

2020年度 学士論文

聴覚-触力覚刺激の特徴量を変化させた時の
時間統合窓の変動

Variation in the window of temporal integration with
changing attributes of audio-tactile stimuli

2016年入学
学籍番号 2116232

佐藤来

Rai Sato

令和2年12月04日

東京芸術大学音楽学部音楽環境創造科

Department of Music Creativity and the Environment,
Faculty of Music,
Tokyo University of the Arts

概要

ゲームや映画などのデジタルコンテンツでは、様々な感覚情報が合わさったマルチモーダル刺激を提示することで、臨場感の向上など体験の質を上げるための工夫が試みられている。しかしながら、マルチモーダル感覚が我々の脳内でどう処理され、そのコンテンツが認知されるのかは未だ研究段階であるため、コンテンツ体験の質を上げるためにどのようにマルチモーダル刺激を呈示すれば良いか、ヒトの認知システムに基づいた具体的な体験設計指標はまだない。

そこで筆者は、音と振動を使ったマルチモーダル刺激の統合処理メカニズムを解明するために、いくつかの実験を行った。本論文では、スマートフォンを持った状態で映画やゲームを体験している際に、音に合わせて振動が同時に呈示される状況で、音と振動の時間統合窓がどのような振る舞いを見せるのかを心理物理学的手法を用いて検討した。

聴覚刺激の周波数とノイズ量をそれぞれ変化させた実験では、聴覚刺激の周波数が低い場合、時間統合窓が小さくなる傾向があることが分かり、実験参加者が聴覚-触力覚刺激の周波数関係によって刺激同士が混ざり合ったと知覚された可能性が示唆された。また、聴覚-触力覚の同時性知覚は、両耳聴下での同時性知覚からその仕組みを説明できる可能性を得られた。

聴覚-触力覚刺激の周波数間隔を変化させた実験では、周波数間隔の増加に伴い時間統合窓が若干小さくなる傾向が見られた。一方で本実験結果は上述の結果と相反するものであり、現在一般に使われる実験手法では、実験参加者の判断バイアスを多分に含まれる可能性が示唆された。

聴覚刺激の強度を変化させた実験では、聴覚強度条件の違いにより時間統合窓の振る舞いに差異が生じることが分かった。また聴覚-触力覚における同時性知覚判断において、刺激の呈示順序にかかわらず両刺激のいずれの判断が同じ傾向を見せる実験参加者グループと、聴覚刺激が触力覚刺激より後に呈示されていた場合、そうでない場合と比べ両者にずれがあると知覚するのが遅くなる傾向を持つグループが存在することがわかった。

これらの結果により、聴覚刺激の周波数と強度を変化させた際に、時間統合窓がそれぞれ変動することが分かった。しかしながら、具体的な特徴量と時間統合窓の大きさの関係までを調査することはできなかった。

本研究は、聴覚-触力覚マルチモーダル刺激を呈示した時の時間統合窓の変動を解明かすまでの第一歩となつたが、時間統合窓の振る舞いを調査する実験手法には未だ改善の余地があることが分かった。

Abstract

The current digital content, such as a game and film, has functions about making users' experiences improve by providing multimodal stimuli combined several sensory information; however, it is not clear how we process the signals and how we perceive this multimodal content in our brain. Therefore, there is no design guideline to provide multimodal information to improve the quality of experience based on the human perception system.

The author has done several experiments to elucidate the integration mechanism between auditory and vibrotactile multimodal stimuli. This thesis investigates how the window of temporal integration of sound and vibration behaves employing psychophysical methods. All of the experiments assumed the situation that vibrations and sounds were provided to the research participants simultaneously while they were watching a film or playing a game holding a smartphone with both hands.

In the experiment to control the auditory stimulus's frequency and amount of noise level, it was found that the frequency parameter is correlated with the length of the window of temporal integration. This result suggests that the participants perceived both stimuli are combined due to the frequency relation between audio-tactile stimuli. Furthermore, These findings may suggest that the mechanism of simultaneity perception allows explaining from the mechanism of dichotic listening.

In the experiment to manipulate the frequency separation of audio-tactile stimuli, a slight decrease in the window of temporal integration was observed as the frequency interval increased. However, this result was incompatible with my previous study; it may indicate that the current experimental method to investigate the window of temporal integration considerably includes participants' judgment bias.

In the experiment to changing the auditory stimulus's intensity, the difference of the intensity affects the behavior variation of the window of temporal integration. Furthermore, the result suggests that there are two groups on the simultaneity perception judgment, which are the group to be able to perceive the stimuli's delay almost in the same way both the sound was presented first and the vibration was, and the one who tended to become more insensitive

to the delay judgment when the vibration was presented first.

According to the results, to change the auditory stimulus's frequency and intensity was concerned with the variation in the window of temporal integration. However, This thesis could not investigate the relationship between the attributes and the window distinctly.

This study will be the first step to discover the behavior of the window of temporal integration; yet, the experimental method to investigate the window still has room for improvement.

目次

第 1 章	序論	1
第 2 章	研究の背景と射程	3
2.1	マルチモーダル	3
2.1.1	マルチモーダル知覚	4
2.1.2	同時性知覚	6
2.2	聴覚-触力覚マルチモーダル知覚	10
2.2.1	感覚器の構造	11
2.2.2	聴覚-触力覚マルチモーダルの実例	14
2.2.3	聴覚-触力覚マルチモーダルの問題点	19
2.3	本論文の射程	20
第 3 章	聴覚刺激の周波数とノイズ量を変化させた時の同時性知覚の変動	23
3.1	背景	23
3.2	実験の概要	24
3.2.1	実験参加者	24
3.2.2	実験装置	24
3.2.3	実験刺激	24
3.2.4	実験手続き	26
3.3	結果	27
3.4	考察	29
3.4.1	聴覚の処理過程と主観的同時点の関係	29
3.4.2	聴覚刺激の周波数と丁度可知差異の関係	30
3.5	結論	30
第 4 章	聴覚-触力覚刺激の周波数間隔を変化させた時の時間統合窓の変動	31
4.1	背景	31
4.2	実験手法	33

