のもの (4 章) を使用した。以下では、本番の試行の構成と提示音圧レベルについて述べる。

本番では、同一刺激音を 700 ms (約 86 BPM) ごとに 80 回繰り返した音列を使用した。一つの刺激音に対し一つの音列が用意されたため、全 40 音列であった。これらの全 40 音列より、ランダムに 8 音列を連結を連結したものを 1 試行とし、全 5 試行が参加者に提示された。1 試行内は休憩を挟まず続けさせ、必要に応じて試行間に休憩を挟んだ。実験所要時間は約 1 時間であった。

提示音圧レベルは実験 1 や実験 2 と同様に 76 dB とした。確かに、周波数を変化させるとラウドネスも変化して知覚されることが知られているため、この実験 3 で周波数を 440 Hz から 3520 Hz へ変更する際に、提示音圧レベルを変更することも考えられた。だが、3.2 節の予備実験 2 にて、周波数が 440 Hz の時に提示音圧レベルを 76 dB から 84 dB に変化させても、得られた結果に有意差がないことが明らかとなった。この時の提示音圧レベルより等ラウドネス曲線を用いて phon 値を概算すると (図 7.1.1 参照)、約 75 phon から約 82 phon に変化させた時に有意差がないことになる。一方、実験 1 と実験 2 における周波数が 440 Hz で提示音圧レベルが 76 dB の場合の phon 値を概算すると約 75 phon であり、もし実験 3 において周波数が 3520 Hz で提示音圧レベルを 76 dB のまま行えば、phon 値は約 79 phon となる。従って、これらの条件でも有意差がない可能性が高いと推察され、実験 1 と実験 2 と同様の提示音圧レベルを用いることにした。

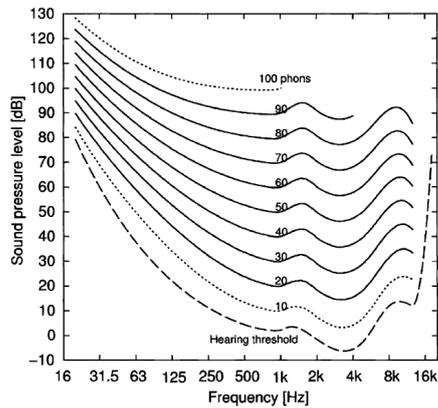


FIG. 11. Estimated equal-loudness-level contours drawn by the equation (6). The dashed line shows the threshold of hearing shown in Fig. 6. The contour at 100 phons is drawn by a dotted line because data from only one institute are available at 100 phons. The contour at 10 phons is also drawn by a dotted line because of the lack of experimental data between 20 phons and the hearing thresholds.

(a) 原文の図

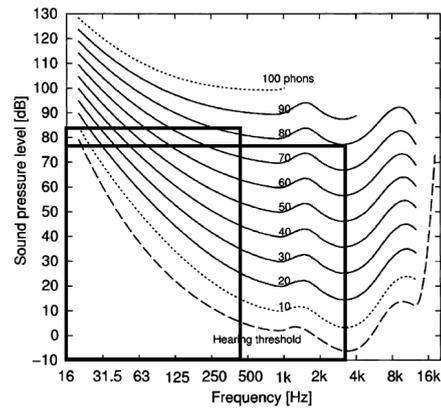


FIG. 11. Estimated equal-loudness-level contours drawn by the equation (6). The dashed line shows the threshold of hearing shown in Fig. 6. The contour at 100 phons is drawn by a dotted line because data from only one institute are available at 100 phons. The contour at 10 phons is also drawn by a dotted line because of the lack of experimental data between 20 phons and the hearing thresholds.

(b) 本論の条件時の線を重ねた図

図 7.1.1: 等ラウドネス曲線 (ISO 226:2003) に重ねた、各条件の時の線

ISO 226:2003 の元となった、Suzuki & Takeshima の論文より引用 [29] した。それに各条件の時の線を重ねた

### 7.1.4 実験装置

実験装置は、本論共通のもの (4 章) を使用した。

### 7.1.5 実験環境と日程

実験は、常時も突発的にも大きな音がしない、静かな部屋で、2021 年 1 月 27 日から 2 月 3 日までの間に一人ずつ行われた。

## 7.2 分析と結果

仮説の検証のため、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを分析した。

主な分析の流れは本論共通の手順 (4 章) にて記載した通りである。

### 7.2.1 生データの確認およびデータ処理

生データの確認を行い、実験の意図に沿わないものや、他と大きく異なるデータがないかを確認した。

#### 同調のズレ量のバイオリンプロット

参加者ごとに各刺激音  $A$  に対する 60 個の同調のズレ量  $Asy(A, i)$  を算出し、これらをバイオリンプロットで示した (付録 C)。バイオリンプロットが何かについては 5.2 節を参照してほしい。

これらのバイオリンプロットより、実験1と実験2と同様に、殆どの参加者の殆どの刺激音の同調のズレ量の分布は単峰分布であることが確認された。多くの刺激音の分布は、両端が細くなる(正規分布のような)形を示した。だが、たまに、長方形に近い(尖度が負のような)形や、左右非対称の(歪度の絶対値が大きいような)形もあった。例えば、長方形に近い形であったのは、参加者Cの刺激音40や参加者Zの刺激音35などであった。また、左右非対称の形であったのは、参加者Aの刺激音15や刺激音31などであった。

また、実験1と実験2と同様に、この実験3でも、ごくまれに、ある刺激音の同調のズレ量の分布の端が非常に長かった。それらの端の同調のズレ量の値は、 $\pm 200$  ms 以上であり、タッピングのミスより偶発的に得られた可能性が高いと考えられた。加えて、実験2と同様に、参加者Zに限って、ある刺激音のバイオリンプロットにおいて同調のズレ量が  $+200$ – $+300$  ms の間でも太くなることがあった(参加者Zの刺激音34と刺激音39)。これは実験2の参加者Lよりは細かった。これもやはり裏拍にタッピングしてしまったことによると推察された。

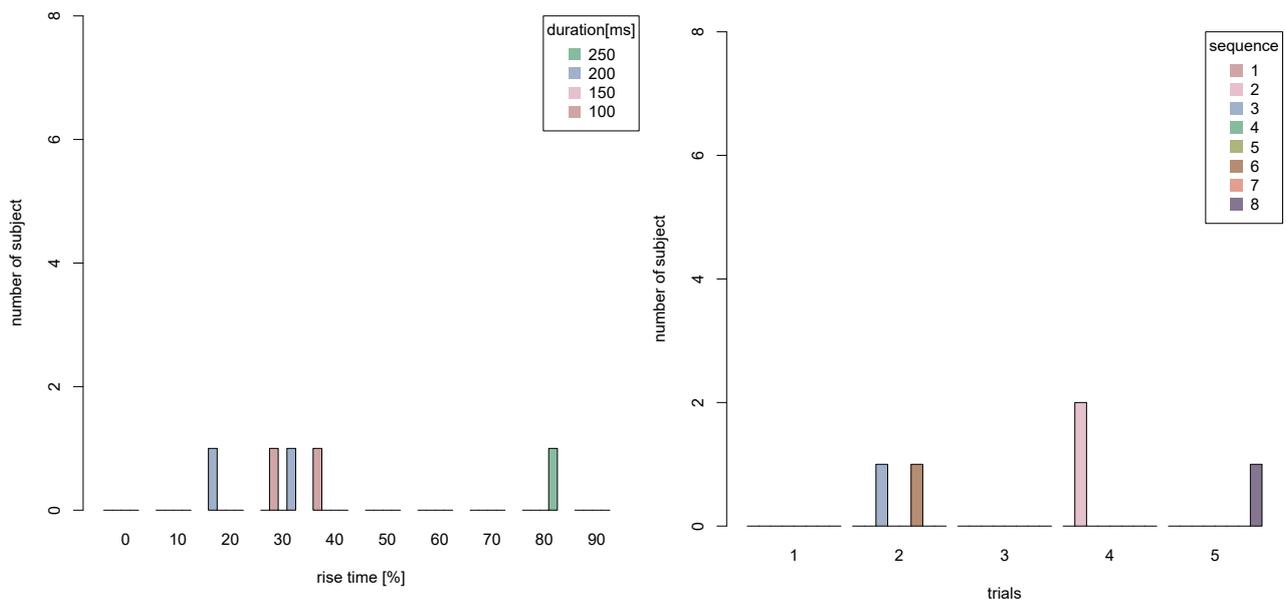
タッピングのミスであれ、裏拍にタッピングしてしまったのであれ、この  $\pm 200$  ms 以上の同調のズレ量のタッピングが、多くの参加者に共通して、特定の刺激音や特定の提示順にあるかを確認した。実験1と実験2と同様に、 $\pm 200$  ms 以上の同調のズレ量であったタッピングが属する刺激音と提示順を調べ、このタッピングが一つでもあった参加者の数を刺激音ごとにあるいは提示順ごとに数え、棒グラフで表したものが図7.2.1である。図の詳しい見方はキャプションを参照してほしい。これらの図より、 $\pm 200$  ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数は最大で2人であることが明らかであった(この実験3の参加者は全8人であった)。

以上のことより、この実験3でも、極めてまれに  $\pm 200$  ms 以上の同調のズレ量のタッピングが確認されたが、これらは刺激音や提示順に依存しないことが確認され、偶発的に得られたタッピングであると考えられた。これらのタッピングには、集中力の低下やタッピングの際の手指の動作ミスに起因するものもあるが、誤って裏拍にタッピングしてしまったことに起因するものもあると推察された。いずれにせよ、このようなデータは実験意図に沿わないものであるため、本論では、 $\pm 200$  ms 以上の同調のズレ量のタッピングは削除した。

### 参加者ごとの刺激音に対する同調のズレ量の平均値の変化量

参加者ごとに、各刺激音Aに対する同調のズレ量の平均値を算出し、縦に並べて図7.2.2に示した。この図では、各刺激音に対する同調のズレ量の平均値の変化量と変化範囲を知ることができる。図の詳しい見方はキャプションを参照してほしい。前者の変化量は数値として表7.2.1にも示した。

まず、表7.2.1より、各刺激音に対する同調のズレ量の平均値の変化量が、参加者によって異なることが明らかである。参加者Cの変化量が最も小さく、28 ms であった。一方、参加者



(a) 刺激音ごと

横軸の“rise time[%]”は刺激音の立ち上がり時間の割合を表し、縦軸は参加者の数を表し、凡例の“duration[ms]”は刺激音の持続時間を表す。各棒グラフは±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。横軸は刺激音の立ち上がり時間の割合であり、それぞれの立ち上がり時間の割合の箇所には、左から右へ、持続時間が短い刺激音から長い刺激音に対する棒グラフが並んでいる。

(b) 提示順ごと

横軸の“trials”は試行を表し、凡例の“sequence”は音列を表す。各棒グラフは±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数を表す。実験で提示された試行が横軸に示され、各試行の箇所には、左から右へ、各試行内で提示された順の各音列に対する棒グラフが並んでいる。すなわち、左から右へと、提示された時系列で並んでいる。

図 7.2.1: ±200 ms 以上の同調のズレ量のタッピングが一つでもあった参加者の数

E の変化量が最も大きく、150 ms であった。確かに、図 6.2.2 より、参加者 E の縦方向への拡がり具合が、他の参加者と大きく異なることが明らかである。参加者 E を除き、他の参加者は、縦方向への拡がり具合が類似していた。参加者 C のみ縦方向への拡がり具合が少し狭いが、他の参加者と全く異なるとはいえない。また、参加者 E を除き、他の参加者は、白丸のシンボルが拡がる範囲が類似していた。

以上のことより、参加者 E のみ、各刺激音に対する同調のズレ量の平均値の変化量が非常に大きいことが明らかになった。参加者 E は実験 1 にも参加したが、実験 1 ではこのような反応が見られなかった。加えて、前項のバイオリンプロットでも、参加者 E が他の参加者と異なる傾向は見られなかった。従って、現段階では、参加者 E の変化量の大きさが実験 1 よりも周波数が高い刺激に対する反応であるのか、あるいはタッピング課題の遂行に問題があったのかを判断することができない。そのため、ここでは参加者 E のデータを削除するかの判断を保留し、次項の標準偏差の分析後に判断した。

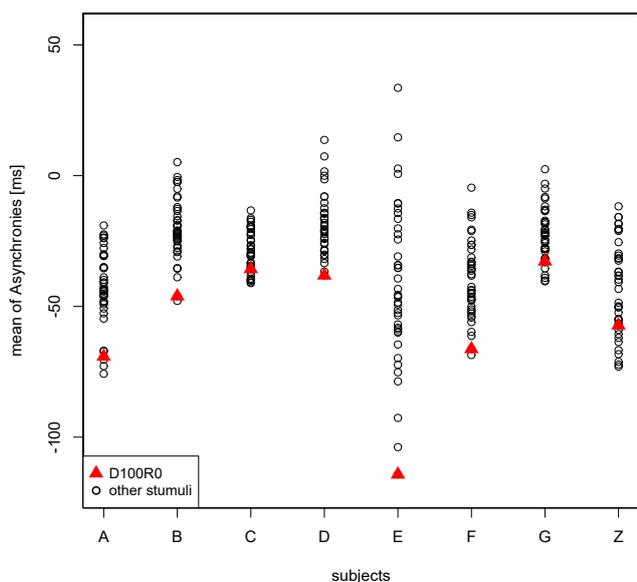


図 7.2.2: 参加者ごとの、各刺激音に対する同調のズレ量の平均値

図の横軸の“subjects”は参加者番号を表し、縦軸の“mean of asynchronies[ms]”は同調のズレ量の平均値を表す。白丸のシンボルが、各刺激音 A に対する同調のズレ量の平均値を、赤色塗りつぶしの三角のシンボルが、 $\Delta PC_{A,D100R0}$  の基準の刺激音である D100R0 音に対する同調のズレ量の平均値を示す。

表 7.2.1: 参加者ごとの、各刺激音に対する同調のズレ量の平均値の変化量 (最大値 – 最小値)

参加者	変動量 [ms]
A	57
B	53
C	28
D	52
E	148
F	64
G	43
Z	61

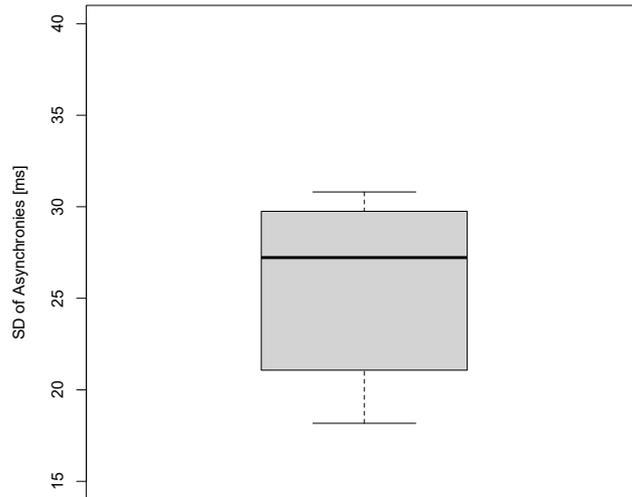


図 7.2.3: 参加者ごとの、全刺激音で平均した同調のズレ量の標準偏差の箱ひげ図

縦軸の“SD of Asynchronies[ms]”が同調のズレ量の標準偏差を表す。この箱ひげ図では、中央値が箱の中の太線で表され、四分位範囲(第1四分位から第3四分位の差分)すなわちデータのばらつきが箱の大きさで表される。また、外れ値は「第1四分位・第3四分位から四分位範囲の1.5倍以上離れた値」と定義され、白い丸のシンボルで表される。そして、外れ値を除いた最小値と最大値がひげの端で表される。

### 同調のズレ量の標準偏差の確認

まず、参加者ごとに、全刺激音で平均した、同調のズレ量の標準偏差を算出し、箱ひげ図として図 7.2.3 に示した。この箱ひげ図の見方はキャプションを参照してほしい。この箱ひげ図より、外れ値がないことが明らかである。すなわち、全刺激音で平均した標準偏差が他の参加者と大きく異なる参加者がいなかったことが確認された。

前項にて、参加者 E のみ、各刺激音に対する同調のズレ量の平均値の変化量が非常に大きかったことが確認された。だが、参加者 E の標準偏差は外れ値にはならなかった。すなわち、他の参加者と大きく異なることが明らかになった。以上のことから、参加者 E のタッピング課題の遂行に問題はなく、前項で確認された変化量の大きさも個人差の範囲内であると判断し、データの削除は行わなかった。

次に、刺激音ごとに、全参加者で平均した、同調のズレ量の標準偏差を算出し、棒グラフとして図 7.2.4 に示した。この図の見方はキャプションを参照してほしい。

この図より、実験 1 と同様に、全参加者で平均した同調のズレ量の標準偏差は、刺激音によってばらつくが、持続時間や立ち上がり時間の割合が増加するにつれて標準偏差も増加するという傾向は見られなかった。

また、音楽訓練を受けた人の SMS 課題における標準偏差は IOI の 2-4% であると言われる。

実際、実験1の音楽群における、刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差は、値にして24-33msであり、割合にしてIOI(700ms)の3.4-4.7%であった。そして、この実験3の音楽群における、同調のズレ量の標準偏差は、値にして23-30msであり、割合にしてIOIの3.3-4.3%であった。従って、先行研究と比較すると近い値が得られたといえる。加えて、音楽群に対し行った実験1と実験3の標準偏差を比較すると、この実験3の方が僅かに小さいことが確認された。

以上のように生データの確認を行った。同調のズレ量の中に、偶発的であつ他と大きく異なるデータや実験意図に沿わないデータが含まれていたため、それらのみ削除した。参加者Eの各刺激音に対する同調のズレ量の平均値の変化量が他の参加者よりも顕著に大きかったが、参加者Eの全刺激音で平均した標準偏差が他の参加者と大きく異ならなかったため、データの削除は行わなかった。その他に、大きく異なるデータは見られなかった。

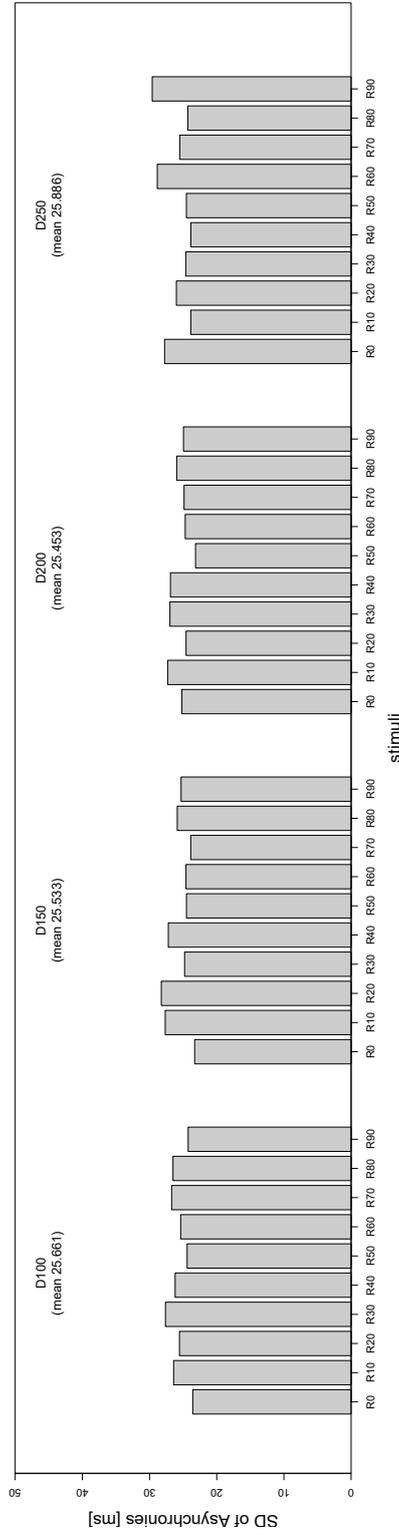


図 7.2.4: 刺激音ごとの、全参加者で平均した同調のズレ量の標準偏差の棒グラフ

全 40 個の刺激音 (持続時間 4 水準 × 立ち上がり時間の割合 10 水準) に対する、全参加者で平均した同調のズレ量の標準偏差が棒グラフで表されている。それらは、10 本ごとに一つのまとまりになっており、一つのまとまりの 10 本は同じ持続時間の刺激音である。10 本の棒グラフの上方に、その持続時間の値が“D + 値”で略記された。合わせて、その下に、その 10 個の刺激音で平均をとった、全参加者で平均した同調のズレ量の平均値の標準偏差の値も示された。その 10 本の棒グラフは、それぞれ、左から順に、立ち上がり時間の割合が小さい刺激音から大きい刺激音に対する値を表す棒グラフが並んでいる。横軸に立ち上がり時間の割合の値が“R + 値”と略記された。