4.2.1	実験参加者	33
4.2.2	実験装置	33
4.2.3	実験刺激	33
4.2.4	手続き	36
4.3	結果	36
4.4	考察	37
4.4.1	時間統合窓の「V字型」傾向	37
4.4.2	実験方法における判断バイアス	39
4.5	結論	40
第 5 章	聴覚刺激の強度を変化させた時の時間統合窓の変動	41
5.1	背景	41
5.1.1	信号検出理論	42
5.1.2	ベイズ統計学とベイズ推定による信号検出モデル	44
5.2	実験手法	48
5.2.1	実験参加者	48
5.2.2	実験装置	48
5.2.3	実験刺激	48
5.2.4	手続き	48
5.3	結果	49
5.4	考察	52
5.4.1	同時性知覚実験における個人差	52
5.4.2	聴覚強度条件を変化させた時の時間統合窓の違い	54
5.5	結論	56
第 6 章	考察	57
6.1	刺激の特徴量と時間統合窓の変動	57
6.1.1	時間統合窓の変動傾向について	57
6.1.2	各実験で得られた時間統合窓の大きさの違いについて	59
6.2	新たな実験手法についての検討	60
6.2.1	心理物理学的実験手法と判断バイアス	60
6.2.2	時間統合窓の大きさを求める心理物理学手法以外の手法	62
第 7 章	結論	64
参考文献		66

付録 A	事後確率の導出	74
付録 B	各個人の実験結果	77
付録 C	平滑化スプラインによる回帰曲線の残差平方和と決定係数	94
付録 D	各実験の実験参加同意書	96

第1章

序論

我々はモノを見たり、音を聞いたり、実際にモノに触れたりすることで、外の世界の様子を知ることができる。例えば、コンサートホールで音楽を聞く時、演奏されている音楽はもちろんのこと、演奏者の動きや打楽器から伝わる振動、さらには会場の匂いなど、色々な感覚情報が合わさり「コンサートホールで音楽を聞く」というイベントを知覚している。このように我々が体験するイベントの多くは、様々な感覚が合わさったマルチモーダルな体験である。

我々が外の世界をどのように捉えているかに少し注意してみると、一度はマルチモーダルな感覚情報がそれぞれの感覚器においてバラバラに処理されたのち、脳のどこかで統合されて知覚されるものだと分かる。この統合処理プロセスはマルチモーダル知覚に関する研究が盛んになっており、様々な分野の研究者が検討を行なっているが、いまだ統一した結果が出されているとは言えない。

このメカニズムが解明されるとどうなるのか。我々にとって大きな利点となるものの一つに、デジタルコンテンツでの体験の向上がある。テレビや映画、ゲーム、そして最近では Virtual Reality や Augmented Reality など、我々の周りには様々なデジタルコンテンツが溢れしており、誰でも気軽にそれらを体験することができる。これらデジタルコンテンツのほとんどは、我々が知覚している世界の再表示であり、その際にヒトの統合処理メカニズムを踏まえた体験表示を行えることが、デジタルコンテンツの体験向上の鍵となる [1]。

本論文は、聴覚と触力覚間の統合処理過程のメカニズムの一端を解明し、両感覚を使ったデジタルコンテンツの体験向上のための設計指標を提案することを大きな目標としている。例えば、ゲームのプレイヤーが銃を打った時にゲームコントローラが震えることで、実際に発弾した時に似た感覚が体験者に呈示されたり、スマートフォンから音と振動が出て我々に電話が来たことを知らせたりなど、身の回りには多くの聴覚-触力覚刺激を使ったデジタルコンテンツの表現がある。しかしながら、この時の聴覚-触力覚の統合処理プロセスは他の感覚同士でのプロセスよりも研究された例が少ない。そのため、多くの聴覚-触力覚を使ったコンテンツの制作時の、ヒトの知覚メカニズムに則った体験設計指標というのはほとんど語られず、現在のデジタルコンテンツにおいて

ては、コンテンツ製作者が、自身の経験や勘などを頼りに体験設計を行なっていることがほとんどである。

筆者自身も音と振動を使ったコンテンツを作ることがある。制作を重ねるうちに、我々が聞いている音が人によって違うのと同様、振動の知覚され方も異なることに気づいた。また、音と振動が同時/非同時に出されたという時間同期感覚もかなり個人差があることが分かった。この時間同期感覚はデジタルコンテンツの体験向上と密接に関わってくる要素である。もあるイベントが現実世界なら音と振動が一緒に出るべきものであるのに、デジタルコンテンツ上で再現した時に、何らかの要因によってずれて知覚されてしまったら、このずれに違和感を覚え、体験の質の低下につながってしまう。

そこで、このずれによる違和感をコントロールし、デジタルコンテンツの体験のクオリティを担保することができる設計指標を制作するために、スマートフォンを持った状態で映画やゲームを体験している際に、音に合わせて振動が同時に呈示されるという限定的な状況で、音と振動の同時/非同時感覚がどのような振る舞いを見せるのかをいくつかの実験を通して検証を行った。

本研究は、前述のように聴覚-触力覚の時間同期プロセスを解明し、両感覚を使ったデジタルコンテンツ制作時の設計指標を制作することを目的としている。これは、Virtual Reality や Augmented Reality などの最先端表現技術の発展とともに、急速に活用されるであろう [2] 聴覚-触力覚の感覚呈示表現を行う際に、その体験クオリティを保証する重要な基準の一つとなると考えられる。また現在用いられている時間同期プロセスを解明するための実験手法は、その振る舞いを認知バイアスによる影響を受けず正確に記述できるものとは言えない。本研究では、個人差による影響を従来手法よりできるだけ排除した、新しい実験手法を提案した。本論文においてはその有用性を十分に検討したとは言えないが、今後検証を重ねることで、聴覚-触力覚の時間同期プロセスを解明する際の、新たな手法の一つとして用いることができるのではないかと考えている。