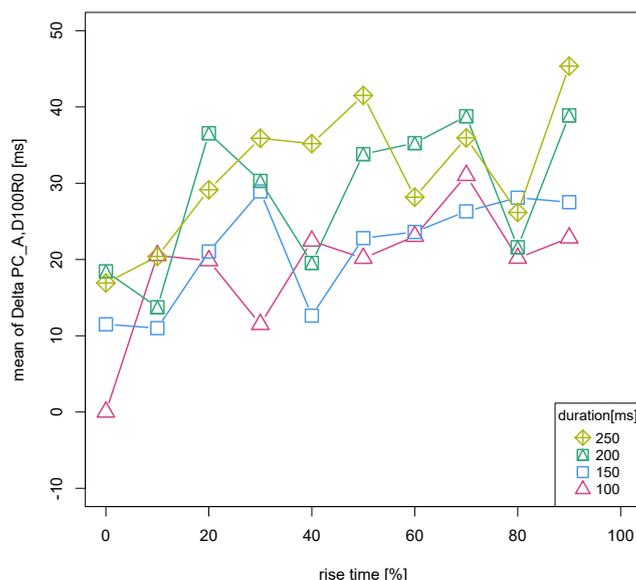


図 7.2.5: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」  
 横軸の“rise time [%]”は刺激音の立ち上がり時間の割合を表し、縦軸の“mean of Delta PC\_A,D100R0”  
 は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の立ち上がり時間  
 の割合を増加させた時の (横軸)、PAT (縦軸) が示されている。

## 7.2.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

仮説の検証のため、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるかを分析した。

まず、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を図に示した。それが図 7.2.5 である。

次に、各持続時間におけるこの関係を最も近似する 4パラメータロジスティック回帰モデルを求め、求められた回帰モデルの予測線を先ほどの図に重ねた。それが図 7.2.6 である。この図より、どの持続時間でも、立ち上がり時間の割合が増加するにつれ、 $\Delta PC_{A,D100R0}$  が急激に増加した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれて  $\Delta PC_{A,D100R0}$  の増加量が徐々に小さくなるか、あるいは  $\Delta PC_{A,D100R0}$  が増加しなくなる事が明らかであった。ロジスティック回帰モデルの予測線の形としては、S 字の上の方と、S 字曲線全体の形であった。加えて、持続時間が増加するにつれ、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」は大きく変化しないまま、 $\Delta PC_{A,D100R0}$  が増加する傾向もみられた。

そして、4パラメータロジスティック回帰モデルにおける、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を強く反映する 2つの指標を算出した。その値を表 7.2.2 に示し

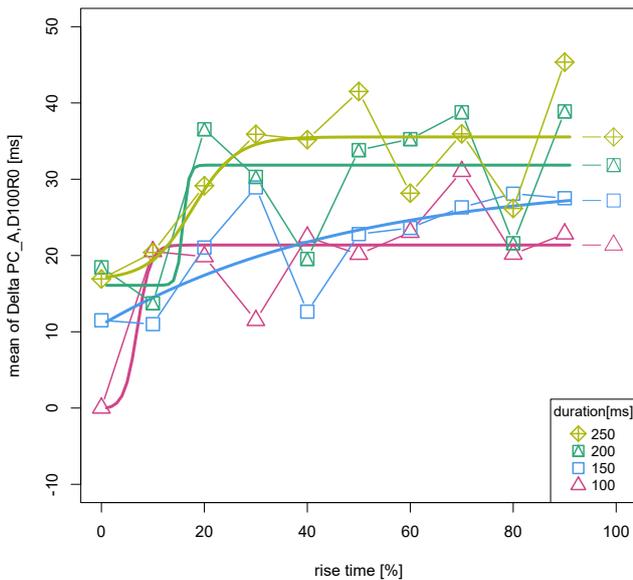


図 7.2.6: 持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」と、それを最も近似する4パラメータロジスティック回帰モデルの予測線  
 横軸の“rise time [%]”は立ち上がり時間の割合を表し、縦軸の“mean of Delta  $PC_{A,D100R0}$ ”は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  の変化 (縦軸) が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

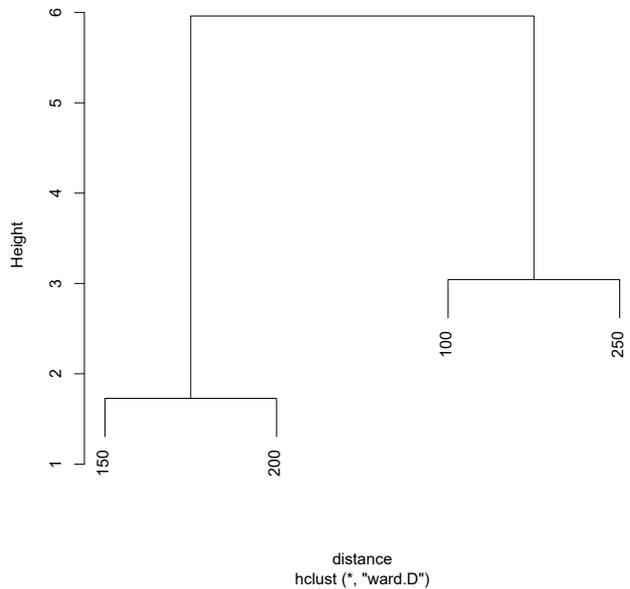


図 7.2.7: 近似した4パラメータロジスティック回帰モデルにおける2つの指標を用いたクラスター分析の結果

た。これらの2つの指標の値を用いてウォード法によるクラスター分析を行い、持続時間を分類した。その結果が、図7.2.7である。この図より、持続時間が、(1)150 msと200 ms、(2)100 msと250 msのグループに分けられると示された。すなわち、この実験3では、持続時間が短いグループと持続時間が長いグループとに分類されず、実験1で見られた「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なることが確認されなかった。

そこで、まず、グループごとの2つの指標の値を比較した。「パラメータの一つである傾き  $b$  の値」は持続時間グループ(1)と持続時間グループ(2)とでばらばらであり、一方、「立ち上がり時間の割合が0%の時の予測値 (PAT) と立ち上がり時間の割合が100%の時の予測値 (PAT) との差分」はグループ(2)よりもグループ(1)の方が小さかった。そのため、「立ち上がり時間の割合が0%の時の予測値 (PAT) と立ち上がり時間の割合が100%の時の予測値 (PAT) との差分」がより小さいグループと、それがより大きいグループに分けられる可能性も示唆された。

表 7.2.2: 近似した4パラメータロジスティック回帰モデルにおける2つの指標の値

持続時間 [ms]	パラメータの一つである傾き $b$ の値	立ち上がり時間の割合が0%の時の予測値 (PAT) と立ち上がり時間の割合が100%の時の予測値 (PAT) との差分
100	-0.801	21.365
150	-0.023	15.911
200	-1.743	15.754
250	-0.235	18.375

表 7.2.3: 実験3の分散分析表

	SS	df	MS	F-ratio	p-value	$\eta_p^2$
参加者	116 677.378	7	16 668.197			
持続時間	8236.463	3	2745.488	8.038	0.001	0.535
参加者 × 持続時間	7173.258	21	341.584			
立ち上がり時間の割合	13 728.459	9	1525.384	5.938	0.000	0.459
参加者 × 立ち上がり時間の割合	16 183.023	63	256.873			
持続時間 × 立ち上がり時間の割合	6995.328	27	259.086	1.573	0.043	0.184
参加者 × 持続時間 × 立ち上がり時間の割合	31 132.731	189	164.723			
計	200 126.639	319	627.356			

### 7.2.3 分散分析

先行研究との比較のため、 $\Delta PC_{A,D100R0}$  を従属変数とし、持続時間、立ち上がり時間の割合を独立変数とした、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内計画の分散分析を行った (表 7.2.3)。

その結果、持続時間と立ち上がり時間の割合との交互作用 ( $F(9, 54) = 1.573, p = 0.043, \eta^2 = 0.184$ ) が5%水準で有意であった。交互作用が有意であったため、主効果はあまり意味をなさないが、先行研究との比較のために主効果も報告する。立ち上がり時間の主効果 ( $F(9, 54) = 5.938, p = 0.000, \eta^2 = 0.459$ )、持続時間の割合の主効果 ( $F(3, 21) = 8.038, p = 0.001, \eta^2 = 0.535$ ) が5%水準で有意であった。

交互作用の下位検定の結果、有意な組み合わせがあった。そのため、さらに下位検定として Shaffer の多重比較を行った。その結果、調整後  $p$  値が5%水準で有意であった組み合わせは、持続時間と立ち上がり時間の割合の各水準との間では、以下の1つであり、

- 立ち上がり時間の割合40%の時の、持続時間100msと250msの間 ( $t(7) = 3.802, adj.p = 0.040$ )

立ち上がり時間の割合と持続時間の各水準との間では、以下の1つであった。

- 持続時間250msの時の、立ち上がり時間の割合0%と40%の間 ( $t(7) = 5.649, adj.p = 0.035$ )

## 7.3 考察

以下では、分析の流れに沿って、考察を行い、仮説の検証を行う。加えて、実験1や実験2の結果と比較することで、周波数の違いがPATにどのように影響するかも検討する。

### 7.3.1 生データの確認

まず、実験1と実験2と同様に、各参加者の各刺激音に対する同調のズレ量のバイオリンプロットより、殆どの参加者の殆どの刺激音の同調のズレ量の分布が単峰分布であることが明らかとなった。この結果は先行研究と一致するものであった。

また、多くの刺激音の分布は両端が細くなる形を示したが、たまに、長方形に近い形や左右非対称の形もあった。このようなPATの分布の変化が何に起因するかは今後の研究が俟たれる。

そして、実験1と同様に、同調のズレ量の平均値と、同調のズレ量の標準偏差からは、個人差が大きいことが確認された。これも先行研究と一致する結果であった。

### 7.3.2 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」

分析より、持続時間が(1)150 msと200 ms、(2)100 msと250 msによって、「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」が異なる可能性が示唆された。だが、そう言い切ることはできないと考えられた。なぜなら、一つ目に、グループごとの2つの指標の値より、「立ち上がり時間の割合が0%の時の予測値(PAT)と立ち上がり時間の割合が100%の時の予測値(PAT)との差分」のみでグループ分けされた可能性があったからであり、二つ目に、回帰モデルの予測線より、どの持続時間でも「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」が同じである可能性があったからだ。2.6章で述べたように、クラスター分析に用いた2つの指標は、4パラメータロジスティック回帰モデルにおいて、持続時間ごとの「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」の差異を強く反映する指標であった。以上のことを踏まえると、実験3のクラスター分析の結果が実験1や実験2と異なったのは、先述した、持続時間が増加するにつれ、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」は大きく変化しないまま、 $\Delta PC_{A,D100R0}$ が増加する傾向を反映したと考えることができる。むしろ、この実験3では全ての持続時間で「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」が同じであったと考えることが妥当であろう。この傾向は、まさに、実験1で示された統合の時間窓の範囲外(持続時間が200 ms以上)での傾向と一致し、かつ、実験2における統合の時間窓の範囲外(持続時間が100 ms以上)での傾向と一致した。本論では、統合の時間窓を直接測定していない。だが、実験1における統合の時間窓の範囲外の傾向が、実験1よりも統合の時間窓が短い場合として行われた実験2とこの実験3においては、実験1よりも

より短い持続時間以上で見られた。以上のことより、PAT 知覚に統合の時間窓が影響していると考えることができ、本論の仮説である「PAT の知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持されたといえる。

### 7.3.3 分散分析

$\Delta PC_{A,D100R0}$  を従属変数とし、持続時間、立ち上がり時間の割合を独立変数とした、Greenhouse-Geisser の  $p$  値補正付き二要因参加者内計画の分散分析を行った。その結果、持続時間と立ち上がり時間の割合との交互作用が 5% 水準で有意であったが、下位検定の結果、統計的有意となった組み合わせは 2 つのみであり、その他の殆どの組み合わせは有意ではないことが明らかとなった。

この実験 3 では、先述した通り、全ての持続時間が、同じ「立ち上がり時間の割合の増加と PAT( $\Delta PC_{A,D100R0}$ ) の移動との関係」であり、加えて、持続時間が長くなると、立ち上がり時間の割合によらず、PAT( $\Delta PC_{A,D100R0}$ ) が時間的後方へ移動するだろうと考えられた。これらのことより、分散分析にて交互作用が有意になったにも関わらず、殆どの組み合わせでは有意ではなかったことは、測定された PAT( $\Delta PC_{A,D100R0}$ ) では交互作用がないはずにもかかわらず、データ上のばらつきが大きいことが反映されたからである可能性が考えられた。

先行研究における分散分析の結果と本論の結果との比較は、考察 (8.1 節) にて行う。

## 7.4 結論

本論の仮説である、「PAT の知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことを検証するため、持続時間と立ち上がり時間の割合を変化させた、周波数がより高い刺激音を用いて、音楽群の参加者に対しタッピング課題を行った。その結果、以下のことが明らかとなった。

- 持続時間が 100 ms 以上での「立ち上がり時間の割合の増加と PAT( $\Delta PC_{A,D100R0}$ ) との関係」が、実験 1 の音楽群の持続時間 200 ms 以上の場合の関係と、また、実験 2 の一般群の持続時間が 100 ms 以上の場合の関係と一致した。

実験 1 の音楽群の持続時間が 200 ms 以上は、統合の時間窓の範囲外であると仮定された。実験 2 とこの実験 3 は、実験 1 よりも統合の時間窓が短い場合として行われた。この条件下で、実験 1 における統合の時間窓の範囲外の傾向が、実験 2 とこの実験 3 において、実験 1 より短い持続時間以上で見られたことより、本論の仮説である「PAT の知覚は統合

の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持されたと結論づける。

- より具体的には、持続時間が 100 ms 以上において、立ち上がり時間の割合が増加するにつれ、 $PAT(\Delta PC_{A,D100R0})$  が急激に時間的後方へ移動した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれて PAT の移動量が徐々に小さくなるか、あるいは PAT が移動しなくなった。回帰モデルの予測線の形としては、S 字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) と、S 字曲線全体の形 (独立変数の変化に対する従属変数の変化量が、徐々に大きくなり、次第に変化量が一定になり、そして変化量が徐々に小さくなる、S 字の全体) であった。
- 統合の時間窓の範囲外 (持続時間 100 ms 以上) において、持続時間が増加するにつれ、「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」は大きく変化しないまま、 $PAT(\Delta PC_{A,D100R0})$  が時間的後方へ移動する可能性が示された。これは周波数による差異の可能性も考えられた。だが、本論では、統合の時間窓の範囲外で持続時間を多くの水準で変化させておらず、周波数も実験 1 と実験 3 の 2 水準でしか変化させていないため、結論付けることはできない。

以上より、本論の仮説が支持された。次章にて考察を行い、仮説と PAT 知覚について更なる検討を行う。

## 第8章 全体考察

本論は、PAT 知覚を扱い、PAT が音の何の要素によってどのように移動するかを基礎研究として調査するものであった。PAT の先行研究と統合の時間窓の先行研究より明らかになっていることを照らし合わせ、「PAT の知覚も統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」と仮説だて、これを検証した。

本章では、前章までの実験結果を踏まえ、全体考察を行った。

まず、8.1 節にて、5 章から 7 章で行った 3 つの実験の結果を比較し、仮説の検討の結論を再度述べた。

次に、8.2 節にて、本論の実験で変化させた「立ち上がり時間の割合」ではなく、「立ち上がり時間」の場合での分析結果を検討した。すなわち、『「立ち上がり時間」の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係』が、今まで述べてきた『「立ち上がり時間の割合」の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係』と異なるかどうかを検討した。

続いて、8.3 節にて、持続時間を取り上げ、「持続時間の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」について検討した。合わせて、周波数が異なる時に  $PAT(\Delta PC_{A,D100R0})$  が移動するかについても検討した。

最後に、8.4 節にて、先行研究との比較を行った。

### 8.1 3 つの実験の結果と仮説の検討の結論

本節では、5 章から 7 章で行った 3 つの実験で用いた刺激音を再度述べる。続いて、3 つの実験の結果を再度述べ、明らかになったことをより明確にする。最後に、全体考察として、仮説の検証の結論を加える。

#### 3 つの実験の刺激音の再確認

まず、3 つの実験で用いた刺激音を再度確認する。刺激音はサイン波であった。周波数は実験 1 と実験 2 では 440 Hz で、実験 3 では 3520 Hz であった。全ての実験にて、持続時間と「立ち上がり時間の割合 (持続時間に占める立ち上がり時間の割合であり、数式では (立ち上がり

表 8.1.1: 3つの実験における持続時間と立ち上がり時間の割合

(a) 実験 1(440 Hz)

(b) 実験 2(440 Hz) と実験 3(3520 Hz)

	持続時間 [ms]				
	100	150	200	250	300
立ち上がり時間の割合 [%]	0	0	0	0	0
	10	10	10	10	10
	20	20	20	20	20
	30	30	30	30	30
	40	40	40	40	40
	50	50	50	50	50
	60	60	60	60	60
	70	70	70	70	70
	80	80	80	80	80
	90	90	90	90	90

	持続時間 [ms]				
	100	150	200	250	—
立ち上がり時間の割合 [%]	0	0	0	0	—
	10	10	10	10	—
	20	20	20	20	—
	30	30	30	30	—
	40	40	40	40	—
	50	50	50	50	—
	60	60	60	60	—
	70	70	70	70	—
	80	80	80	80	—
	90	90	90	90	—

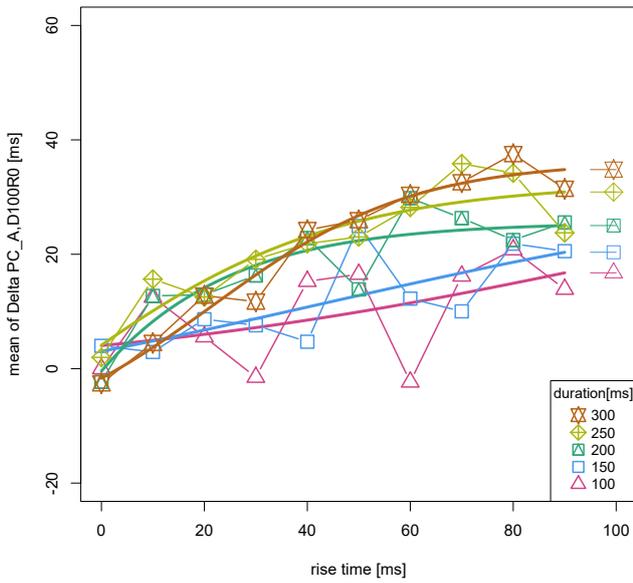
時間)/(持続時間)×100[%]である)」を変化させた。各実験で用いた刺激音の、持続時間と立ち上がり時間の割合を表 8.1.1 に示した。提示音圧レベルは 76 dB で一定であった。

実験 1 では、音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行い、PAT を測定した。この実験 1 の統合の時間窓の長さを基準とし、実験 2 と実験 3 では、実験 1 よりも統合の時間窓が短い場合の PAT を測定した。実験 2 では、実験 1 と同様の刺激音を用いて、音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行い、PAT を測定した。実験 3 では、周波数が高い刺激音を用いて、音楽群の参加者に対しタッピング課題を行い、PAT を測定した。

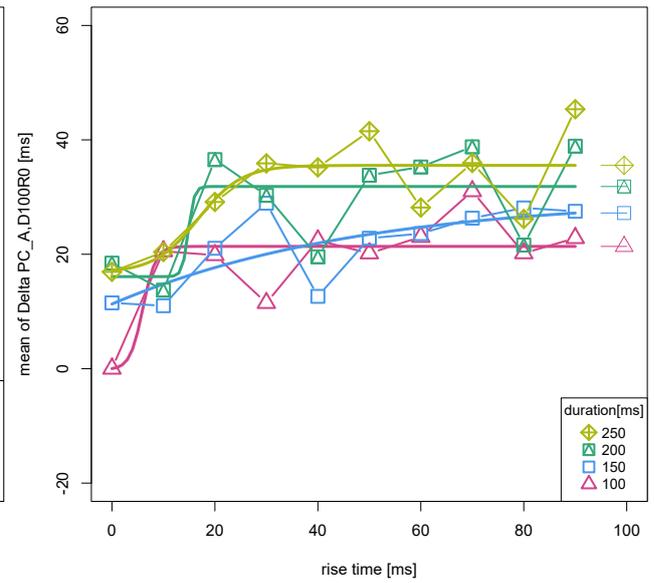
### 3つの実験の分析と考察の結果の再確認

続いて、3つの実験より、分析と考察の結果を再度述べる。また、各実験の章で示した、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」とそれを最も近似する 4パラメータロジスティック回帰モデルの予測線の図を図 8.1.1 に再掲した。本節では、実験 2 の全参加者のロジスティック曲線の近似の図は再掲せず、参加者グループごとのロジスティック曲線の近似の図のみを再掲した。

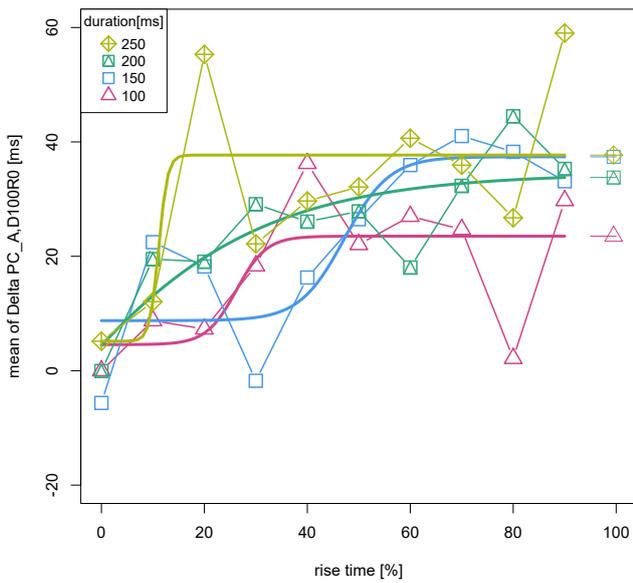
まず、実験 1 の分析より、「立ち上がり時間の割合の増加と PAT の変化との関係」が、持続時間 150 ms 以下と 200 ms 以上とで異なることが示された。より具体的には、持続時間が 150 ms 以下では、立ち上がり時間の割合が増加するにつれ、PAT はゆるやかに時間的後方へ移動した。その移動量は持続時間が 200 ms 以上の場合のその移動量よりも小さかった。一方、持続時間が 200 ms 以上では、立ち上がり時間の割合が増加するにつれ、PAT が急激に時間的後方へ移動した。その移動量は持続時間が 150 ms 以下の場合よりも大きかった。加えて、立ち上がり時間



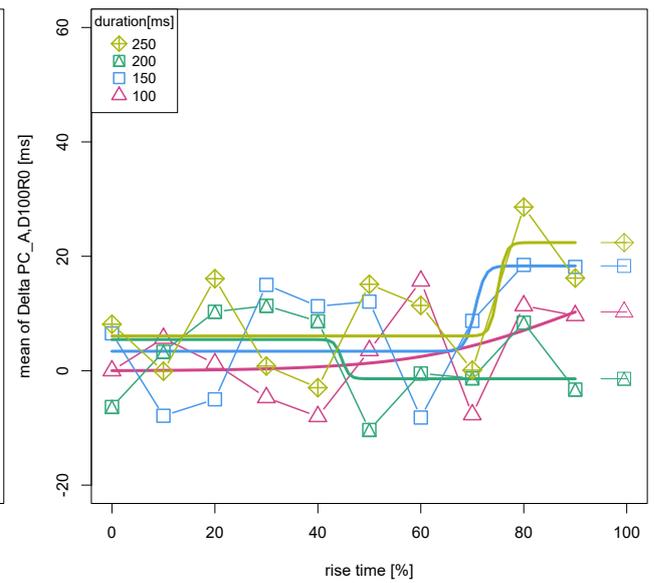
(a) 実験 1



(b) 実験 3



(c) 実験 2 の参加者 3 人のグループ



(d) 実験 2 の参加者 7 人のグループ

図 8.1.1: 【再掲】持続時間ごとの、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」と、それを最も近似する 4 パラメータロジスティック回帰モデルの予測線  
 横軸の “rise time [%]” は立ち上がり時間の割合を表し、縦軸の “mean of Delta PC\_A,D100R0” は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の「立ち上がり時間」を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  (縦軸) が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。ここでは、実験 2 の全参加者のロジスティック曲線の近似の図は再掲せず、参加者グループごとのロジスティック曲線の近似の図のみを再掲した。それに伴い、実験 1 と実験 3 の縦軸範囲を、実験 2 の参加者グループごとの図の範囲に変更した。

の割合が60–70%以上になると、立ち上がり時間の割合が増加するにつれてのPATの移動量が徐々に小さくなった。

この150–200 msは、統合の時間窓が存在するとされる時間範囲に重なる。刺激音の持続時間が統合の時間窓の範囲内から範囲外までに及ぶ以上、統合の時間窓の影響を受けないとは考えにくい。そこで、この音楽群の参加者の統合の時間窓の長さが150–200 msの間にあると考えると、本論の仮説が支持される可能性が示された。だが、統合の時間窓を直接測定していないため、本当に仮説が支持されるかどうかは、実験2と実験3にて統合の時間窓の長さが異なる時の刺激音のPATとの比較をして検討する必要がある。

実験2より、一般群の参加者は、持続時間ごとの「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」が異なる2つのグループに分けられることが明らかになった。一つ目のグループは、実験2の参加者10人中の7人が属した。すなわち、一般群の大半がこのグループに属した。一般群の大半は、立ち上がり時間の割合が70%以降でのみ、立ち上がり時間の割合が増加するとPATが僅かに時間的後方へ移動した。それ以前(立ち上がり時間の割合が0–70%)では、立ち上がり時間の割合が増加してもPATが移動しなかった。これは実験1の音楽群と全く異なる結果であった。二つ目のグループは、実験2の参加者10人中の3人が属した。彼らは、どの持続時間(持続時間100 ms以上)でも、立ち上がり時間の割合が増加するにつれ、PATが急激に時間的後方へ移動した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれてのPATの移動量が徐々に小さくなるか、あるいは移動しなくなった。これは、実験1の音楽群の持続時間200 ms以上での関係と同じであった。

実験1の音楽群の持続時間が200 ms以上は、統合の時間窓の範囲外であると仮定された。この実験2は、実験1よりも統合の時間窓が短い場合として行われた。その結果、二つ目の参加者グループでは、持続時間が100 ms以上での「立ち上がり時間の割合の増加とPATの移動との関係」が、実験1の音楽群の持続時間200 ms以上での関係と一致した。このことより、本論の仮説が支持されると考えられた。だが、この結果は10人中3人の参加者の結果を平均して得られたものであり、不安定である可能性も考えられた。

実験3より、周波数を高くした刺激音に対する、持続時間ごとの「立ち上がり時間の割合の増加とPAT( $\Delta PC_{A,D100R0}$ )の移動との関係」は、どの持続時間(100 ms以上)でも、立ち上がり時間の割合が増加するにつれ、PATが急激に時間的後方へ移動した。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれてPATの移動量が徐々に小さくなるか、あるいはPATが移動しなくなることが明らかであった。これは、実験1の音楽群の持続時間200 ms以上での関係と、また、実験2の一般群の持続時間が100 ms以上での関係と同じであった。

すなわち、実験3での持続時間が100 ms以上での「立ち上がり時間の割合の増加とPATとの関係」が、実験1の音楽群の持続時間200 ms以上での関係と、また、実験2の一般群の持続

時間が 100 ms 以上での関係と一致した。実験 2 と実験 3 は、実験 1 よりも統合の時間窓が短い場合として行われた。このことより、仮説が支持されると考えられた。

## 仮説の検証の結論

最後に、全体考察として、仮説の検証の結論を加える。以上の 3 つの実験の結果より、本論の仮説である「PAT の知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持されたと結論づける。

より具体的には、音の持続時間が統合の時間窓の範囲内 (実験 1 の持続時間 150 ms 以下) では、立ち上がり時間の割合が増加するにつれ、PAT はゆるやかに時間の後方へ移動する。その移動量は、音の持続時間が統合の時間窓の範囲外の場合のその移動量よりも小さい。ロジスティック回帰モデルの予測線の形としては、S 字曲線の下の方の形 (独立変数の変化に対する従属変数の変化量が徐々に大きくなる) か、あるいは S 字の真ん中の部分 (独立変数の変化に対する従属変数の変化量がほぼ一定であり直線に近い形となる部分) である。一方、音の持続時間が統合の時間窓の範囲外 (実験 1 での持続時間 200 ms 以上と、実験 2 と実験 3 での持続時間 100 ms 以上) では、立ち上がり時間の割合が増加するにつれ、PAT が急激に時間の後方へ移動する。その移動量は、音の持続時間が統合の時間窓の範囲内の場合の移動量よりも大きい。加えて、立ち上がり時間の割合がある一定以上になると、立ち上がり時間の割合が増加するにつれての PAT の移動量が徐々に小さくなるか、PAT が移動しなくなる。ロジスティック回帰モデルの予測線の形としては、S 字の上の方 (独立変数の変化に対する従属変数の変化量が徐々に小さくなる部分) か、あるいは S 字曲線全体の形 (独立変数の変化に対する従属変数の変化量が、徐々に大きくなり、次第に変化量が一定になり、そして変化量が徐々に小さくなる、S 字の全体の部分) である。

以上のことより、本論の仮説が支持され、より具体的には、「PAT 知覚も統合の時間窓の影響を受け、統合の時間窓の範囲内と範囲外では、立ち上がり時間の割合の増加に対する PAT の変化量と曲率が異なる」と結論付ける。

## 8.2 「立ち上がり時間の割合」ではなく「立ち上がり時間」の場合の分析結果

本論の実験では刺激音を「立ち上がり時間の割合」で変化させ、分析では『「立ち上がり時間の割合」の増加と  $\Delta PC_{A,D100R0}$  の変化』を検討し、仮説の検証を行ってきた。本節にて、「立ち上がり時間の割合」ではなく「立ち上がり時間」であった場合での分析結果を検討した。すなわち、『「立ち上がり時間」の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係』である場合にも、「立ち上がり時間の割合」より導いた結論が成り立つのかを検討した。

『「立ち上がり時間」の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係』を明らかにするため、5章から7章で行った分析と同様に分析を行った。なお、実験2に対しては、全参加者平均での分析は行わず、参加者グループごとでの分析のみを行った。

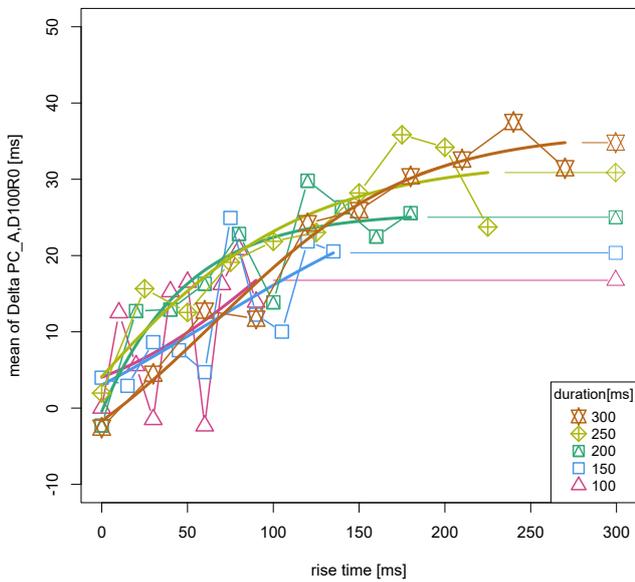
まず、実験で変化させた「立ち上がり時間の割合」を「立ち上がり時間」に変換した。次に、3つの実験ごとに、持続時間ごとの『「立ち上がり時間」の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係』を図に示した。そして、各持続時間におけるこの関係を最も近似する4パラメータロジスティック回帰モデルを求め、先ほどの図に重ねた。それが、図 8.2.1 である。この図より、「立ち上がり時間」におけるロジスティック回帰モデルの予測線は、「立ち上がり時間」に変換したことによって予測線の長さが持続時間ごとに異なるが、今まで見てきた『「立ち上がり時間の割合」の増加と  $\Delta PC_{A,D100R0}$  の変化との関係』を近似したロジスティック回帰モデルの予測線の形と変わらないように見えた。

次に、「立ち上がり時間」に対する回帰モデルにおける2つの指標を用いてクラスター分析を行った。その結果は、5章から7章で示した「立ち上がり時間の割合」に対するクラスター分析の結果と一致した。

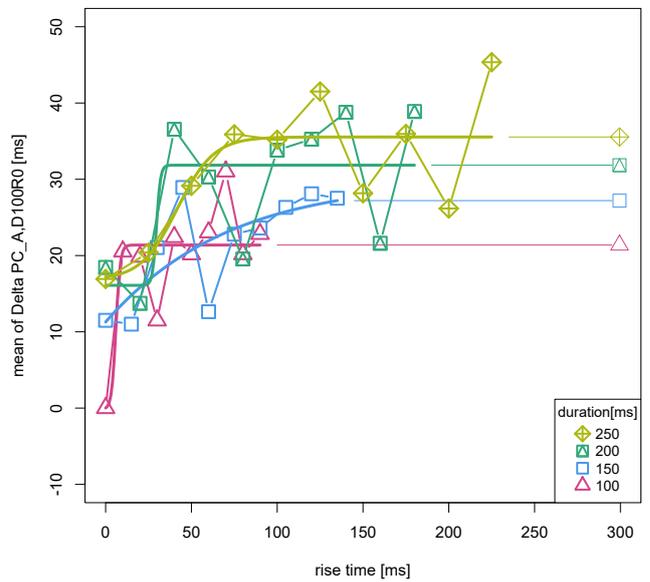
続いて、以上の傾向を確かめるべく、『「立ち上がり時間」の増加と  $\Delta PC_{A,D100R0}$  の変化との関係』を近似したロジスティック回帰モデルの予測線と、『「立ち上がり時間の割合」の増加と  $\Delta PC_{A,D100R0}$  の変化との関係』を近似したロジスティック回帰モデルの予測線とが一致するかどうかを確かめた。

まず、前者の「立ち上がり時間」に対する回帰モデルの予測線を、持続時間ごとに横方向に引き伸ばし、「立ち上がり時間の割合」に対する回帰モデルの予測線と重ねた。それが図 8.2.2 である。図の見方はキャプションを参照してほしい。この図より、両者はよく一致しており、その差異は僅かであることが明らかである。

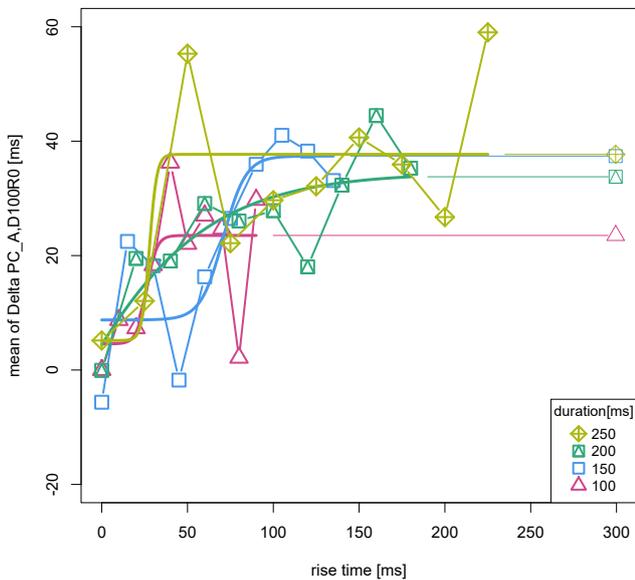
次に、これを数値として確かめるべく、各持続時間における「立ち上がり時間の割合」に対する回帰モデルの予測式と「立ち上がり時間」に対する回帰モデルの予測式との予測値の差分を算出した。その際、「立ち上がり時間の割合」を10%刻みに変化させた値と、それを「立ち上



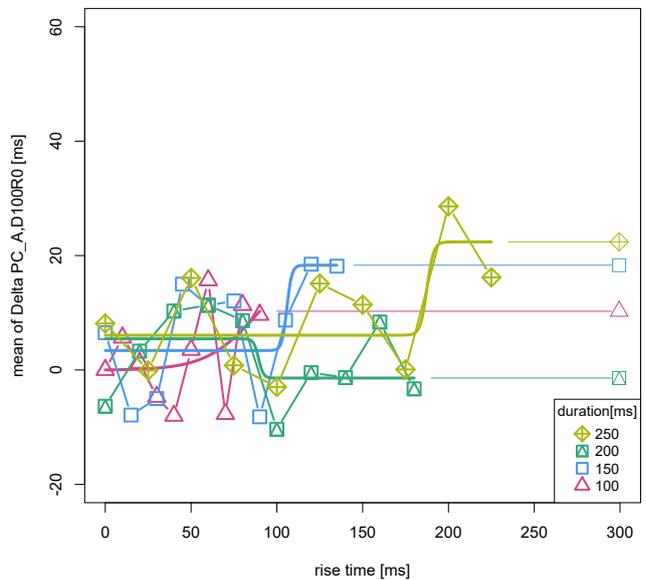
(a) 実験 1



(b) 実験 3



(c) 実験 2 の参加者 3 人のグループ



(d) 実験 2 の参加者 7 人のグループ

図 8.2.1: 持続時間ごとの、『立ち上がり時間』の増加と  $\Delta PC_{A,D100R0}$  の変化との関係』と、それを最も近似する 4 パラメータロジスティック回帰モデルの予測線  
 横軸の “rise time [ms]” は「立ち上がり時間」を表し、縦軸の “mean of Delta PC<sub>A,D100R0</sub>” は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある持続時間の音 (シンボル) における、その音の「立ち上がり時間」を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  (縦軸) が示されている。この変化に、太線によって、ロジスティック回帰モデルの予測線が重ねられている。