本章では研究の動機と目的について述べた。第 2 章では研究の背景としてマルチモーダル知覚の仕組みと聴覚-触力覚マルチモーダルの先行研究を述べ、本研究の射程について記す。第 3 章では聴覚刺激の周波数とノイズ量を変化させた時の時間統合窓の変動について、第 4 章では聴覚-触力覚刺激の周波数間隔を変化させた時の時間統合窓の変動について、第 5 章では聴覚刺激の強度を変化させた時の時間統合窓の変動に関する研究を記す。第 6 章で全体的な考察を行い、第 7 章で結論を述べる。

第2章

研究の背景と射程

本章では、まずマルチモーダルと呼ばれる現象について説明し、そのような現象がどのように脳内で処理されているかを解説する。その後本論文のテーマとしている音と振動によるマルチモーダル知覚についての実例及び先行研究を紹介し、最後に本研究の射程を述べる。

2.1 マルチモーダル

一般に生物はその種の生存に適した感覚能力を持ち、そこから得られる情報を処理しながら生きている。例えば、人間の感覚には五感と呼ばれる視覚・聴覚・嗅覚・味覚・触覚(もしくは皮膚感覚)の他に、平衡感覚、運動感覚、内臓感覚などが存在する。これらの感覚の種類は感覚モダリティと呼ばれており、色・音・味・匂い・手触りなど、それぞれ特有の知覚体験を生み出す。そのため従来の研究においては、単一の感覚モダリティから得られる知覚体験を取り扱ったユニモーダル知覚研究が主流であった[3]。

しかしながら、私たちが普段体験する出来事の多くは複数の感覚モダリティが密接に関連した知覚体験である。例えばビールを飲む時、私たちはグラスに注がれるビールの色や泡の形、ビールが注がれる時の音、ホップの香り、グラスの舌ざわりや持ち上げた時の重さ、口元に運んだ時の冷たさや香り、グラスが唇に触れた時の硬さ、口の中に広がる苦味やのどごし感など、多くの感覚から得られる情報を統合して、「ビールを飲む」という状況を知覚している。その一方で、私たちの脳においては、目に入った情報は視覚系、耳から入った情報は聴覚系、手や指から入った情報は触覚系によって、というように一度バラバラに処理される。そしてこれらの情報は、脳内で並列的に、そして統合的に処理することで外界を知覚している。この情報処理の過程では感覚情報間の相互作用が生じることがあり、それに伴って知覚体験が変容することもある。

そのため、最近の研究ではユニモーダル知覚だけではなく、複数の感覚モダリティ間の情報処理過程や相互作用から得られる知覚体験を取り扱う研究が盛んになってきている。このうち、多くの感覚情報が合わさった状態のこと、および複数の感覚モダリティの相互作用によって生じる知覚体験や知覚現象は、総称してマルチモーダル(Multimodal)と呼ばれる。なお、特定のモダリ

ティが他のモダリティの知覚に影響を及ぼす場合には、特にクロスモーダル (Crossmodal) と呼ばれる。

2.1.1 マルチモーダル知覚

前述のように、我々は同時多発的に入力される複数の感覚情報を並列的に処理して外界を知覚している。取り込まれた感覚情報は、中枢神経系の初期の処理レベルでは独立に処理されたのち、単一のイベントを構成する情報が相互に関係付けられ統合される。つまり、脳が知覚世界を作り上げるためにには、入力段階で異なる情報処理を経た各感覚情報を正しい組み合わせで適切に対応づけないといけない。従って、異なる感覚モダリティに由来する情報がどのようにして統合されるのかを明らかにすることは人間の認知システムを理解する上で重要になってくる。[4]。

各モダリティの情報はどこで統合されるのだろうか。この疑問を解く鍵となる重要な場所の一つとして、中脳に位置する上丘^{じょうきゅう}という部分がある。上丘の深層部には、異なる感覚モダリティから入力を受ける神経細胞が存在しており [5]、哺乳類の上丘深層部では 2 種類の感覚刺激の同時入力に対して発火する多感覚ニューロンが発見されている [6] など、この部分を起点にした多感覚統合の神経メカニズムの研究が行われてきた。これらの研究から、複数の感覚モダリティからの情報統合に関して、三つの法則 (空間法則・時間法則・逆有効性の法則) が提案されている [7]。

空間法則は、異なるモダリティの感覚情報が同じ場所から呈示された時に統合が最も強く生じることを指している。一方時間法則は、異なるモダリティの感覚情報がほぼ同時に呈示された時に神経細胞が最もよく活動し、情報統合が生じることをいう。逆効力の法則は、異なるモダリティの感覚情報のそれぞれの効力が弱い場合に、両者の統合が強く働くことを指している。上丘には個別の感覚刺激が弱い時ほど効率よく情報を統合し、その刺激の存在を強調する機能があるとされている [7]。

次に、身近なマルチモーダル現象について考えてみよう。夏祭りの屋台などでよく見かける様々な種類の味が並ぶかき氷屋のシロップが、実は香料と着色料を変化させることで味の違いを生み出しているということは、マルチモーダルの例としてしばしば語られる。このような視覚と味覚の相互作用に関する研究には、以下のものがある。砂糖入りの梨ジュースを赤色・黄色・緑色・無色に着色したときに甘味に違いが生じるかどうかを調べると、赤色に着色したジュースが一番甘いと評価され、緑色に着色したジュースが一番甘くないと評価される [8]。砂糖を溶かした水溶液を用いた実験では、緑に着色した水溶液は甘さの感度が上がり、黄色は甘味の感度が下がった。また、赤色に着色した場合は甘さの感度に関して影響はなかった [9]。

また、腹話術効果と呼ばれる錯覚 [10] もマルチモーダルの例として有名である。人形の口を手で開閉しながら唇をほとんど閉じたままで喋ると、あたかも人形が喋っているのかのように知覚される錯覚である。テレビを見ているとき、スピーカーは画面の外にあるにも関わらず、画面上の話者の口から音声が聞こえてくるように知覚する現象もこれにあたる。これは視覚刺激と聴覚