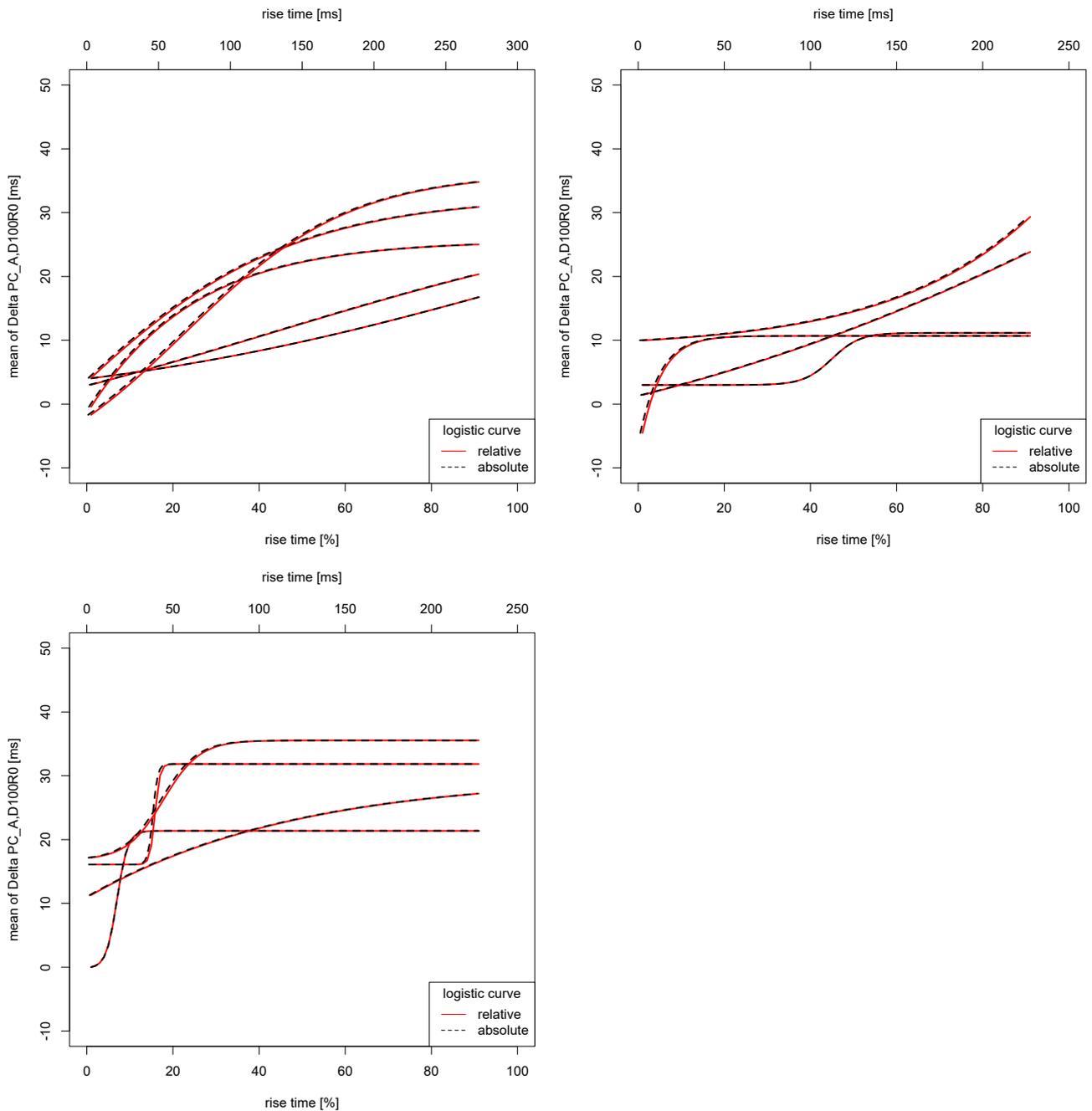


図 8.2.2: 「立ち上がり時間」に対する回帰モデルの予測線を横方向に引き伸ばし、「立ち上がり時間の割合」に対する回帰モデルの予測線と重ねた図  
 横軸の“rise time [%]”は「立ち上がり時間の割合」を表し、縦軸の“mean of Delta PC\_A,D100R0”は  $\Delta PC_{A,D100R0}$  の平均値を表す。引き伸ばされた「立ち上がり時間」に対する回帰モデルの予測線(凡例における“absolute”)を黒色の点線で示し、「立ち上がり時間の割合」に対する回帰モデルの予測線(凡例における“relative”)を赤色の実線で示した。

がり時間」に変換した値を入れた。その結果、「立ち上がり時間の割合」と「立ち上がり時間」との予測値の差は最大で、実験1で1.4 ms、実験2で0.94 ms、実験3で0.98 msであった。すなわち、殆ど差異がなかった。

以上のことより、本論にて「立ち上がり時間の割合」で分析し結論付けたことは、「立ち上がり時間」であっても成立すると考えられる。すなわち、『PAT 知覚も統合の時間窓の影響を受け、統合の時間窓の範囲内と範囲外では、「立ち上がり時間」の増加に対する PAT の変化量と曲率が異なる』ことが成立すると考えられる。

## 8.3 持続時間や周波数の変化と PAT の変化との関係

実験3にて、持続時間が増加するにつれ、「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」は大きく変化しないまま、 $PAT(\Delta PC_{A,D100R0})$  が時間の後方へ移動する可能性が示された。そこで、本節にて、「持続時間の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」について検討する。また、実験1と実験3の比較を通して、周波数が異なる時に  $PAT(\Delta PC_{A,D100R0})$  が移動するかについても検討する。

「持続時間の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」を明らかにするために分析を行った。5章から7章で行った「立ち上がり時間の割合の増加と  $PAT(\Delta PC_{A,D100R0})$  の移動との関係」と同様に分析を行うのが望ましいが、本論で用いた持続時間は5水準あるいは4水準のみであるため、ロジスティック回帰モデルでの近似が非常に不安定になることが容易に想定される。そこで、本節では、図より目視で考察を行った。そのため、より安定した結果である、実験1と実験3のみを比較した。

まず、立ち上がり時間の割合ごとの「持続時間の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を図8.3.1に示した。この図の見方はキャプションを参照してほしい。

この図より、実験1では、立ち上がり時間の割合によって、持続時間を増加させた時の  $\Delta PC_{A,D100R0}$  の変化が大きく異なるが、実験3では、類似していることが明らかである。すなわち、先述したように、実験1では「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によって異なるが、実験3では「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が持続時間によらず同じであることが再度確認された。

さらに、この図より、実験3では、立ち上がり時間の割合によらず、持続時間が増加するにつれて  $\Delta PC_{A,D100R0}$  が増加する傾向も明らかであった。一方、実験1では、立ち上がり時間の割合が50%以上(右図)では、持続時間が増加するにつれて  $\Delta PC_{A,D100R0}$  が増加する傾向がありそうだが、左図では言及しにくかった。これは、実験1では、立ち上がり時間の割合が増加するにつれて  $\Delta PC_{A,D100R0}$  が大きく変化することにより、持続時間が増加するにつれて  $\Delta PC_{A,D100R0}$  がどのように変化するのが把握しにくくなっているためであると考えられた。

そこで、立ち上がり時間の割合の全水準を平均したときの、持続時間の増加(横軸)に対する  $\Delta PC_{A,D100R0}$ (縦軸)を図に示した。これが図8.3.2である。

この図より、次の二つのことが確認された。一つ目に、実験1でも実験3でも、持続時間が増加するにつれ  $\Delta PC_{A,D100R0}$  が増加する傾向が見られた。より具体的には、持続時間が100msから250msまでは、持続時間が増加するにつれ  $\Delta PC_{A,D100R0}$  がほぼ直線的に増加したが、持続時間が250ms以上では、持続時間が増加しても  $\Delta PC_{A,D100R0}$  が増加しない可能性が示唆された。持続時間が100msから250(300)msまで増加した時の、 $\Delta PC_{A,D100R0}$  の増加量は約10ms

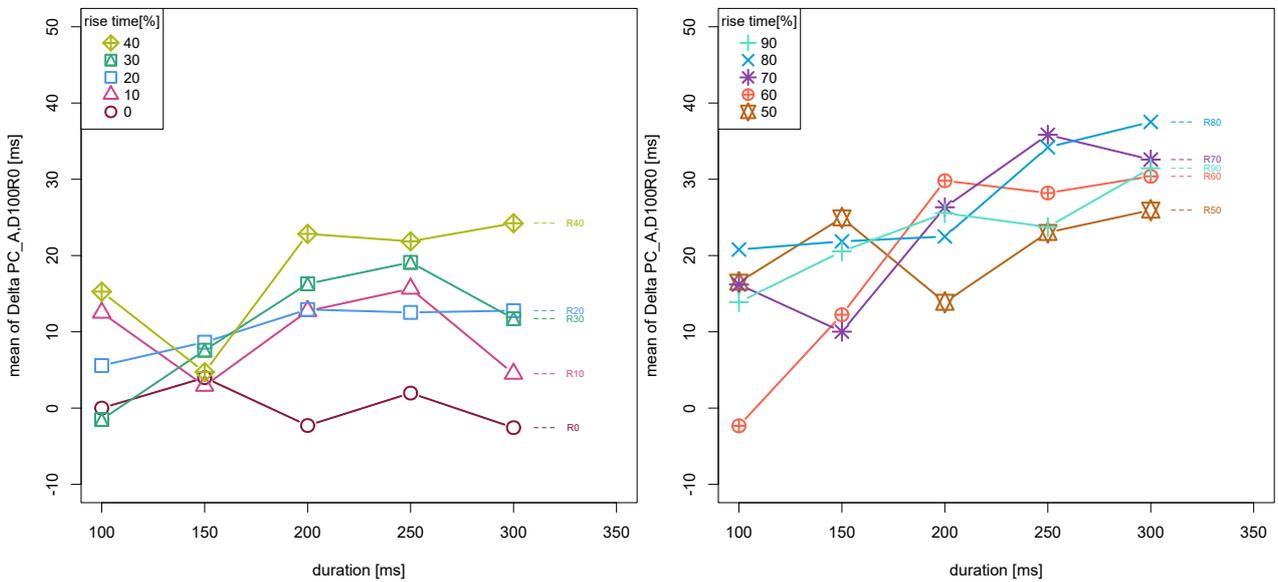
であった。二つ目に、周波数が高くなると  $\Delta PC_{A,D100R0}$  が増加する傾向が見られた。すなわち、440 Hz の刺激音を用いた実験 1 よりも、3520 Hz の刺激音を用いた実験 3 の方が、どの持続時間でも  $\Delta PC_{A,D100R0}$  が約 10 ms 大きかった。

以上より、次のことが明らかになった。

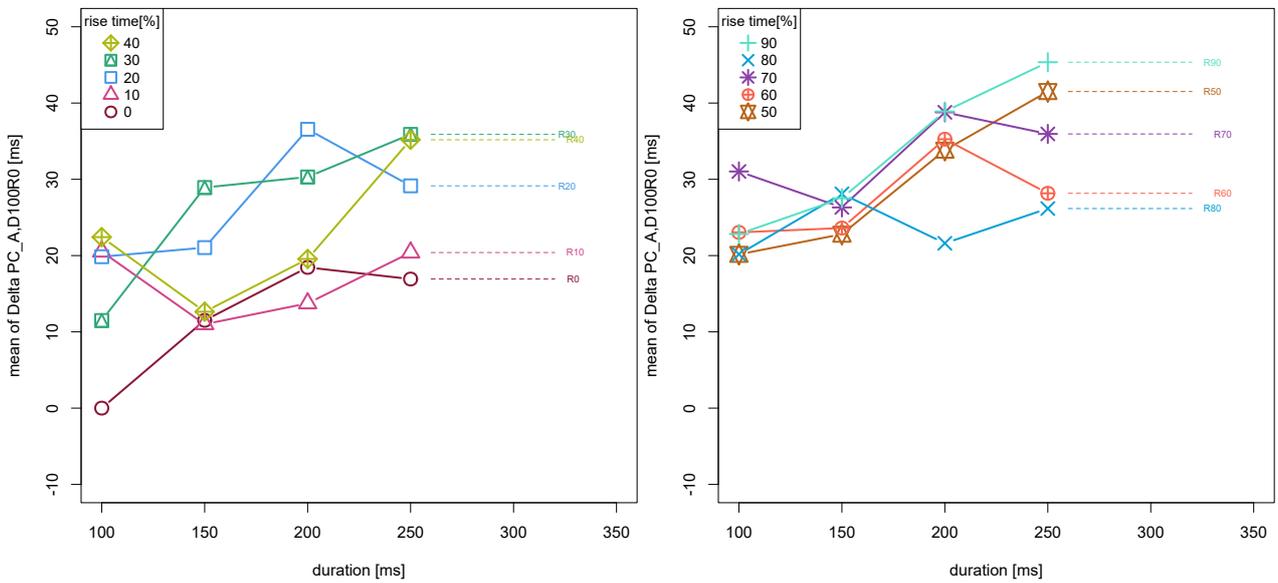
まず、持続時間に関して述べる。立ち上がり時間の割合の全水準で平均した場合に、持続時間が 100 ms から 250 ms まで増加するにつれて PAT が約 10 ms 程度、時間的後方へ移動した。持続時間の増加に対する PAT の移動は、直線的であった。このことより、持続時間が増加するにつれて PAT が時間的後方へ僅かに移動する可能性が示唆された。

次に、周波数について述べる。周波数を 440 Hz と 3520 Hz とで変化させた時、立ち上がり時間の割合の全水準で平均した場合に、持続時間に寄らず、PAT が 10 ms 程度、時間的後方へ移動することが明らかとなった。このことより、周波数が高くなると PAT も時間的後方へ僅かに移動する可能性が示唆された。

だが、これらは本論の仮説に直接関わることではないため、厳密な分析が行えなかった。本論で示せたのは可能性にすぎず、結論づけるためには、より精密に実験を行う必要がある。これは今後の研究が俟たれる。



(a) 実験 1



(b) 実験 3

図 8.3.1: 立ち上がり時間の割合ごとの、持続時間の増加に対する  $\Delta PC_{A,D100R0}$  の変化  
 横軸の “duration[ms]” は持続時間を表し、縦軸の “mean of Delta PC\_A,D100R0” は  $\Delta PC_{A,D100R0}$  の平均値を表す。ある立ち上がり時間の割合の音 (シンボル) における、その音の持続時間を増加させた時の (横軸)、 $\Delta PC_{A,D100R0}$  (縦軸) が示されている。立ち上がり時間の割合の水準が多いため、視認性向上のため、便宜的に 2 つの図に分けた。さらに、図の右側に、各立ち上がり時間の割合の値を “R + 値” と略記した。例えば、立ち上がり時間の割合が 0% の場合は “R0” と記した。

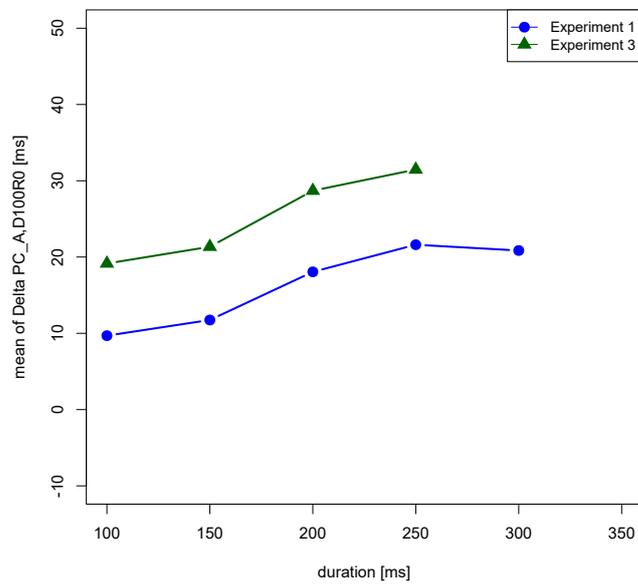


図 8.3.2: 立ち上がり時間の割合を平均した、持続時間の増加に対する  $\Delta PC_{A,D100R0}$  の変化  
 横軸の “duration[ms]” は持続時間を表し、縦軸の “mean of Delta PC\_A,D100R0” は  $\Delta PC_{A,D100R0}$  の  
 平均値を表す。立ち上がり時間の割合の全水準を平均したときの、持続時間の増加 (横軸) に対する  
 $\Delta PC_{A,D100R0}$  (縦軸) が示された。丸のシンボルが実験 1 の結果を示し、三角のシンボルが実験 3 の結  
 果を示す。

## 8.4 先行研究と本論の実験との比較

殆どの先行研究では、持続時間や立ち上がり時間が増加するにつれ PAT が移動するかどうかを結論付けるために、分散分析が行われた。そのため、まず、それらの先行研究と本論との分散分析の結果を比較した。

次に、持続時間や立ち上がり時間の割合と  $\Delta PC_{A,D100R0}$  との相関係数を算出し、相関係数を報告した研究と本論の結果を比較した。

### 分散分析の結果の比較

本論と同じく、分散分析の要因に、立ち上がり時間の割合あるいは立ち上がり時間と、持続時間とを含めた研究は、P.Vos ら [35] と Danielsen ら [5] のみである。従って、この2つの研究と本論との分散分析の結果を比較した。

参考のため、2章にて、PAT の先行研究で用いられた刺激音と分散分析の結果をまとめた表 2.4.1 とこの表における分散分析の結果の列を行列にした表 2.4.2 に、本論の3つの実験を追加したものを表 8.4.1 と表 8.4.2 に示した。必要に応じて参照してほしい。

本論の音楽群の実験1と実験3の結果は、P.Vos らの結果や Danielsen らの実験1の結果とは一致したが、Danielsen らの実験2とは一致しなかった。一方、本論の実験2は、P.Vos らや Danielsen らと一致しなかった。

このように分散分析結果が先行研究と一致しないこともあった。だが、これにより、本論と先行研究とで、持続時間や立ち上がり時間の増加に対する PAT の移動の傾向が全く異なったとは言えない。それは次の二つの理由からである。

まず一つ目に、得られた PAT の移動の傾向が分散分析の結果に反映されないことが多くあるからだ。例えば、Danielsen らは、実験1と実験2で PAT の移動の傾向は同様であったにもかかわらず、実験1の分散分析では有意であった主効果や交互作用が実験2では有意とならなかったと報告した。また、本論の実験1でも、分散分析の結果が『持続時間 150 ms 以下と 200 ms 以上とで、「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」が異なる』ことを反映していると考えられたが、これを正確に反映した場合に有意になるはずの組み合わせが有意にならなかったことがあった。

二つ目に、分散分析は線形モデルを用いた分析手法であるが、PAT の移動の傾向は非線形であるからだ。本論では持続時間や立ち上がり時間の増加と PAT ( $\Delta PC_{A,D100R0}$ ) の変化との関係を近似するのに、非線形であるロジスティック回帰モデルを前提とした。実際にこの非線形のモデルを適用し、求められた回帰モデルの予測線は殆どが非線形であり、線形で近似できそうなものは僅かであった。特に統合の時間窓の範囲外の場合で非線形であった。すなわち、直線回帰や分散分析などの等分散正規分布を仮定する線形モデルの分析方法は、PAT 知覚における

「持続時間の増加と PAT の移動との関係」や「立ち上がり時間の増加と PAT の移動との関係」を分析するためには、適切ではないと考えられる。

## 相関係数の結果の比較

続いて、Bechtold ら [1] の算出した相関係数との比較を行う。Bechtold らは、持続時間と立ち上がり時間が異なる 9 つの刺激音を用いて実験を行った。そして、持続時間あるいは立ち上がり時間と参加者ごとの PAT との相関係数を算出した<sup>1</sup>。その結果、持続時間との相関係数も立ち上がり時間との相関係数も低かったと報告した。

そこで、本論でも、各実験において持続時間や立ち上がり時間の割合と  $\Delta PC_{A,D100R0}$  との相関係数を算出した。

まず、持続時間と  $\Delta PC_{A,D100R0}$  との相関係数を算出した。この際、参加者ごとに、持続時間ごとの、立ち上がり時間の割合の全水準で平均した  $\Delta PC_{A,D100R0}$  を算出した。その後、これと持続時間とのピアソンの積率相関係数を算出した。その値は、実験 1 では  $r = 0.343$  であり、実験 2 では  $r = 0.122$  であり、実験 3 では  $r = 0.244$  であった。確かに Bechtold の結果と一致して、相関係数が低かった。

次に、立ち上がり時間の割合と  $\Delta PC_{A,D100R0}$  との相関係数を算出した。この際、参加者ごとに、立ち上がり時間の割合ごとの、持続時間の全水準で平均した  $\Delta PC_{A,D100R0}$  を算出した。その後、これと立ち上がり時間の割合とのピアソンの積率相関係数を算出した。その値は、実験 1 では  $r = 0.487$  であり、実験 2 では  $r = 0.205$  であり、実験 3 では  $r = 0.238$  であった。確かに Bechtold の結果と一致して、相関係数が低かった。

以上のように、本論の実験より得られた、持続時間あるいは立ち上がり時間の割合と PAT との相関係数は低い値であった。これは先行研究と一致する結果であった。だが、本論では、持続時間が増加するにつれ、立ち上がり時間の割合が増加するにつれ、PAT が移動することが示された。従って、相関係数の結果を根拠に、持続時間や立ち上がり時間が増加するにつれ PAT が移動するか否かを議論することはできないといえる。

---

<sup>1</sup>Bechtold は、持続時間あるいは立ち上がり時間と、Bechtold の言うところの “360 PAT estimates” との相関係数を算出したことが次の文章から明らかである。

“As suggested by earlier research, onset rise time was positively correlated with the 360 PAT estimates. Yet, this correlation was weak [ $r = 0.143, t(358) = 2.748, p = 0.006$ ].”

“the correlation between duration and the 360 PAT estimates was weak [ $t(358) = 2.685, p = 0.007, r = 0.140$ ]”  
そして、この “360 PAT estimates” は、各刺激音に対する参加者ごとの PAT であることが次の文章から明らかである。

“The experiment yielded 360 (9 stimuli  $\times$  40 participants) valid PAT estimates.”

表 8.4.1: PAT の先行研究で用いられた刺激音と分散分析の結果

\*1 ... 「論文」では、各論文の名称として、第一著者と出版年を記載した。その論文中で2つ以上の実験を行っているものは、どの実験かも記載した。

\*2 ... 「課題」では、PAT 測定課題を以下の方法で略記した。

タッピング課題は「タッピング」、同時調整課題は「同時調整」、等時調整課題は「等時調整」と記された。

\*3 ... 「刺激音」では、タッピング課題においてはすべての刺激音を、等時調整課題と同時調整課題では、テスト音を記載した。

\*4 ... 「時間範囲」では、統合の時間窓 (150–300ms) を 200 ms とした場合に、用いられた持続時間が統合の時間窓の範囲内か範囲外か、それとも範囲内から範囲外までかを示した。

「内」の表記は用いられた持続時間が 200 ms 以内であることを、「外」の表記は持続時間が 200 ms 以上であることを、「内外」の表記は持続時間が 200 ms 以内から 200 ms 以上までであることを表す。

\*5 ... 「分散分析の結果」では、分散分析の各要因を略記した。略記の仕方は計画の列に記載したものをを用いた。また、各要因の主効果については要因の略記のみとし、交互作用は2つの要因を“x”でつないで略記した。そして、有意は「SIG.」有意でないは「N.S.」と略記した。論文中に何も報告されなかった要因は有意ではなかったと推測し、「(N.S.)」と記載した。

例) 立ち上がり時間 “R” の主効果が有意であった場合、「“R” SIG.」と記載した。

立ち上がり時間 “R” と持続時間 “D” の交互作用が有意であった場合、「“R” × “D” SIG.」と記載した。

論文*1	課題*2	刺激音*3	時間範囲*4	立ち上がり時間	持続時間	周波数	提示音圧レベル	値	分散分析の計画	分散分析の結果*5
本論実験 1 (音楽群)	タッピング	合成音	内外	0–90% (0–270 ms)	100–300 ms	[440 Hz]	[76 dB]	ΔPC –2.56–37.51 ms	二要因 (立ち上がり時間 “R”、 持続時間 “D”)	“R” SIG. “D” SIG. “R” × “D” SIG.
本論実験 2 (一般群)	タッピング	合成音	内外	0–90% (0–225 ms)	100–250 ms	[440 Hz]	[76 dB]	ΔPC –4.46–29.04 ms	二要因 (立ち上がり時間 “R”、 持続時間 “D”)	“R” SIG. “D” N.S. “R” × “D” N.S.
本論実験 3 (音楽群)	タッピング	合成音	内外	0–90% (0–225 ms)	100–250 ms	[3520 Hz]	[76 dB]	ΔPC 0–45.35 ms	二要因 (立ち上がり時間 “R”、 持続時間 “D”)	“R” SIG. “D” SIG. “R” × “D” SIG.
P.Vos1995[35] 実験 2	タッピング	合成音	内外	0–80% (0–240 ms)	1–300 ms	[500 Hz]	参加者が調整	同調のズレ量 目視: –55–100 ms	二要因 (立ち上がり時間 “R”、 持続時間 “D”)	“R” SIG. “D” SIG. “R” × “D” SIG.
Danielsen2019[5] 実験 2	タッピング	合成音	内外	0–50 ms	1–400 ms	100–3000 Hz	参加者が調整	同調のズレ量 –26.71–26.48 ms	三要因 (立ち上がり時間 “R”、 持続時間 “D”、 周波数 “F”)	“R” N.S. “D” N.S. “F” N.S. (全て N.S. で傾向止まり)
	同時調整	合成音	内外	0–50 ms	1–400 ms	100–3000 Hz	参加者が調整	ΔPC 0.42–34.46	同上	同上
Collins2006[3]	等時調整	合成音	内	0–100 ms	[200 ms]	130–2093 Hz	記載なし	ΔPC 目視: –10–80 ms	三要因 (立ち上がり時間 “R”、 周波数 “F”、 繰り返し “Rep”)	“R” SIG. “F” N.S. “Rep” N.S. “R” × “F” N.S. 全ての交互作用 N.S.
J.Vos1981[34] 実験 1	等時調整 (持続時間を調整)	合成音	—	5–80 ms	調整値	[400 Hz]	[82 dB]	Δt (≈ΔPC) –1.6–34.9 ms	一要因 (立ち上がり時間の ペアの種類 “Rpair”)	“Rpair” SIG.
J.Vos1981[34] 実験 2	等時調整 (持続時間を調整)	合成音	—	5–80 ms	調整値	[400 Hz]	37, 57, 77 dB	Δt (≈ΔPC) –53.3–58.1 ms	四要因 (ベース音の立ち上がり時間 “Rbase”、 テスト音の立ち上がり時間 “Rtest”、 音量 “L”、 繰り返し “Rep”)	“Rbase” SIG. “Rtest” SIG. “Rbase” × “L” SIG. “Rtest” × “L” SIG. その他は記載なし (“L” N.S.)
Danielsen2019[5] 実験 1 (London2019[15] 実験 1)	タッピング	楽器音	内外	0–43 ms	1–487 ms	538–8199 Hz	参加者が調整	同調のズレ量 –21.57–16.89 ms	三要因 (立ち上がり時間 “R”、 持続時間 “D”、 周波数 “F”)	“R” SIG. “D” SIG. “F” SIG. “D” × “F” SIG. “D” × “R” SIG. その他は記載なし
	同時調整	楽器音	内外	0–43 ms	1–487 ms	538–8199 Hz	参加者が調整	ΔPC –0.50–23.44 ms	同上	同上
Gordon1987[8] (Gordon1984[9])	同時調整、 等時調整	楽器音	外	45–122 ms	[320 ms]	[311 Hz]	[–90 dB]	両課題平均 ΔPC –21–28 ms	一要因 (楽器の種類 ≈ 立ち上がり時間 “R”)	“R” SIG.
Bechtold2018[1]	同時調整	楽器音 (奏法変化)	外	60–173 ms	平均 564ms ± SD 77ms	[311 Hz; Eb4]	参加者が調整 (刺激音の RMS: –17.1––36.9 dB)	ΔPC –23––83 ms	— 相関係数 (立ち上がり時間 “R” と ΔPC、 持続時間 “D” と ΔPC)	— “R” と ΔPC 弱い “D” と ΔPC 弱い
Nymoen2017[19]	同時調整	楽器音	—	記載なし	記載なし	記載なし	記載なし	言及ないが多分 ΔPC 目視: 0–20 ms	—	—
Polfremam2013[22]	同時調整 (ベース音も 調整可能)	楽器音、 正弦波	—	記載なし	記載なし	[130 Hz; C4]	参加者が調整 (実験中も 調整可能)	ΔPC –0.46–51.21 ms	—	—
Camp2011[2]	同時調整	楽器音	内外	0.07–1666 ms	2–4700 ms	77–465 Hz, 22000 Hz	記載なし	ΔPC –6.2–46.5 ms	様々な分析	—
Wright2008[37]	同時調整	楽器音	外	ばらばら (記載なし)	[300 ms]	[311 Hz; Eb4]	参加者ごとの 再生環境	ΔPC の中央値 –14.0–40.2 ms	—	—

表 8.4.2: 先行研究の分散分析結果

表 2.4.1 の分散分析の列にてまとめたものを行列に変換したものである。

分散分析が行われ、有意であったと報告された要因には“SIG.”と、有意ではないと報告された要因には“N.S.”と表記した。分散分析が行われたが論文中に何も報告されなかった要因は、その要因は有意ではなかったと推測し、“(N.S.)”と記載した。

統制された要因は“[-]”と表記した。また、提示音圧レベルを参加者が調整した場合には“好み”と表記した。J.Vos らが用いた課題では持続時間の長さを調整したため、彼らの持続時間のセルに“調整”と記載した。

検討されなかった交互作用は“(空白)”とした。

\*1 ... 「論文」では、各論文の名称として、第一著者と出版年を記載した。その論文中で2つ以上の実験を行っているものは、どの実験かも記載した。

論文*1	刺激音	立ち上がり時間 (R)	持続時間 (D)	周波数 (F)	提示音圧レベル (L)	D×R	D×F	F×L	R×F	R×L	F×L
本論実験1	合成音	SIG.	SIG.	[-]	[-]	SIG.					
本論実験2	合成音	SIG.	N.S.	[-]	[-]	N.S.					
本論実験3	合成音	SIG.	SIG.	[-]	[-]	SIG.					
P.Vos ら [35]	合成音	SIG.	SIG.	[-]	好み	SIG.					
Danielsen 実験2[5]	合成音	N.S.	N.S.	N.S.	好み	N.S.	N.S.		N.S.		
Collins[3]	合成音	SIG.	[-]	N.S.	[-]					N.S.	
J.Vos ら 実験1[34]	合成音	SIG.	調整	[-]	[-]						
J.Vos ら 実験2[34]	合成音	SIG.	調整	[-]	(N.S.)						SIG.
London ら [15] / Danielsen 実験1[5]	楽器音	SIG.	SIG.	SIG.	好み	SIG.	SIG.			(N.S.)	
Gordon[9, 8]	楽器音	SIG.	[-]	[-]	[-]						

# 第9章 結論

## 9.1 本論の概略と結論

音楽において、リズムは重要である。特に、アンサンブルでは、他者とリズムを合わせることで、簡略化すると、他者の音と自分の音が鳴った瞬間を合わせることは重要である。この他者と合わせる「音が鳴った瞬間」は、“Perceptual Attack Time; PAT”と呼ばれる。このPATは、「時間の流れの中に配置された、知覚された音のアタックの瞬間」と定義される。

PATの先行研究では、音の何の要素を変化させるとPATが変化するのが調査されてきた。それらの研究より、持続時間が長い音の方が、また、立ち上がり時間が長い音の方がPATが時間的後方に位置することが報告された。また、持続時間と立ち上がり時間との交互作用も報告された。加えて、周波数や提示音圧レベルが異なるとPATが変化する可能性も示された。

だが、PATの先行研究では、持続時間や立ち上がり時間がどのくらい長くなるとPATがどのくらい時間的後方へ移動するのか、すなわち、持続時間あるいは立ち上がり時間の増加とPATの移動との関係は殆ど調査されていない。加えて、周波数や提示音圧レベルがPATをどのように変化させるかについては結論が出ていない。

一方、音の知覚は聴覚情報処理を経て形成されるため、PATの知覚も聴覚情報処理の何らかの影響を受けるはずである。特に、PATの先行研究では、実験で刺激音の持続時間を1msから400msあるいは500ms程度まで変化させている。この持続時間の範囲は「統合の時間窓」の時間範囲と重なっているため、「統合の時間窓」の影響を受ける可能性が考えられる。「統合の時間窓」とは、聴覚情報が保持される時間範囲の一つである。この時間範囲で保持された情報はまとめて処理されるが、その時に、これらの情報が統合処理される。この処理により、「統合の時間窓」の範囲内であれば、持続時間が増加するにつれ、最小可聴閾が低下しラウドネスが上昇して知覚されるという現象が生じるが、「統合の時間窓」の範囲外ではこの現象は生じず、持続時間が増加してもラウドネスも最小可聴閾も変化しない。統合の時間窓の先行研究より、統合の時間窓の長さは150–300msの範囲のどこかにあると考えられる。また、提示音圧レベルが増加するにつれ、かつ、周波数が高くなるにつれ、統合の時間窓が短くなる可能性が示唆されている。加えて、音楽訓練を受けると統合の時間窓が長くなる可能性も示唆されている。

そこで、本論では、統合の時間窓の先行研究とPATの先行研究より明らかになっていることを照らし合わせ、「PATの知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり

時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」と仮説だてた。

PAT 知覚は複雑な事象であり、未だ確立した知見が少ない研究領域である。そのため、本論は PAT 知覚に関する基礎研究として行われた。

仮説の検証のため、PAT を測定する実験を行う。PAT の測定課題には、タッピング課題、同時調整課題、そして等時調整課題があるが、先行研究からは、これらの3つの課題により得られる PAT が異なるという確証が得られていない。だが、本論の条件では、タッピング課題を用いるのが適切であると考えられた。そこで、予備実験において、3つの PAT 測定課題を用いて PAT を測定した。得られた値に有意差がないことが確認されたため、PAT 測定の課題としてタッピング課題を用いることに問題がないと判断した。その際、提示音圧レベルを2水準(76 dB、84 dB)で変化させたが、有意差が見られなかった。さらに、本実験で持続時間と立ち上がり時間の割合を変化させる水準を検討し、サンプルサイズも検討した。

その後、本実験を行った。本論では統合の時間窓を直接測定しない。そのため、PAT の知覚に統合の時間窓が関係しているのかを確認するには、統合の時間窓がある長さの時の刺激音の PAT と、統合の時間窓が異なる長さの時の刺激音の PAT とを比較する必要がある。先行研究より、統合の時間窓の長さを変化させ得るのは、提示音圧レベル、周波数、そして音楽訓練を受けたことがあるかどうか(音楽訓練の有無)である。このうちの提示音圧レベルは、予備実験で用いた音圧レベル2水準では有意差がないことが確認された。そこで、実験1を基準とし、統合の時間窓が異なる長さの時として、実験2にて音楽訓練の有無を取り上げ、実験3にて周波数を取り上げ、PAT 測定を行った。また、各実験では、刺激音は最も単純な合成音であるサイン波とし、持続時間と「立ち上がり時間の割合(持続時間に占める立ち上がり時間の割合であり、数式では(立ち上がり時間)/(持続時間)×100[%]である)」を変化させた。

以下にて各実験の詳細を述べる。

まず、実験1では、音楽訓練を受けたことがある参加者(音楽群)に対しタッピング課題を行った。刺激音はサイン波とし、その持続時間を5水準(100, 150, 200, 250, 300 ms)で変化させ、立ち上がり時間の割合を10水準(0, 10, 20, 30, 40, 50, 60, 70, 80, 90%)で変化させた。また、周波数は440 Hzとした。これは、現在の音楽における基準となる周波数だからであり、JIS規格でも楽器用標準周波数はト音記号の音名“イ”の音の周波数で440Hzであると定められている[13]からである。提示音圧レベルは予備実験にて有意差が見られなかったため、参加者間で固定し、76 dBとした。

次に、実験2にて、統合の時間窓の長さが異なる場合として音楽訓練の有無を取り上げた。実験1の音楽群よりも統合の時間窓が短いと考えられる、音楽訓練を受けたことがない参加者(一般群)に対しタッピング課題を行った。刺激音はサイン波とし、その持続時間を4水準(100, 150,

200, 250 ms) で変化させ、立ち上がり時間の割合を 10 水準 (同上) で変化させた。また、周波数は実験 1 と同様に 440 Hz とした。提示音圧レベルも実験 1 と同様に、参加者間で固定し、76 dB とした。

最後に、実験 3 にて、統合の時間窓の長さが異なる場合として周波数を取り上げた。実験 1 の 440 Hz の刺激音よりも統合の時間窓が短いと考えられる、より高い周波数の刺激音を用いて、実験 1 と同様の音楽群に対しタッピング課題を行った。刺激音はサイン波とし、実験 2 と同様に、その持続時間を 4 水準 (同上) で変化させ、立ち上がり時間の割合を 10 水準 (同上) で変化させた。また、周波数は 3520 Hz とした。提示音圧レベルは実験 1 や実験 2 と同様に、参加者間で固定し、76 dB とした。

各実験の分析は以下のように進められた。まず、ある刺激音を基準とし、その基準の刺激音  $R$  と、その基準音を含む全ての刺激音  $A$  との  $\Delta PC_{A,R}$  を PAT として算出した。次に、ある持続時間の音 (シンボル) における、その音の立ち上がり時間の割合を増加させた時の (横軸)、PAT (縦軸) を図に示すことで、持続時間ごとの「立ち上がり時間の割合の増加と PAT ( $\Delta PC_{A,R}$ ) の移動との関係」を示した。そして、各持続時間におけるこの関係を最も近似する 4 パラメータロジスティック回帰モデルを求めた。4 パラメータロジスティック回帰モデルとは、予測値がデータに合わせた最小値から最大値までの間を S 曲線の形で変化していくモデルである。だが、そのパラメータの値によっては、S 字の真ん中の部分や S 字曲線の下の方の形、あるいは、S 字の上の方のみを当てはめることが可能である。続いて、求められた回帰モデルの予測線を先ほどの図に重ね、その予測線の形が持続時間によって異なるかを視覚的に確認した。さらに、求められた 4 パラメータロジスティック回帰モデルにより「立ち上がり時間の割合が 0 の時の予測値 (PAT) と立ち上がり時間の割合が 100 の時の予測値 (PAT) との差分」と「パラメータの一つである傾き  $b$  の値」を算出し、これらの値が持続時間によって異なるかをクラスター分析を行って確認した。