刺激が異なる位置から同時に呈示されると、まるで聴覚刺激が視覚刺激の位置から呈示されたように知覚されてしまう現象であると言える。

腹話術効果を生じさせるためには、視覚刺激と聴覚刺激の呈示位置が空間的にある程度接近している必要がある。Jack and Thurlow[11]は発話している顔の映像とその音声を異なる位置から5分間提示し、音声が顔の位置から聞こえている時間を測定した。その結果、顔の映像とその音声が水平方向に20度離れている条件では202秒間、40度では約80秒間の腹話術効果が生じたが、両者が50度離れている時には腹話術効果はほとんど生じなかった。また、話者映像と聴覚情報は時間的にもある程度接近している必要があり、音声が映像に対して200ms遅延した環境においては腹話術効果が観測されるが、300msでは消失すると報告した。

腹話術効果に見られるように、異種感覚モダリティ間の情報統合には両者の時間的整合性と空間的整合性が重要である。これは、同じイベントに由来する各感覚情報は自然な環境下では同じ時空間位置を共有すると考えられるからである[12]。一方、これらのイベントをカメラやマイクロフォンを使って収録し再提示する場合など人工的に感覚情報を呈示する環境下においては、その前提条件は大きく崩れる。

視聴覚間における情報統合の時間的整合性と空間的整合性に関する研究として、放送時に生じる映像信号と音声信号のずれの問題(リップシンク: Lip-synching)が代表的である[13]。この問題に関して、国際電気通信連合の無線通信部は、放送における映像音声の相対タイミングに関する勧告を行っている[14]。放送において映像と音声の間にずれが発生する原因として主に3つ挙げることができる。第一に映像処理過程における遅延である。日本での放送では、1秒間に29.97フレーム伝送されるため、1フレームの映像を送るには約33秒必要になる。また、中継現場や他局からの映像との同期を取るために使用するフレームシンクロナイザや映像効果を付加するためのビデオエフェクタなどは内部の処理をフレーム単位で行うため、そのような装置を通過した映像はそれぞれの機材の処理にかかる分の遅延が含まれることになる。第二の理由として、中継番組で映像と音声が別々の回線で伝送される場合がある点である。例えば海外からの中継で、映像は通信衛星経由であるのに対し、音声は地上回線経由で伝送されることがある。また極端な例では、映像は地球の西周りで、音声は東周りで伝送されることもある。第三に、テレビ受信機に存在する遅延である。映像受信機内では様々な画像処理が行われており、それぞれの処理に対して遅延が付加される。映像の口の動きと音声の再生時刻が一致していない場合には、不自然な感じを強く受け、放送の品質に大きな影響を与えるため、上記遅延要因を加味した放送における映像音声の相対タイミングの許容値が決められることになった[15]。

感覚情報処理において時間感覚と空間感覚は相互依存しており、この関係は時空相待じくうそうたいと呼ばれる[16]。二つの刺激を空間的に離れた場所に継時的かつ短時間ずつ呈示する際、2刺激間の空間間隔の知覚は2刺激間の時間感覚に依存する。これらは時間が空間に及ぼす影響を見たタウ効果[17]や空間が時間に影響を及ぼす影響を見たカッパ効果(S効果)[18]などで知られている。このように、感覚情報処理と対象とする感覚情報の時間的整合性・空間的整合性は極めて密接に結び

ついており、時間知覚処理は空間知覚を組み入れたメカニズムが存在することが予測される。

2.1.2 同時性知覚

視覚的刺激の下でも聴覚的刺激の下でも時間知覚は成立する。知覚された時間の長さに多少の違いはあっても、知覚された時間を感覚モダリティの違いによって全く異質なものであると感じることはない。そこで感覚モダリティ間に共通な、時間知覚を成立させるための何らかの過程を考えることができる。もしこのような共通な過程が存在するのならば、それはどのようなメカニズムを持っているのだろうか。

そもそも時間とは何か。我々は日常生活で普段何気なく時間という言葉を使う。特に音楽の世界では、時間の流れを持ったメロディーや和音の進行によって多種多様な表現が生み出されるよう、音楽と時間は分かちがたく存在している [19]。しかしながら、佐藤は「時間、空間といったあまりに広い概念は、本来定義できないものである」 [20] と述べるように、時間の定義は難しい [16]。

そこで難波は、研究意図によって焦点を絞った時間の定義を提案し、「『精神物理学^{*1}における時間』という拘束条件の下で、刺激に対する計測された時間を時間事象として取り扱うこと」 [16] を進言した。ここで時間は、電波時計などから提供される標準化された単位による時間であり、これは物理的に確かに観測できる。時間そのものは刺激ではないが [22]、上記の提案により、心理物理学的実験において、時間事象(計測された時間)からこれに対応する感覚量の関係を研究することができる。

さて、時間事象から生み出される感覚とは何か。難波 [16] はこれを時間事象に対応する時間意識と名付けた。これは観測者の「心理的あるいは主観的時間」 [12] ということができる。例えば Gibson は「事象を基礎的事実として、また時間を事象からの抽象概念として考えてみるべきである。事象は知覚される、しかし時間は知覚されない」(古崎訳、p.109)[23] と言及している。これに対する解釈として、松田は「時間は事象の変化や継起を基礎に置くものであり、したがって、心理的時間は、様々な感覚モダリティで捕らえられた事象の、変化や継起の側面の抽出と統合という認知処理過程を経て心理的に構成されたものと考えることができるだろう」 [24] とまとめている。

このように、時間知覚を心理物理学的手法を用いて考える際、呈示される刺激の時間事象の変化とそれに対応する時間意識との関係を明らかにする必要がある。ここで、どのような刺激が時間事象となるのか、またこの時間事象がどのような処理過程によって時間意識に変換されるのかが問題となる。

^{*1} 精神物理学 (psychophysics: 本文中では以降心理物理学と明記している) は 19 世紀に G. T. Fechner によって提唱された、精神と身体との相互関係を研究する学問である。これは「刺激の物理的性質と、その刺激によって生じる感覚・知覚などの心理的過程との量的関係を研究する実験科学の分野を指す」 [21] と定義されている。

事象時間・脳時間・主観時間

時間事象と時間意識の例のように、物理学における時間と知覚される時間は異なるものである。人間の時間に関する知覚について議論する際には、事象時間、脳時間、主観時間の三つに時間を区別して考える必要がある [25]。藤崎はこれら時間の定義を、「事象時間とは外界で事象が起きた時間であり、脳時間とはその事象によって引き起こされた脳活動のタイムコースであり、主観時間とは事象のタイミングについて実際に観察者が知覚した内容である」 [26] と述べている。

事象時間、脳時間、主観時間の違いを、自分の目の前で拍手をした場合を例にとって考えてみる。事象時間とは外界で事象が起きた時間である。しかしながら、事象が発生し様々な感覚刺激がその時点で同時に発生していたとしても、刺激が感覚器に届くまでの間に、その時間関係は事象からの距離に応じてずれてしまうことに注意する必要がある。自分の目の前で拍手をした場合、その事象が発生した時点で、拍手音(聴覚情報)と手の動き(視覚情報)、手に伝わる感覚(触覚情報)がそれぞれの感覚器へ伝搬される。音の伝搬速度は常温で約 340 m/s であるのに対し、光の伝達速度は約 30 万 km/s であるため、光の方が圧倒的に速い。例えば頭部の前方約 30 cm の位置で拍手したとすると、拍手音は 0.8 ms 後に鼓膜へ届くのに対し、視覚情報は 0.000001 ms 後に網膜へと届く計算になる [4]。ちなみに、この例の場合、触覚情報は外界で刺激が伝搬しないので、事象時間は発生しない。

次に脳時間を見てみる。脳内においても、神経伝達や処理の過程で感覚刺激間に時間のずれが生じる。網膜の錐体細胞の変換プロセスは、明所視で約 50 ms である [27]。また、聴覚の蝸牛における変換プロセスは約 40 μ s で、脳への信号の伝搬に約 10 ms 要するとされている [28]。皮膚上の機械受容器から神経細胞への変換には約 2 ms かかり [29]、体性感覚の神経伝達速度は、平均して 50 m/s 程度であるとされている [30]。そのため、片腕の長さを約 50 ms とすると、拍手時に生じた触覚情報は約 10 ms で脳に到達する。このように、たとえ全ての感覚情報が同時に感覚器に到達していたとしても、処理の初期の段階から時間のずれが生じてしまうことを意味する。

主観時間とは、「(感覚情報が) 同時である」や「(感覚情報のうち) どれが先」といった、事象のタイミングについての観察者の知覚内容である。目の前で拍手をすれば、上記のように各感覚情報が脳へと到達する時刻は異なるにもかかわらず、様々なタイムラグを脳の高次の部分が調整し、その音が手を叩くのと同時に鳴ったと知覚される [26]。

もし脳の高次処理の部分のどこかに同時性の判断を担う領域があるなら、この領域に張ってあるゴールに最初に到達する神経的信号がどの感覚情報なのかを見張っていれば良いことになる。このように情報が脳内の一つの場所に集められることで意識が形成されるという考えはカルテジアン劇場(Cartesian Theater)モデルと呼ばれる(図 2.1)。このモデルでは、全ての情報はある脳内領域を通過するという想定になっている。しかしながら、実際にはこのような劇場もゴールも我々の脳内には存在しない。感覚情報は主観的な知覚体験であり、時間のずれは確かに存在するが、それらの微妙なずれが一定の幅の時間窓にあれば、注意がこれらを結びつけて、時間的な同