以上の 3 つの実験とその分析により、本論の仮説である「PAT の知覚は統合の時間窓の影響を受ける。その影響により、立ち上がり時間の増加と PAT の移動との関係が、音の持続時間が統合の時間窓の範囲内か範囲外かで異なる」ことが支持された。より具体的には、統合の時間窓の範囲内 (実験 1 の持続時間 150 ms 以下) では、立ち上がり時間の割合が増加するにつれ、PAT はゆるやかに時間的後方へ移動する。その移動量は、音の持続時間が統合の時間窓の範囲外の場合のその移動量よりも小さい。ロジスティック回帰モデルの予測線の形状としては、S 字曲線の下の方の形か、あるいは S 字の真ん中の部分である。一方、統合の時間窓の範囲外 (実験 1 の持続時間が 200 ms 以上、実験 2 と実験 3 の持続時間が 100 ms 以上) では、立ち上がり時間の割合が増加するにつれ、PAT が急激に時間的後方へ移動する。その移動量は、音の持続時間が統合の時間窓の範囲内の場合の移動量よりも大きい。加えて、立ち上がり時間の割合があ

る一定以上になると、立ち上がり時間の割合が増加するにつれてのPATの移動量が徐々に小さくなるか、PATが移動しなくなる。ロジスティック回帰モデルの予測線の形状としては、S字の上の方か、あるいはS字曲線全体の形である。

従って、「PAT知覚は統合の時間窓の影響を受ける。その影響により、統合の時間窓の範囲内と範囲外では、立ち上がり時間の割合の増加に対するPATの変化量と、曲率が異なる」と結論付けた。

この結論は「立ち上がり時間の割合」に対して分析し明らかになったことであるが、「立ち上がり時間」に対して分析しても同様の結論が得られることが確認された。

以上のようなPAT知覚における持続時間と立ち上がり時間との交互作用の詳細な記述は先行研究では行われておらず、本論によって初めて明らかになったことである。また、聴覚情報処理システムの一つである「統合の時間窓」がPAT知覚に影響を与えることも、本論が新規に明らかにした。

## PATの移動量を音価に変換する

各実験結果を、音楽におけるテンポやリズムの観点から考えてみる。

実験と同様に700msごとに4分音符が鳴る、BPMが約86(正確にはBPMが85.71)の楽曲を想定しよう。この楽曲では、理論的には、4分音符が1分間に約86個ほど演奏される。すなわち、音価ごとの持続時間を示すならば、4分音符1つが700msの持続時間、8分音符が350ms、16分音符が175ms、32分音符が87.5ms、64分音符が43.75msの持続時間である。付点の場合の持続時間は、付点4分音符が1050ms、付点8分音符が525ms、付点16分音符が262.5ms、付点32分音符が131.25ms、付点64分音符が65.625msの持続時間である。

実験で用いた刺激音の持続時間と立ち上がり時間の割合を、BPMが約86の楽曲での音価に変換してみよう。立ち上がり時間の割合は、3つの実験とも0-90%で変化させたため、ここでは簡略化して0-100%で変化させたと考えよう。一方、持続時間は、実験1では100-300msで、実験2と実験3では100-250msで変化させた。これをBPMが約86の楽曲での音価に変換すると、実験1では「長めの32分音符」から「短めの8分音符」まで変化させ、実験2と実験3では「長めの32分音符」から「短めの付点16分音符」まで変化させたことになる。

続いて、このように刺激音を変化させた時のPATの移動量をBPMが約86の楽曲での音価に変換してみよう。立ち上がり時間の割合を0%から90%まで増加させた時のPATの移動量は、実験1では最大で約36msであり、実験3では最大で約21msであった。これを音価に変換すると、64分音符が43.75msであるから、PATは最大で「短めの64分音符」程度、移動したことになる。僅かではあるが十分にずれを感知できる時間である。

## 9.2 本論より示された可能性及び今後の研究課題

上記の結論以外にも、いくつかの可能性が明らかとなった。これらは可能性に過ぎず、結論付けるためには今後の研究が必要である。

実験1より、統合の時間窓の範囲外において、立ち上がり時間の割合が0-90%まで増加した時のPATの移動量が、持続時間が長くなるにつれて大きくなる可能性が示された。一方、実験3より、統合の時間窓の範囲外において、持続時間が長くなるにつれ、「立ち上がり時間の割合の増加とPATの移動との関係」は大きく変化しないまま、PATが時間的後方へ移動する可能性が示された。このように、同じ統合の時間窓の範囲外でも異なる傾向が示された。これは周波数の差異に起因する可能性が考えられる。だが、本論では、統合の時間窓の範囲外で持続時間を多くの水準で変化させておらず、周波数も2水準でしか変化させていない。加えて、仮説と直接関係がないため、実験も分析も精密には行えなかった。そのため、周波数と持続時間が立ち上がり時間の増加とPATの変化との関係をどのように変化させるかについては、今後、更なる研究が必要である。

実験2より、参加者ごとの、持続時間ごとの「立ち上がり時間の割合の増加とPATの変化との関係」を用いて、一般群の参加者をグループ分けをしたところ、二つのグループに分けられた。一つのグループは10人中7人が属し、実験1の音楽群と全く異なる、持続時間ごとの「立ち上がり時間の割合の増加とPATの変化との関係」を示した。もう一つのグループは、10人中3人のみが属し、実験1の音楽群と同様の関係を示した。これらの参加者グループごとに異なる参加者の性質を探したが、本論では見つからなかった。少なくとも、小学校までにピアノを習っていたかどうかや、普段の音楽聴取の平均時間やジャンル、リズムゲームの経験、ダンスの経験は参加者グループごとに異ならなかった。従って、音楽群と全く異なる関係を示した一般群の参加者の反応が何に起因するのか、また、一般群は周波数が高い刺激音に対しどのような反応を示すのかなど、一般群のPATの知覚については今後の研究が必要である。

予備実験より、提示音圧レベルを2水準で変化させた時のPATに有意差がなく、提示音圧レベルと持続時間や立ち上がり時間の割合との交互作用も有意ではないことが明らかになった。先行研究では、提示音圧レベルとその他の音の要素との交互作用があると報告されたため、先行研究と一致しない結果となった。その理由として、先行研究では提示音圧レベルを約20dB程度以上変化させていたが、本論では8dBのみ変化させたことが影響した可能性が考えられた。また、提示音圧レベルはPATを移動させるのではなく、統合の時間窓の長さを変化させ、そして統合の時間窓の長さが変化した影響によりPATの知覚が変化した可能性も考えられた。従って、提示音圧レベルがPATの知覚をどのように変化させるのかについては、今後の研究が必要である。

考察にて、持続時間が増加するにつれてPATが移動するかを検討した。立ち上がり時間の割

合の全水準で平均した場合に、持続時間が100 msから250 msまで増加するにつれてPATが約10 ms程度、時間的後方へ移動した。このことより、持続時間が増加するにつれてPATが時間的後方へ僅かに移動する可能性が示唆された。これは本論の仮説に直接関わることではないため、立ち上がり時間の割合を平均した場合の分析の結果より示唆されたことであった。従って、結論づけるためには、より精密に実験を行う必要があり、今後の研究が必要である。

また、考察にて、周波数が増加するにつれPATが移動するのかを検討した。周波数を440 Hz(実験1)と3520 Hz(実験3)とで、持続時間に寄らず、PATが10 ms程度、時間的後方へ移動することが明らかとなった。このことより、周波数が高くなるとPATが時間的後方へ僅かに移動する可能性が示唆された。だが、本論の仮説に直接関わることではないため、立ち上がり時間の割合の全水準で平均した場合の分析の結果より示唆されたことであった。従って、結論づけるためには、より精密に実験を行う必要があり、今後の研究が必要である。加えて、本論では周波数を2水準のみで変化させたため、周波数の変化がどのようにPATを移動させるかについても更なる研究が必要である。

3つの実験より、タッピング課題より得られた各参加者の各刺激音のPATの分布を確認したところ、殆どの分布は、両端が細くなる(正規分布のような)形を示したが、まれに、長方形に近い(尖度が負のような)形や、左右非対称の(歪度の絶対値が大きいような)形もあった。本論では分布の算出以外の分析は行わなかった。だが、PATが確率密度分布に従うという考え方もあるため、刺激音が異なるときのPATの分布の変化や、測定課題によるPATの分布の差異などは今後の研究が必要である。

さらに、本論では「PAT知覚は統合の時間窓の影響を受ける。その影響により、統合の時間窓の範囲内と範囲外では、立ち上がり時間の割合の増加に対する $PAT(\Delta PC_{A,R})$ の変化量と、曲率が異なる」との結論が得られたが、なぜこのようなことが生じるのかという生理学的な根拠は調査していない。そのため、今後、聴覚生理学的な研究が行われることを期待する。

最後に、合成音と楽器音とに問わず、音のどこにPATがあるのかが明らかになったならば、音楽の自動生成にPATを揃える機能が組み込まれることで、より人間の知覚に沿った音楽を生成することができるかもしれない。あるいは、演奏分析において各音のPATを自動的に明らかにすることで、より人間の知覚に沿った演奏の在り様を分析ことができるようになるかもしれない。今後のPATの研究の発展を期待する。

# 謝辞

本研究を進め、また本論を書くにあたり、多くの方々にご指導及びご協力を賜りました。

東京藝術大学音楽環境創造科の丸井淳史先生には、指導教員として丁寧にご指導、ご鞭撻いただきました。研究の進みが遅かったり研究の方向性が定まらなかったりした時期にも温かく、また親身にご指導いただきました。丸井先生の指導の下であったから本論が完成したと感じております。心より厚く御礼申し上げます。

東京藝術大学音楽環境創造科の亀川徹先生には、度々、的確なご指導と激励をいただきました。特に、投稿論文のチェックや研究進捗報告の際には、研究に不可欠な鋭いご意見をいただきました。心より御礼申し上げます。

元、東京藝術大学音楽環境創造科の西岡龍彦先生には、本研究を進めるにあたり、音楽の観点からの鋭いご意見をいただきました。音楽の観点を忘れがちであった時に、何度も音楽に立ち返るきっかけをいただきました。心より御礼申し上げます。

東京藝術大学音楽環境創造科の後藤英先生には、本研究を進めるにあたり、多くのご意見をいただきました。特に、音楽や楽器音での本研究の立ち位置を深く考えるきっかけをいただきました。心より御礼申し上げます。

東京藝術大学音楽環境創造科の田村文生先生には、本研究を進めるにあたり、多くのご意見をいただきました。素朴な疑問として本研究の真意を突いてきたため、本研究の根幹を考え直すきっかけをいただきました。心より御礼申し上げます。

東京藝術大学音楽文化学専攻音楽教育の山下薫子先生には、お忙しいところ博士論文審査会に参席いただき、異なる視点から、それまで気が付いていなかった重要なことをご指摘いただきました。心より御礼申し上げます。

そして、心理実験の参加者としてご協力いただきました皆さまにも、心より御礼申し上げます。特に最後の1年半は、コロナ禍にもかかわらず快く実験に参加していただきましたこと、深謝しております。

最後に、長い大学院生活の間、見守り続けてくれた家族に、この場を借りて感謝申し上げます。

本研究ならびに筆者をとりまく全ての方々に、心より御礼申し上げます。

本論の完成をもって、少しでも恩返しのできたのなら幸甚の限りです。

## 参考文献

- [1] Bechtold, T. A. and Senn, O. (2018) “Articulation and Dynamics Influence the Perceptual Attack Time of Saxophone Sounds,” *Frontiers in Psychology*, 9 1–11.
- [2] Camp, M.-A., Kilchenmann, L., Volken, T., and Senn, O. (2011) On measuring and interpreting microtiming 95–110.
- [3] Collins, N. (2006) “Investigating Computational Models of Perceptual Attack Time,” *Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9)*, 923–929.
- [4] Cowan, N. (1984) “On short and long auditory stores,” *Psychological Bulletin*, 96 341–370.
- [5] Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M. R., and London, J. (2019) “Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds,” *Journal of Experimental Psychology: Human Perception and Performance*, 45 402–418.
- [6] Eddins, D. A. and Green, D. M. (1995) “temporal integration and temporal resolution,” in *Handbook of Perception and Cognition: Hearing*, San Diego: Academic Press, p. 207–242.
- [7] Garner, W. R. and Miller, G. A. (1947) “The masked threshold of pure tones as a function of duration,” *Journal of Experimental Psychology*, 37 293–303.
- [8] Gordon, J. W. (1987) “The perceptual attack time of musical tones,” *The Journal of the Acoustical Society of America*, 82 88–105.
- [9] Gordon, J. W. (1984) “Perception of Attack Transients in Musical Tones,” Ph.D. dissertation.
- [10] Hirsh, I. J. (1959) “Auditory Perception of Temporal Order,” *Journal of the Acoustical Society of America*, 31 759–767.
- [11] Hirsh, I. J. and Sherrick, C. E. (1961) “Perceived order in different sense modalities,” *Journal of Experimental Psychology*, 62 423–432.
- [12] Janker, P. M. (1996) “The range of subjective simultaneousness in tapping experiments with speech stimuli,” in *Workshop on the Auditory Basis of Speech Perception* Keele, England, UK July 15-19, 1996, Janker1996.
- [13] JIS Z 8106:2000, (2000) 『音響用語』, 日本工業規格.
- [14] Kazazis, S., Esterer, N., Depalle, P., and Mcadams, S. (2017) “A Performance Evaluation of the Timbre Toolbox and the MIRtoolbox on Calibrated Test Sounds,” 2017 *International Symposium on Musical Acoustics*, 1–4.

- [15] London, J., Nymoén, K., Langerød, M. T., Thompson, M. R., Code, D. L., and Danielsen, A. (2019) “A comparison of methods for investigating the perceptual center of musical sounds,” *Attention, Perception, and Psychophysics*, 81 2088–2101.
- [16] Miller, G. A. (1948) “The Perception of Short Bursts of Noise,” *Journal of the Acoustical Society of America*, 20 160–170.
- [17] Morton, J., Marcus, S., and Frankish, C. (1976) “Perceptual centers (P-centers),” *Technical Report 5* 405–408.
- [18] Munson, W. A. (1947) “The Growth of Auditory Sensation,” *Journal of the Acoustical Society of America*, 19 584–591.
- [19] Nymoén, K., Danielsen, A., and London, J. (2017) “Validating Attack Phase Descriptors Obtained by the Timbre Toolbox and MIRToolbox,” *14th Sound and Music Computing Conference*, 214–219.
- [20] Obando Leitón, M. E. (2019) “Temporal integration in cochlear implants and the effect of high pulse rates,” Ph.D. dissertation.
- [21] Plomp, R. and Bouman, M. A. (1959) “Relation between hearing threshold and duration for tone pulses,” *Journal of the Acoustical Society of America*, 31 749–758.
- [22] Polfreman, R. (2013) “Comparing Onset Detection & Perceptual Attack Time,” in the *14th International Society for Music Information Retrieval Conference; ISMIR 2013*.
- [23] Repp, B. H. (2005) “Sensorimotor synchronization: A review of the tapping literature,” *Psychonomic Bulletin and Review*, 12 969–992.
- [24] Repp, B. H. and Su, Y. H. (2013) “Sensorimotor synchronization: A review of recent research (2006-2012),” *Psychonomic Bulletin and Review*, 20 403–452.
- [25] Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., and Münte, T. F. (2001) “Event-related brain potentials to sound omissions differ in musicians and non-musicians,” *Neuroscience Letters*, 308 33–36.
- [26] Scharf, B. (1978) “Loudness,” in *Handbook of perception: Vol.IV Hearing*, San Diego: Academic Press, 187-242.
- [27] Scott, S. K. (1998) “The point of P-centres,” *Psychological Research*, 61 4–11.
- [28] Stevens, J. C. and Hall, J. W. (1966) “Brightness and loudness as functions of stimulus duration,” *Perception & Psychophysics*, 1 319–327.
- [29] Suzuki, Y. and Takeshima, H. (2004) “Equal-loudness-level contours for pure tones,” *The Journal of the Acoustical Society of America*, 116 918–933.
- [30] Takeshima, H. (1988) “Growth of the loudness of a tone burst with a duration up to 10 seconds,” *The Journal of the Acoustical Society of America of Japan*, 9.
- [31] Viemeister, N. F. and Wakefield, G. H. (1991) “Temporal integration and multiple looks,” *Journal of the Acoustical Society of America*, 90 858–865.
- [32] Villing, R. C., Repp, B. H., Ward, T. E., and Timoney, J. M. (2011) “Measuring perceptual centers using the phase correction response,” *Attention, Perception, and Psychophysics*, 73 1614–1629.

- [33] Villing, R., Ward, T., and Timoney, J. (2003) “P-Centre Extraction from Speech: the need for a more reliable measure,” Irish Signals and Systems Conference, 5–10.
- [34] Vos, J. and Rasch, R. (1981) “The perceptual onset of musical tones,” *Perception & Psychophysics*, 29 323–335.
- [35] Vos, P., Mates, J., and van Kruysbergen, N. W. (1995) “The perceptual center of a stimulus as the cue for synchronisation to a metronome: Evidence from asynchronies,” *Quarterly Journal of Experimental Psychology*, 48A 1024–1040.
- [36] Warren, B. M. and Ackroff, J. M. (1976) “Two types of auditory sequence perception,” *Perception & Psychophysics*, 20 387–394.
- [37] Wright, M. J. (2008) “The shape of an instant: Measuring and modeling Perceptual Attack Time with probability density functions (if a tree falls in the forest, when did 57 people hear it make a sound?).,” Ph.D. dissertation.
- [38] Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R. J., and Näätänen, R. (1998) “Temporal window of integration of auditory information in the human brain,” *Psychophysiology*, 35 615–619.
- [39] 玉越勢治, (2011) 「脳波を指標とした聴覚情報処理過程に関する基礎及び応用研究」, 博士論文.
- [40] 玉越勢治・溝部光希・箕浦菜々子・片山順一・八木昭宏, (2011) 「ミスマッチ陰性電位を指標とした時間統合窓の基本特性の検討」, 『生理心理学と精神生理学』, 第 29 巻, 5–12 頁.
- [41] 中島祥好, (1979) 「短音で示された分割時間の精神物理学的研究」, 『日本音響学会誌』, 第 35 巻, 145–151 頁.
- [42] 播磨敏雄, (2015) 「2 音源により生じる音像と先行音効果の関係に関する研究」, 博士論文, 秋田県立大学.
- [43] 萩原啓二, (1969) 「聴覚の時間的積分機構に関する研究」, 『日本耳鼻咽喉科学会会報』, 第 72 巻, 1061–1096 頁.