一性が産み出されるという考え方方が主流となっている [16]。

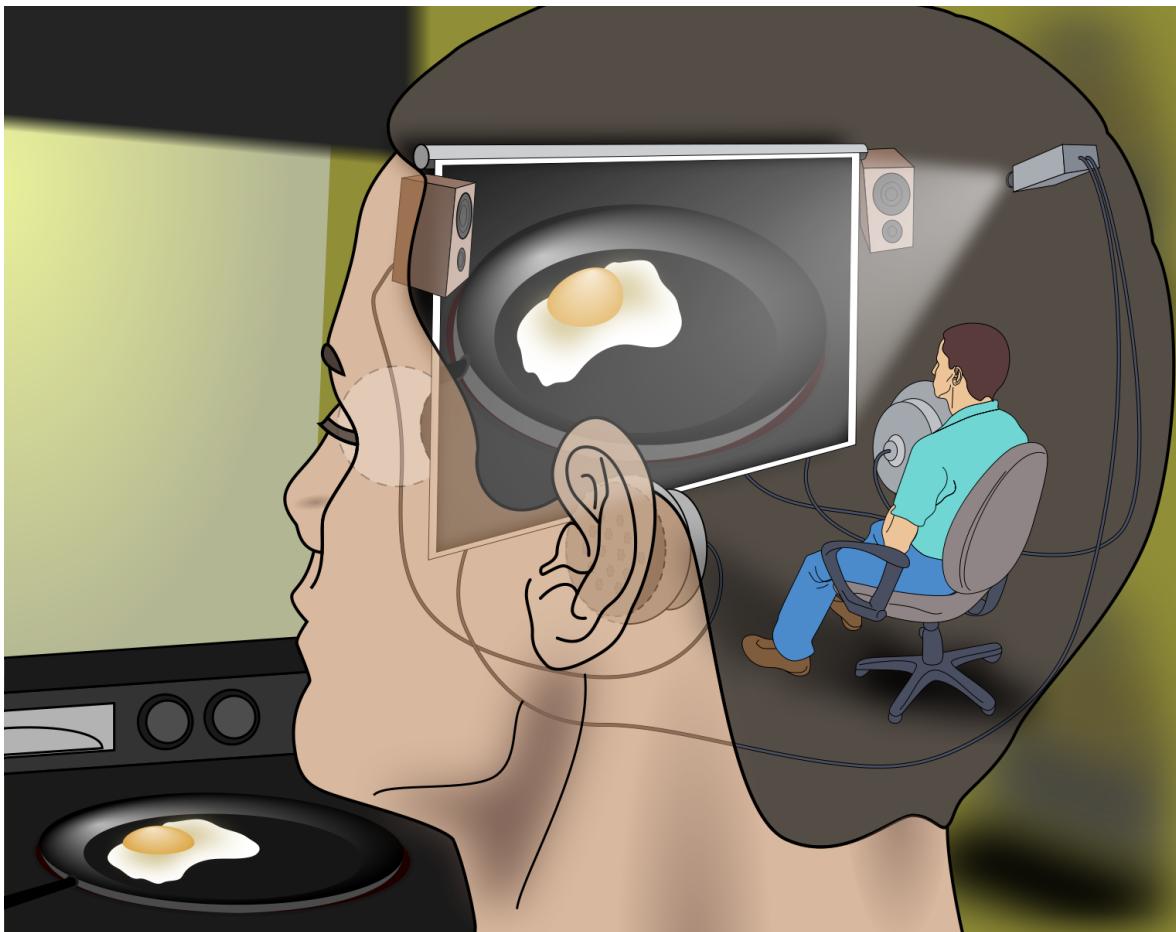


図 2.1 カルテジアン劇場モデルのイメージ図 [31]。経験される対象は脳の中に住んでいる小人の前で上映される。

時間統合窓

異種感覚情報の統合にあたって、呈示タイミングや空間情報といったそれぞれの物理刺激のパラメータが少し異なっていても、ある一定の幅の時間窓内にあればそれらの感覚情報は統合される。この時間的な同一性を産み出す時間窓のことを時間統合窓 (Window of temporal integration) と呼ぶ [32]。前述のように、感覚情報の絶対時間を司る器官がわれわれに存在しないため、時間の知覚、とりわけ感覚間の同時性の知覚については複雑な処理過程が存在する。また、脳内で同時性に関する処理を行う際には、外界での物理的な時間的ずれと体内の神経的な伝達速度の違いをそれぞれ対処する必要がある。これらを上手く処理して、ある外界でのイベントでの同時性を知覚する仕組みを説明する単純な例として、脳はこのような時間的ずれに敏感ではなく、ある時間枠に感覚刺激が複数存在する場合、それらは同時に発生したものだといわば錯覚してしまう、という考え方がある。

図 2.2 はこの概念を図式化したものである。視覚刺激と聴覚刺激はそれぞれの感覚器で処理さ

れ、脳の中でその刺激が発生したことを知覚する(図2.2中の黒色のポール)。時間統合窓の考えでは、脳は時間的ずれに敏感ではないので、複数感覚の刺激の発生がある時間窓の内で起こっていった場合、その中で発生した感覚刺激は同時に起こったものとして知覚されたことになる。

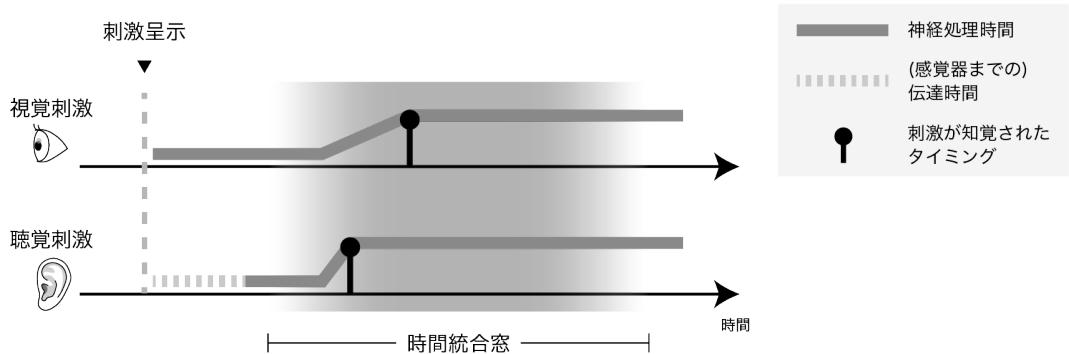


図2.2 時間統合窓の概念図 (Vroomen and Keetels[32]を改変)

同時性知覚の測定

時間統合窓の存在及びその窓の大きさを測定する心理物理学的測定方法は代表的なものとして2種類存在する。1つ目は同時性判断課題 (Simultaneity judgment task) である。これは、2つの感覚情報を呈示した際、それらの刺激が「同時であった」か「非同時であった」かを判断させる方法である。2つ目は、時間順序判断課題 (Temporal order judgment task) である。この課題では、2つの感覚刺激を呈示した際、「どちらの刺激が先であった(もしくは後であった)」かその刺激の呈示順序を判断させる方法である[4][16][33]。

スピーカーから呈示される短音と照明機器から呈示されるフラッシュを実験刺激として使用した時の同時性知覚実験を例とする。同時性判断課題による実験では図2.3の左側のグラフのような結果が表される。横軸が短音とフラッシュの刺激開始時間差 (Stimulus onset asynchrony) であり、刺激開始時間差が正の値の場合は、フラッシュより短音を遅れて呈示されたことを表している。縦軸は同時と判断された割合である。通常同時性知覚の実験を行う際は、刺激開始時間差の値をいくつか設定しそれに対する実験参加者の同時と判断された割合を算出する。その後同時性判断課題ではプロットされた結果から、標準正規分布の確率密度関数を用いて回帰曲線を算出する。この回帰曲線のうち、同時と判断された割合が最大になった点、つまり正規確率密度関数における中央値 μ を主観的同時点 (Point of subjective simultaneity) と呼ぶ。また、正規確率密度関数における標準偏差 σ に 0.6745 を乗じた値を、2つの感覚刺激間の時間的ずれがわかる最小のずれの大きさ、丁度可知差異 (Just noticeable difference) と呼ぶ[34]。丁度可知差異が小さければ、短い時間ずれで2つの感覚刺激を弁別可能になることから、時間的ずれへの感受性を示すものと

して用いられ、この値を時間統合窓の大きさの指標としている [4]。

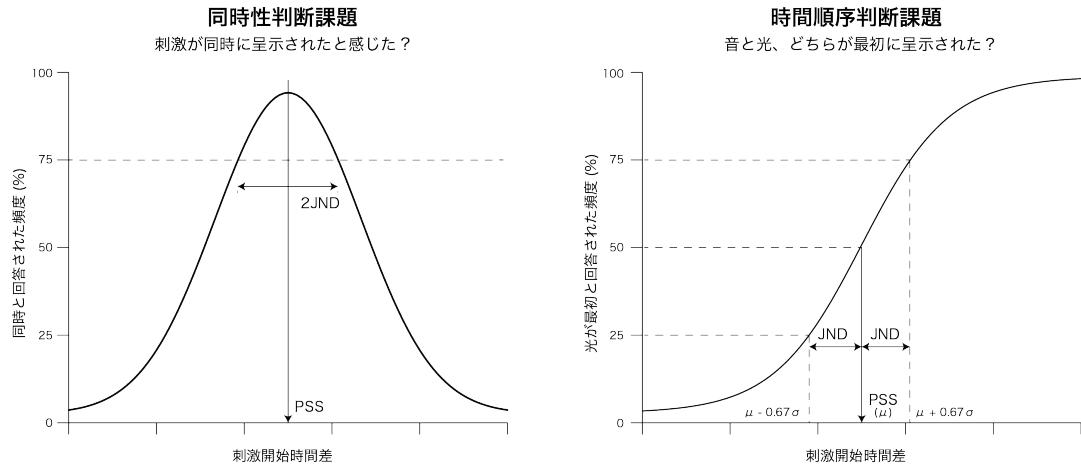


図 2.3 同時性知覚実験に用いられる 2 つの課題と指標 (難波 [16] を改変)。各課題で得られた結果から標準正規分布の確率密度関数または累積分布関数を用いて回帰曲線を算出することで、丁度可知差異 (JND) 及び主観的同時点 (PSS) という指標が得られる。

一方、時間順序判断課題課題の場合は図 2.3 の右側のグラフのような結果が見られる。縦軸はフラッシュが先に呈示されたと判断された割合である。実験参加者に判断を求めた結果は標準正規分布の累積分布関数を用いて回帰曲線が算出される。回帰曲線から縦軸が 50 % となる点が主観的同時点となり、主観的同時点と縦軸が 25 % 及び 75 % を切る点の差が丁度可知差異となる。

2.2 聴覚-触力覚マルチモーダル知覚

今までのマルチモーダル知覚に関する研究は、視覚と聴覚に関するものが多く、聴覚と触覚^{*2}を対象とした研究がなされることは少なかった [35]。これは日常生活でのマルチモーダル現象を考えると不思議に見える。例えば蚊が我々の首の後ろへ回って血を吸おうとしたときに感じる羽音と痛みの感覚や、スマートフォンが音と振動を同時に鳴らして、電話やメールが来たことを知らせる機能など、日常には視覚情報に依拠しない聴覚-触力覚のマルチモーダル現象が、たとえそれがごく小さく気づかないようなものであっても、数多く存在している。

ここでは具体的な議論に入る前に、聴覚と触力覚の生理学的な仕組みを説明したのち、両感覺に見られる構造上の類似点について述べる。

^{*2} 本論文では、触覚は痛みや温度感など物に触れて生じる感覚全般のことを指し、触力覚は触覚のうち機械受容器が担当する触れたり圧迫されたりする感覚のことを指している。

2.2.1 感覚器の構造

聴覚系の構造

私たちが音を聞くことができる原因是、外界の音情報（空気の疎密波）を耳を通じて神経情報に変換し、脳神経系で情報処理することによってである。図 2.4 は音情報を受け取ってから神経情報として変換されるまでの聴覚系の構造を表したものである。

空気の振動である音は耳介で集められ、外耳道を経て鼓膜に到達する。鼓膜の振動は耳小骨を経て蝸牛へと伝わる。蝸牛は細長い管（蝸牛管）がカタツムリの殻のように巻かれた構造になっている。蝸牛管の中には長軸方向に張られた膜によって前庭階、中央階、鼓室階に分けられている。前庭階と中央階を隔てる膜をライスネル膜、中央階と鼓室階を隔てる弾力性のある板状の組織を基底膜と呼ぶ（図 2.5 および図 2.6）。

音波が耳小骨を通り蝸牛に伝わると、前庭階と鼓室階のリンパ液の間に圧力差が生じて基底膜が進行波として振動する。高周波音では進行波が基底膜の基部に近い部分で最大振幅となり、その場所を超えると振幅は急激に減衰する。低周波音では、進行波先端まで進んで最大振幅となる。

このように、入力音の周波数情報は基底膜上の場所情報に変換される。基底膜上有る場所に着目すると、その場所はある周波数範囲の入力音に対して選択的に大きく振動する。このように蝸牛の基底膜振動系は、入力音をその周波数成分にふるい分ける機械的な帯域フィルタ群と見なせる。

聴覚フィルタの中心周波数は基底膜の先端から根元に向かって対数的に並ぶ。この基底膜上の周波数軸の尺度として、心理物理的に測られたヒトの周波数分解能（臨界帯域幅）から着想を得た Bark 尺度と ERB 尺度という二つの尺度がある [36, 37]。