# 付録A 実験1: 各参加者の各刺激音における同調のズレ量のバイオリンプロット

各参加者の各刺激音における同調のズレ量のバイオリンプロットを掲載した。バイオリンプロットとは、ヒストグラムの平滑化の方法の一つであり、横向きヒストグラムをカーネル密度推定を用いて平滑化したものである。横軸の“stimuli”は刺激番号を表し、縦軸の“Asynchronies”は同調のズレ量を表す。刺激番号と、刺激音の持続時間や立ち上がり時間の割合との対応は、表A.1を参照してほしい。

表 A.1: 実験1の刺激番号と、持続時間及び立ち上がり時間の割合 (及び立ち上がり時間) との対応表

刺激番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
持続時間 [ms]	100	100	100	100	100	100	100	100	100	100	150	150	150	150	150	150	150	150	150	150
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	10	20	30	40	50	60	70	80	90	0	15	30	45	60	75	90	105	120	135
刺激番号	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
持続時間 [ms]	200	200	200	200	200	200	200	200	200	200	250	250	250	250	250	250	250	250	250	250
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	20	40	60	80	100	120	140	160	180	0	25	50	75	100	125	150	175	200	225
刺激番号	41	42	43	44	45	46	47	48	49	50										
持続時間 [ms]	300	300	300	300	300	300	300	300	300	300										
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90										
(立ち上がり時間 [ms])	0	30	60	90	120	150	180	210	240	270										

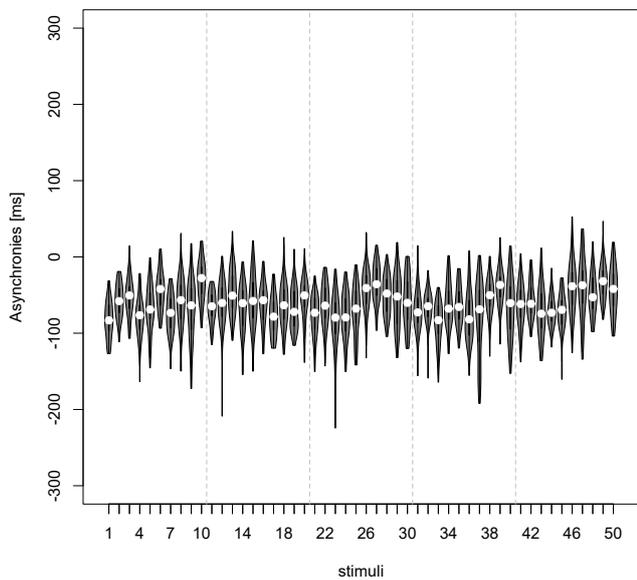


図 A.1: 参加者 A

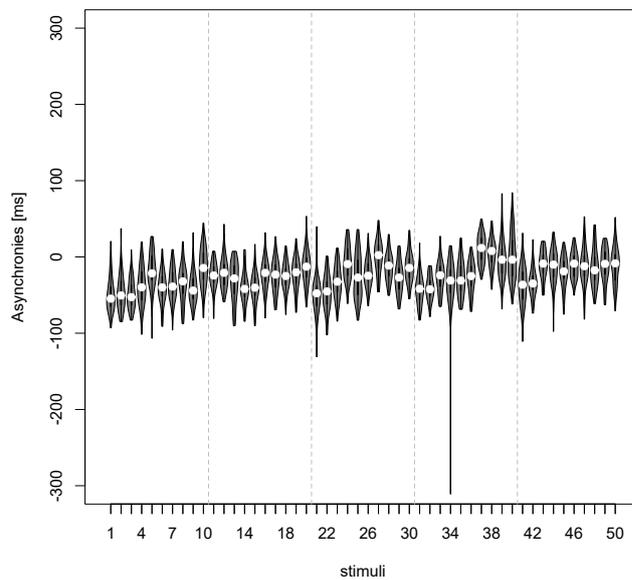


図 A.2: 参加者 B

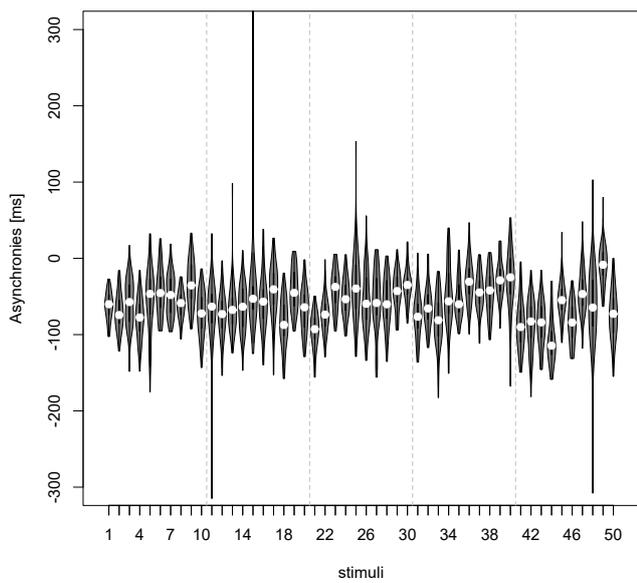


図 A.3: 参加者 C

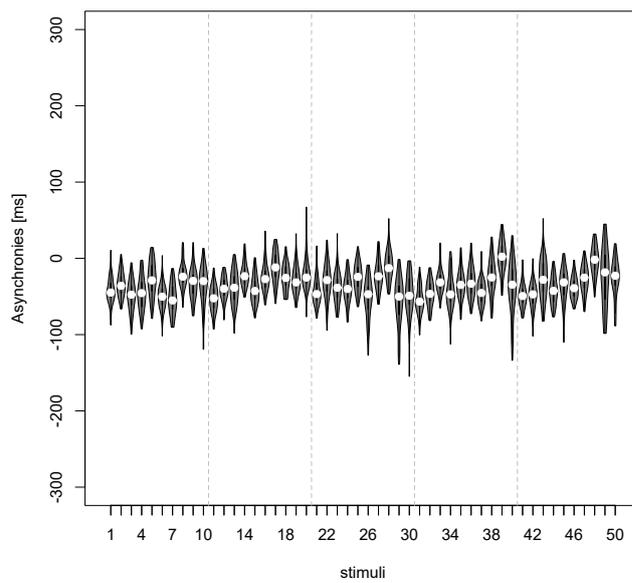


図 A.4: 参加者 D

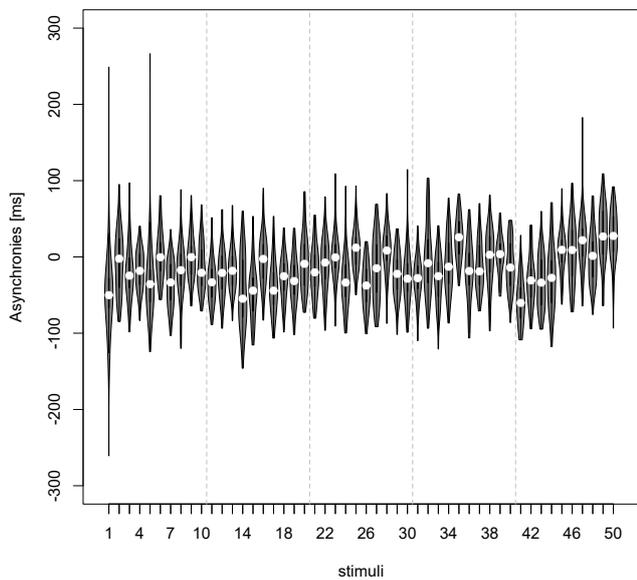


図 A.5: 参加者 E

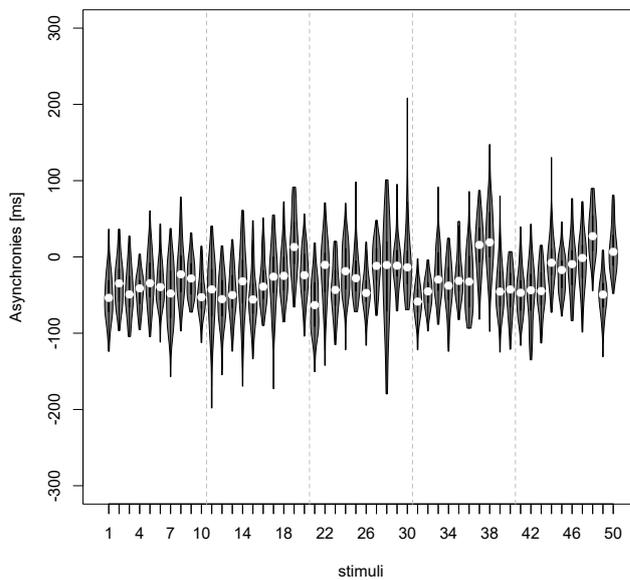


図 A.6: 参加者 F

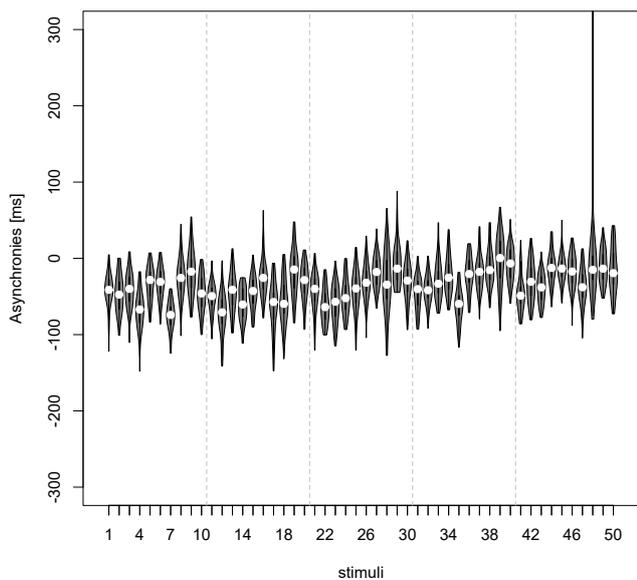


図 A.7: 参加者 G

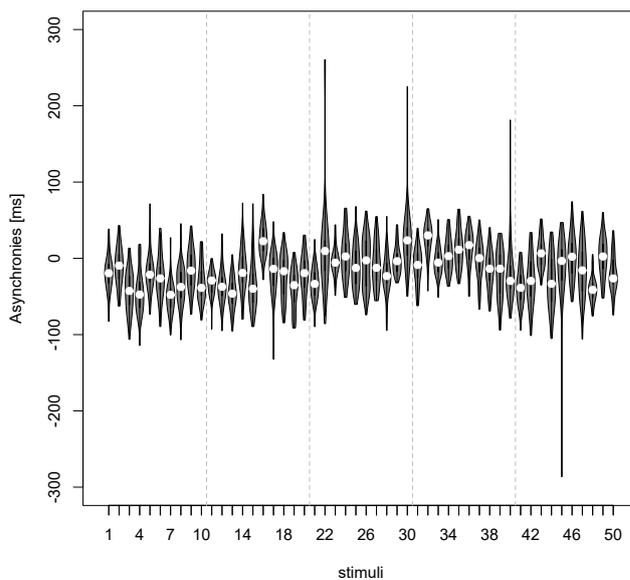


図 A.8: 参加者 Y

# 付録B 実験2: 各参加者の各刺激音における同調のズレ量のバイオリンプロット

各参加者の各刺激音における同調のズレ量のバイオリンプロットを掲載した。バイオリンプロットとは、ヒストグラムの平滑化の方法の一つであり、横向きのヒストグラムをカーネル密度推定を用いて平滑化したものである。横軸の“stimuli”は刺激番号を表し、縦軸の“Asynchronies”は同調のズレ量を表す。刺激番号と、刺激音の持続時間や立ち上がり時間の割合との対応は、表 B.1 を参照してほしい。

表 B.1: 実験2の刺激番号と、持続時間及び立ち上がり時間の割合 (及び立ち上がり時間) との対応表

刺激番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
持続時間 [ms]	100	100	100	100	100	100	100	100	100	100	150	150	150	150	150	150	150	150	150	150
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	10	20	30	40	50	60	70	80	90	0	15	30	45	60	75	90	105	120	135
刺激番号	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
持続時間 [ms]	200	200	200	200	200	200	200	200	200	200	250	250	250	250	250	250	250	250	250	250
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	20	40	60	80	100	120	140	160	180	0	25	50	75	100	125	150	175	200	225

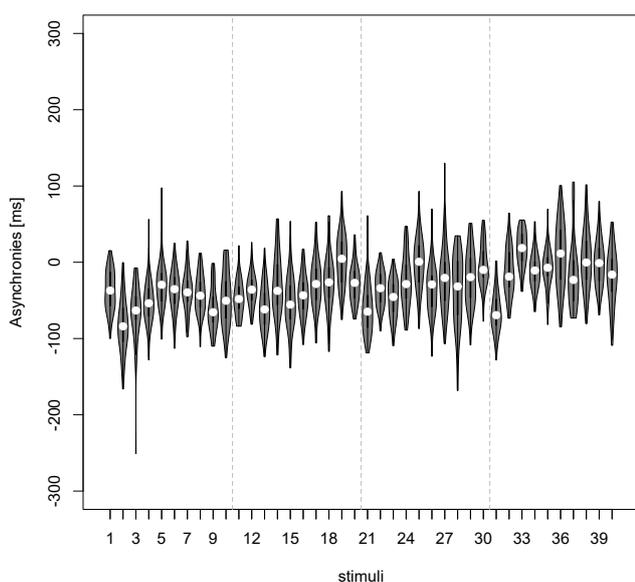


図 B.1: 参加者 H

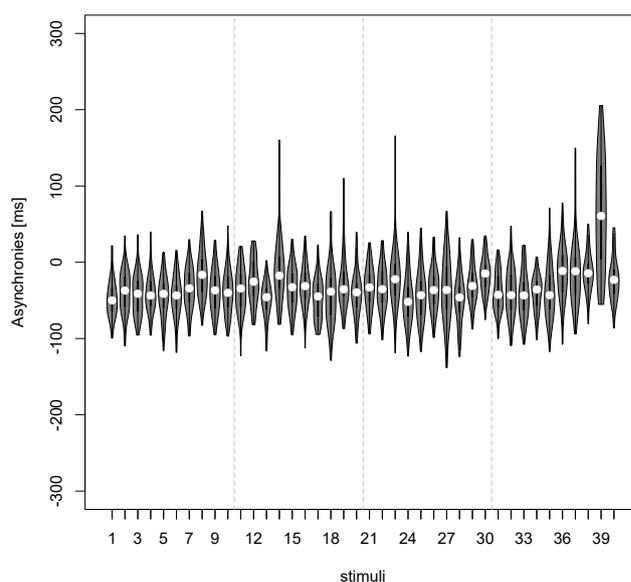


図 B.2: 参加者 I

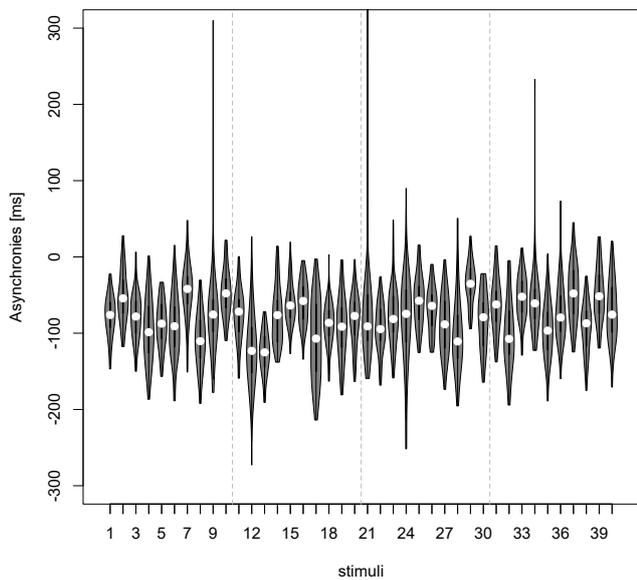


図 B.3: 参加者 J

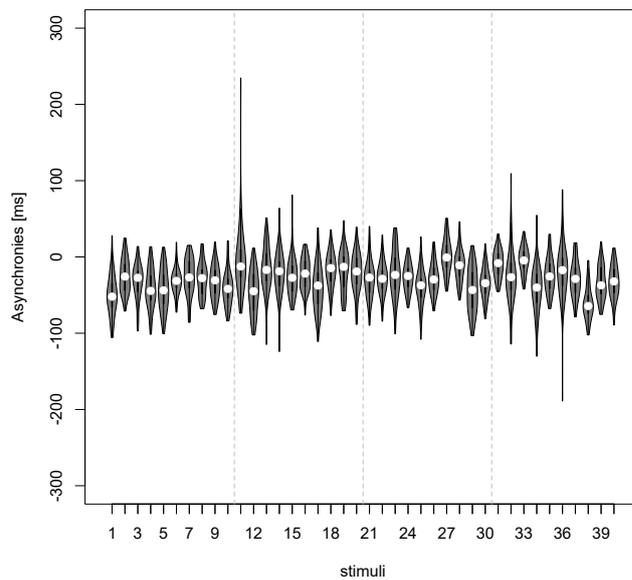


図 B.4: 参加者 K

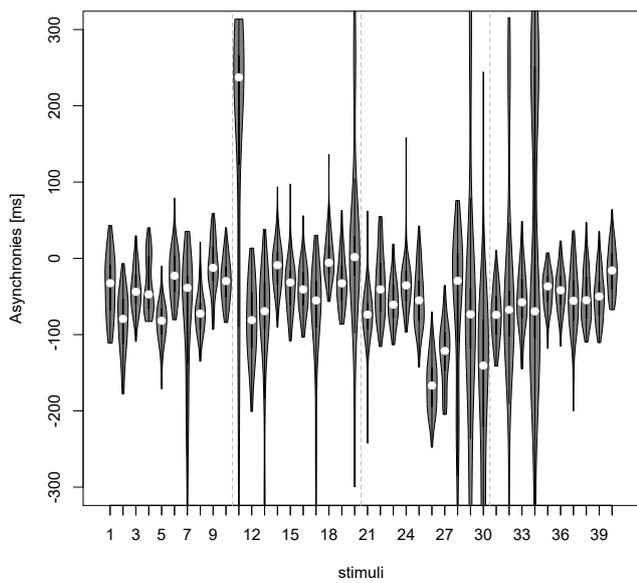


図 B.5: 参加者 L

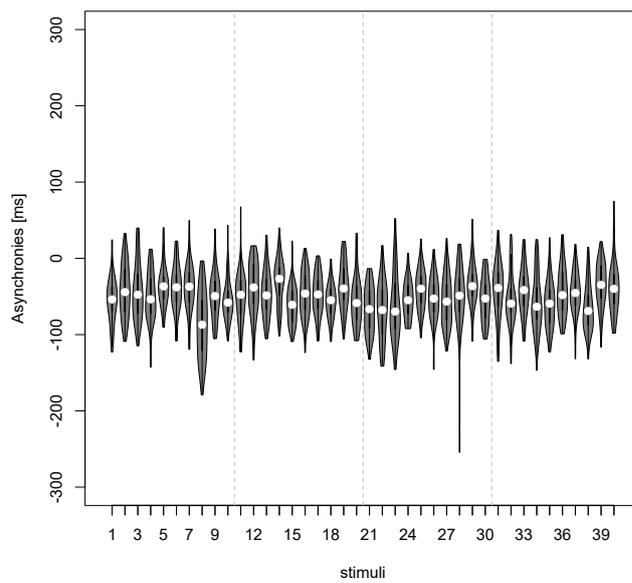


図 B.6: 参加者 M

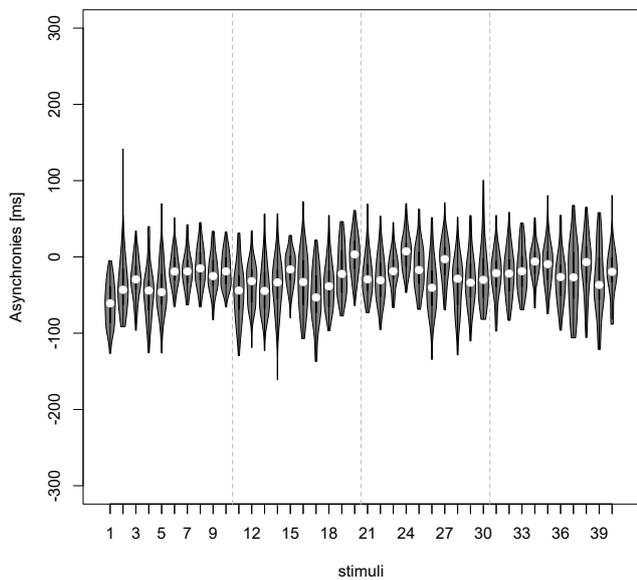


図 B.7: 参加者 N

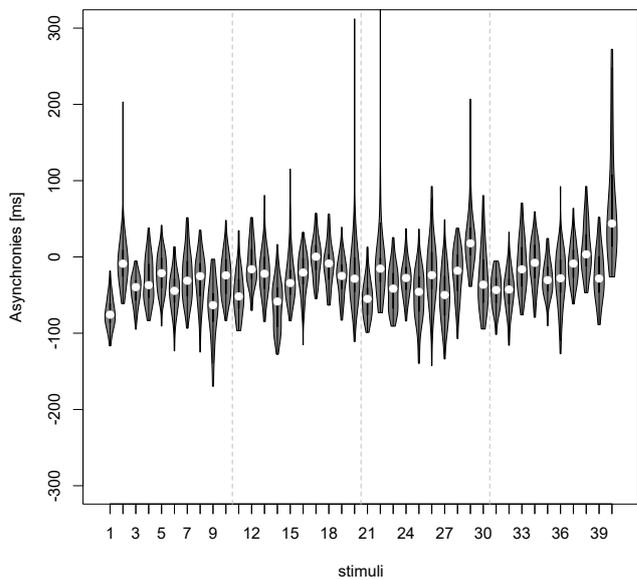


図 B.8: 参加者 O

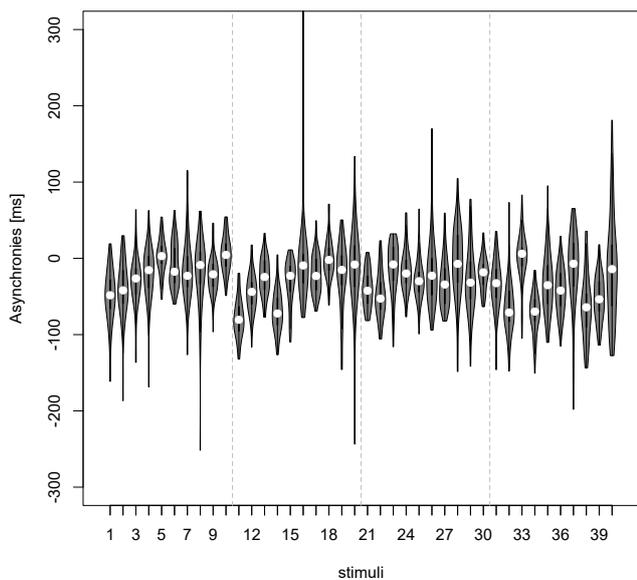


図 B.9: 参加者 P

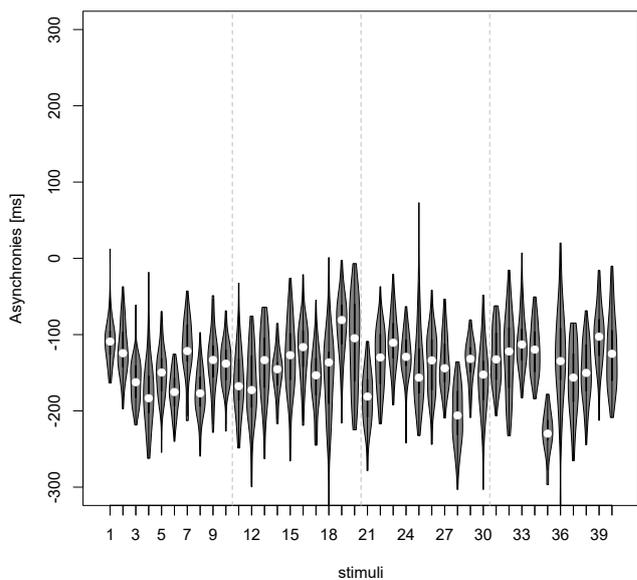


図 B.10: 参加者 Q

# 付録C 実験3: 各参加者の各刺激音における同調のズレ量のバイオリンプロット

各参加者の各刺激音における同調のズレ量のバイオリンプロットを掲載した。バイオリンプロットとは、ヒストグラムの平滑化の方法の一つであり、横向きヒストグラムをカーネル密度推定を用いて平滑化したものである。横軸の“stimuli”は刺激番号を表し、縦軸の“Asynchronies”は同調のズレ量を表す。刺激番号と、刺激音の持続時間や立ち上がり時間の割合との対応は、表 C.1 を参照してほしい。

表 C.1: 実験3の刺激番号と、持続時間及び立ち上がり時間の割合 (及び立ち上がり時間) との対応表

刺激番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
持続時間 [ms]	100	100	100	100	100	100	100	100	100	100	150	150	150	150	150	150	150	150	150	150
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	10	20	30	40	50	60	70	80	90	0	15	30	45	60	75	90	105	120	135
刺激番号	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
持続時間 [ms]	200	200	200	200	200	200	200	200	200	200	250	250	250	250	250	250	250	250	250	250
立ち上がり時間の割合 [%]	0	10	20	30	40	50	60	70	80	90	0	10	20	30	40	50	60	70	80	90
(立ち上がり時間 [ms])	0	20	40	60	80	100	120	140	160	180	0	25	50	75	100	125	150	175	200	225

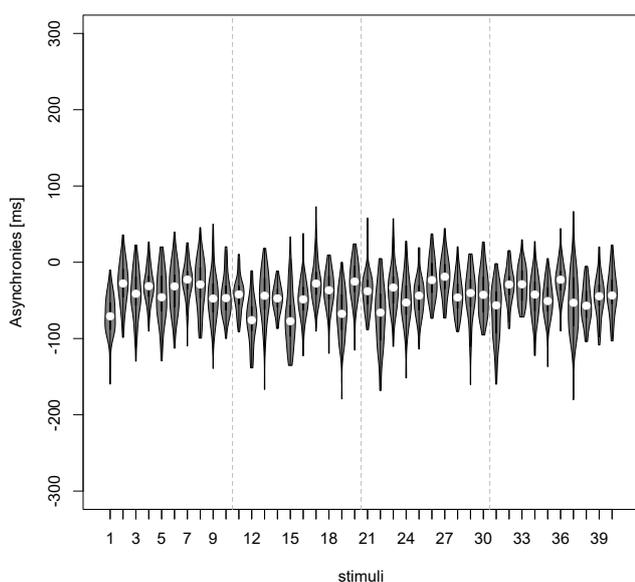


図 C.1: 参加者 A

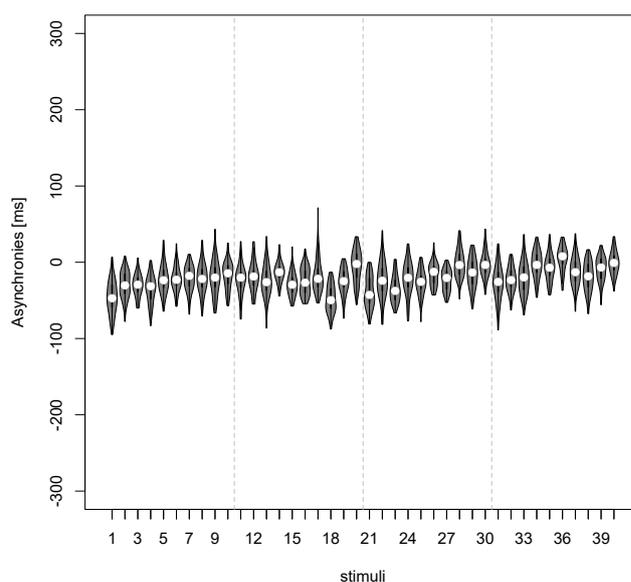


図 C.2: 参加者 B

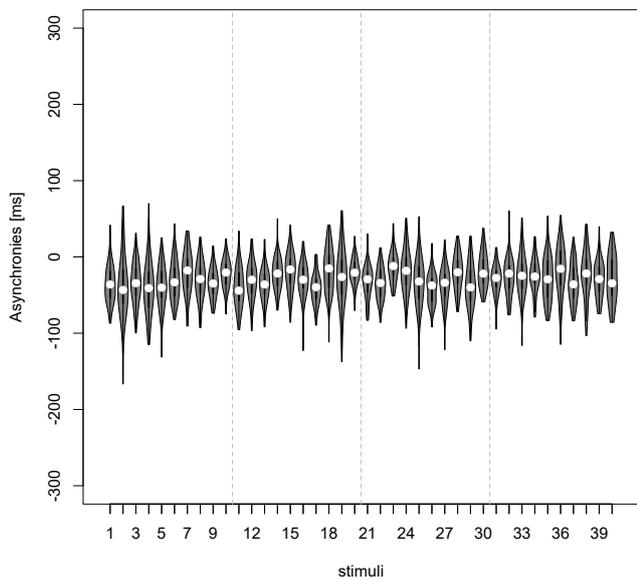


図 C.3: 参加者 C

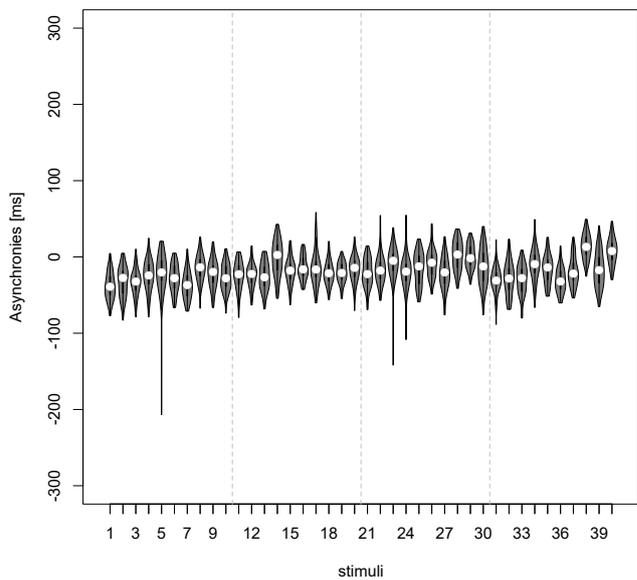


図 C.4: 参加者 D

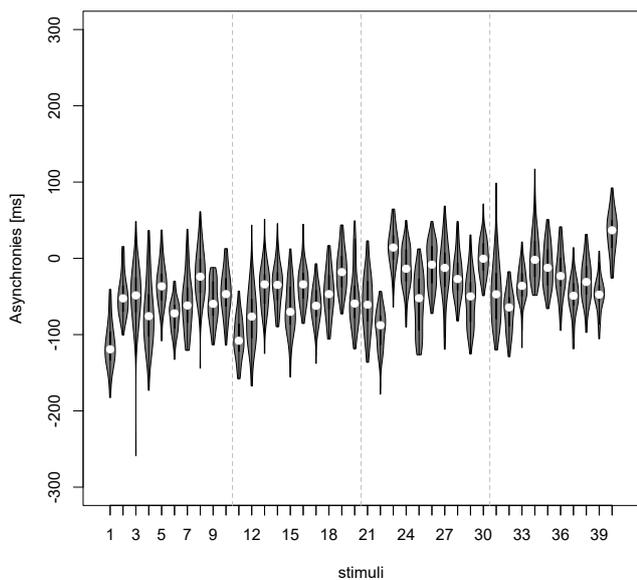


図 C.5: 参加者 E

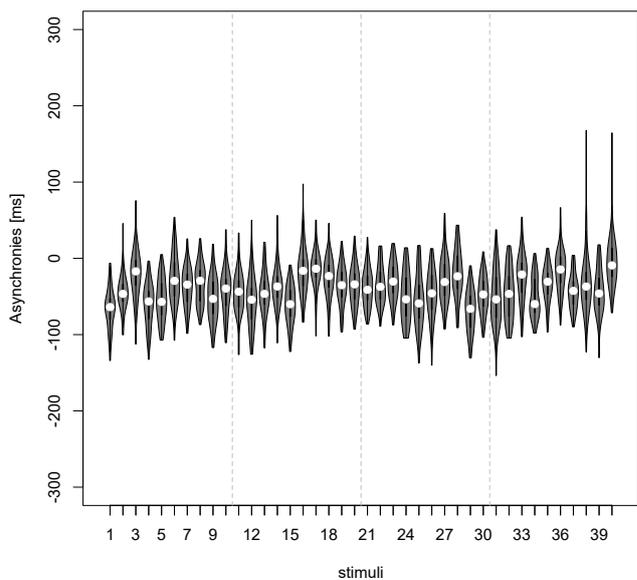


図 C.6: 参加者 F

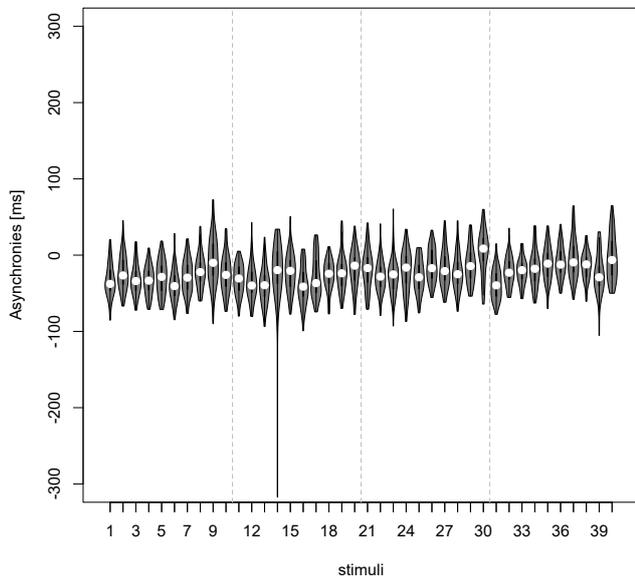


图 C.7: 参加者 G

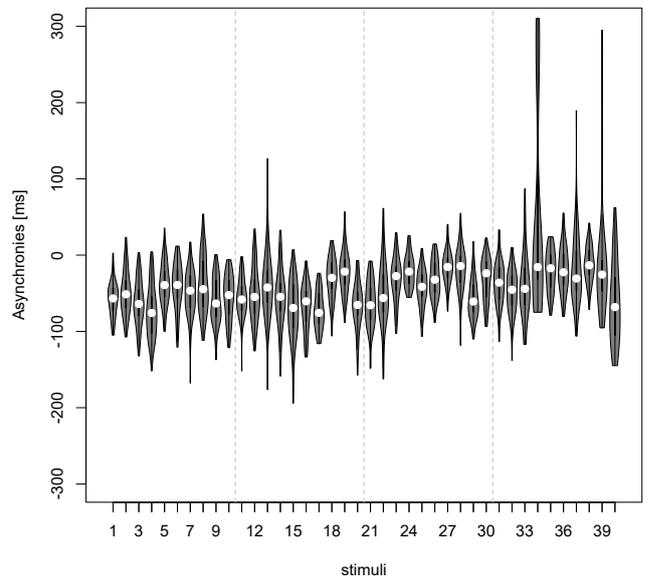


图 C.8: 参加者 Z

## 付録D 各実験における測定された $\Delta PC_{A,D100R0}$ の値と近似した4パラ メータロジスティック回帰モデ ルのパラメータ値

本論の各実験における、 $\Delta PC_{A,D100R0}$  より測定された  $\Delta PC_{A,D100R0}$  の値と、持続時間ごとの「立ち上がり時間の割合の増加と  $\Delta PC_{A,D100R0}$  の変化との関係」を近似した4パラメータロジスティック回帰モデルのパラメータ値を掲載した。

## 測定された $\Delta PC_{A,D100R0}$ の値

実験意図に沿わないデータを削除し、全参加者で平均をとった、各刺激音に対する  $\Delta PC_{A,D100R0}$  の値を掲載した

表 D.1: 実験 1

	持続時間 [ms]					
	100	150	200	250	300	
0	0.000	3.987	-2.286	1.970	-2.565	
10	12.506	2.920	12.724	15.653	4.503	
20	5.578	8.642	12.926	12.543	12.774	
立ち上がり	30	-1.496	7.588	16.298	19.108	11.722
時間の割合	40	15.266	4.712	22.840	21.864	24.237
[%]	50	16.507	24.937	13.872	23.025	25.953
	60	-2.329	12.243	29.808	28.200	30.384
	70	16.199	10.022	26.324	35.834	32.583
	80	20.786	21.854	22.493	34.192	37.513
	90	13.894	20.542	25.562	23.745	31.446

表 D.2: 実験 3

	持続時間 [ms]					
	100	150	200	250	300	
0	0.000	11.506	18.451	16.925	—	
10	20.526	11.004	13.739	20.399	—	
20	19.851	21.047	36.558	29.145	—	
立ち上がり	30	11.488	28.925	30.306	35.892	—
時間の割合	40	22.423	12.640	19.555	35.187	—
[%]	50	20.161	22.803	33.818	41.508	—
	60	23.041	23.610	35.260	28.167	—
	70	31.027	26.314	38.783	35.950	—
	80	20.175	28.101	21.613	26.169	—
	90	22.833	27.515	38.895	45.353	—

表 D.3: 実験 2

(a) 参加者 3 人のグループ

	持続時間 [ms]					
	100	150	200	250	300	
0	0.000	-5.636	-0.057	5.159	—	
10	8.739	22.473	19.507	12.071	—	
20	7.300	18.217	19.058	55.299	—	
立ち上がり	30	18.327	-1.763	29.111	22.169	—
時間の割合	40	36.236	16.294	26.059	29.670	—
[%]	50	22.057	26.490	27.854	32.120	—
	60	26.992	35.951	18.040	40.638	—
	70	24.701	41.025	32.315	35.929	—
	80	2.147	38.267	44.480	26.731	—
	90	29.745	33.122	35.277	59.025	—

(b) 参加者 7 人のグループ

	持続時間 [ms]					
	100	150	200	250	300	
0	0.000	6.525	-6.342	8.131	—	
10	5.658	-7.880	3.306	-0.090	—	
20	1.239	-4.999	10.309	16.087	—	
立ち上がり	30	-4.694	15.008	11.353	0.816	—
時間の割合	40	-8.007	11.278	8.638	-2.969	—
[%]	50	3.518	12.079	-10.359	15.111	—
	60	15.684	-8.188	-0.468	11.413	—
	70	-7.689	8.718	-1.329	0.100	—
	80	11.360	18.495	8.394	28.607	—
	90	9.642	18.145	-3.295	16.193	—

## 持続時間ごとの「立ち上がり時間の割合の増加と $\Delta PC_{A,D100R0}$ の変化との関係」を近似した4パラメータロジスティック回帰モデルのパラメータ値

4パラメータロジスティック回帰モデルの式を再掲する。従属変数を  $y$  とし、独立変数を  $x$ 、最小値を  $c$ 、最大値を  $d$ 、傾きを  $b$ 、中央の位置を  $e$  とする。

$$y = c + \frac{d - c}{1 + \exp(b(x - e))}$$

表 D.4: 実験 1

持続時間 [ms]	b	c	d	e
100	-0.018	-1.375	39.695	102.838
150	-0.018	-10.957	33.306	42.001
200	-0.043	-356.702	25.601	-60.801
250	-0.037	-37.336	32.528	-10.133
300	-0.052	-13.997	36.239	21.755

表 D.5: 実験 3

持続時間 [ms]	b	c	d	e
100	-0.801	-0.161	21.375	6.020
150	-0.023	-202.310	29.828	-108.449
200	-1.743	16.095	31.849	14.828
250	-0.235	16.812	35.544	16.741
—	—	—	—	—

表 D.6: 実験 2

(a) 参加者 3 人のグループ

持続時間 [ms]	b	c	d	e
100	-0.344	4.559	23.520	26.195
150	-0.153	8.754	37.396	70.804
200	-0.023	-69.758	34.500	-39.886
250	-0.467	5.154	37.698	27.802
—	—	—	—	—

(b) 参加者 7 人のグループ

持続時間 [ms]	b	c	d	e
100	-0.068	-0.041	19.432	88.253
150	-0.696	3.403	18.319	105.848
200	0.724	-1.410	5.448	89.322
250	-0.523	6.072	22.399	187.462
—	—	—	—	—