皮膚感覚

皮下には多数の触覚受容器が存在している。手指、手掌などの無毛皮膚部には、形・大きさ・配置の異なる複数種の触覚受容器が埋め込まれており、それぞれの触覚受容器の配置は皮膚構造と深く関わっている。これは、目や耳が独立した単一の組織であることと対照的である。皮膚感覚受容器は、表皮と真皮の境界から皮下組織にかけて存在する。それらには、触力覚情報を受容する機械受容器、温度情報を受容する温度受容器、痛覚情報を伝える侵害受容器がある（図 2.7）。

触力覚は生体の表面に加えられた圧力や振動といった触刺激によって生じる感覚である。この感覚に関する受容器はマイスナー小体、パチニ小体、メルケル盤、ルフィニ終末の 4 種類の機械受容器であり、刺激の応答や反応する周波数にそれぞれ違いがある。

Bolanowski *et al.* は 0.4 Hz から 500 Hz までの間で、4 種類の受容器が受け持っている振動周波数領域を調べた（図 2.8）。皮膚の表面の温度は 30 °C であり、振動呈示時間は 700 ms、振動子の大きさは 2.0 cm² である。その結果、パチニ小体は 40 Hz から 500 Hz の周波数に対して U 字

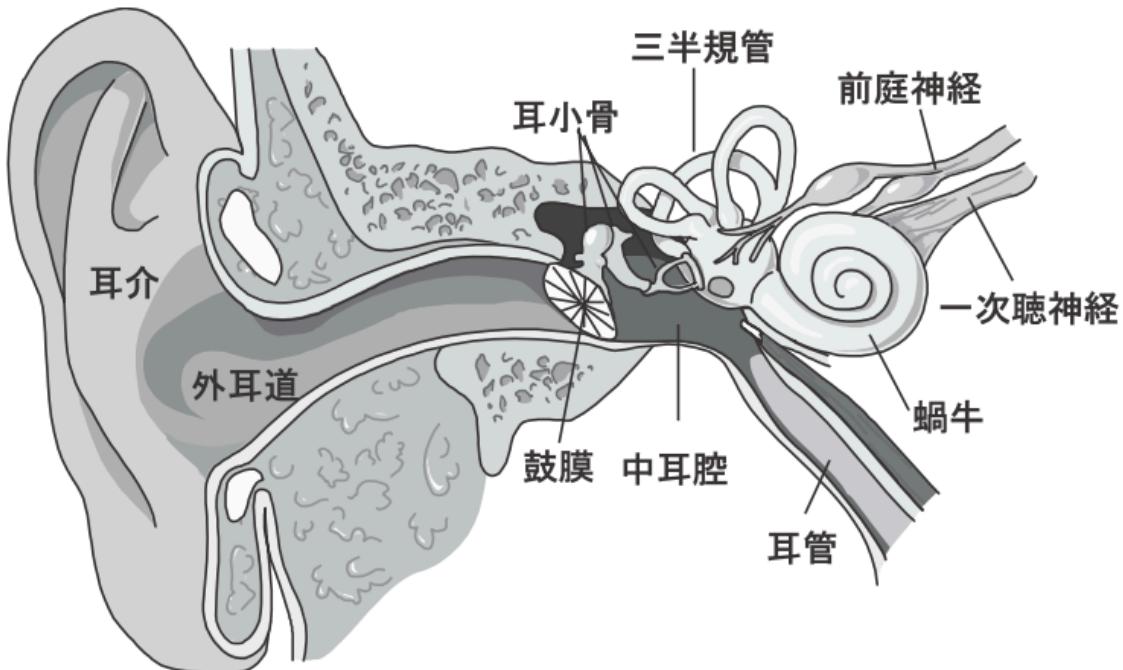


図 2.4 聴覚抹消系の構造 (平原 [37] から引用)

型の応答を示すことがわかった。また、マイスナー小体は 3 Hz から 40 Hz の刺激を処理し、刺激に対してどの周波数であってもほぼ一定に反応する。メルケル盤は 0.4 Hz から 100 Hz の範囲内で働き、ルフィニ終末はパチニ小体とほぼ同じ周波数域で反応するが、感度がかなり低いことがわかった [39]。

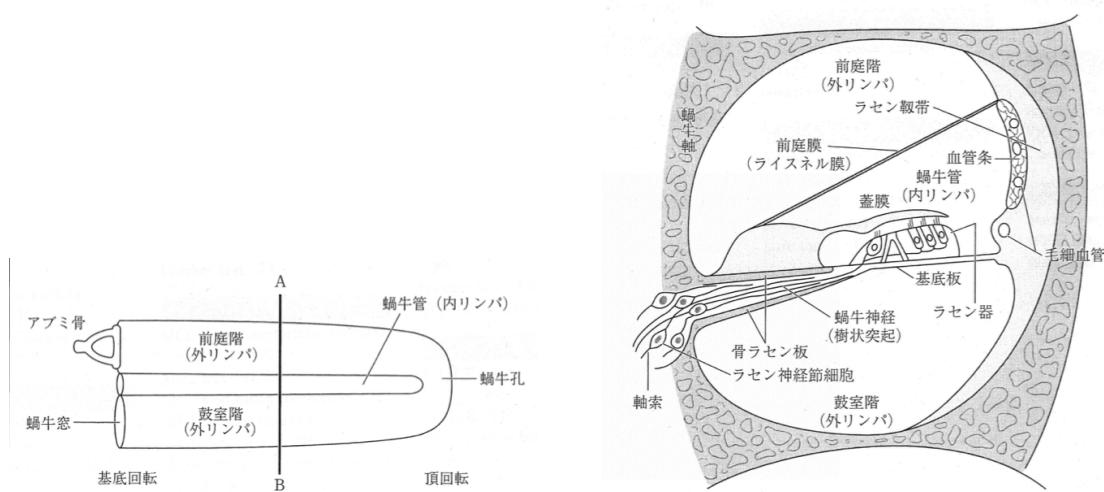


図 2.5 蠕牛の構造 (日本聴覚医学会 [38] から引用) 図 2.6 蠕牛の断面図 (日本聴覚医学会 [38] から引用)

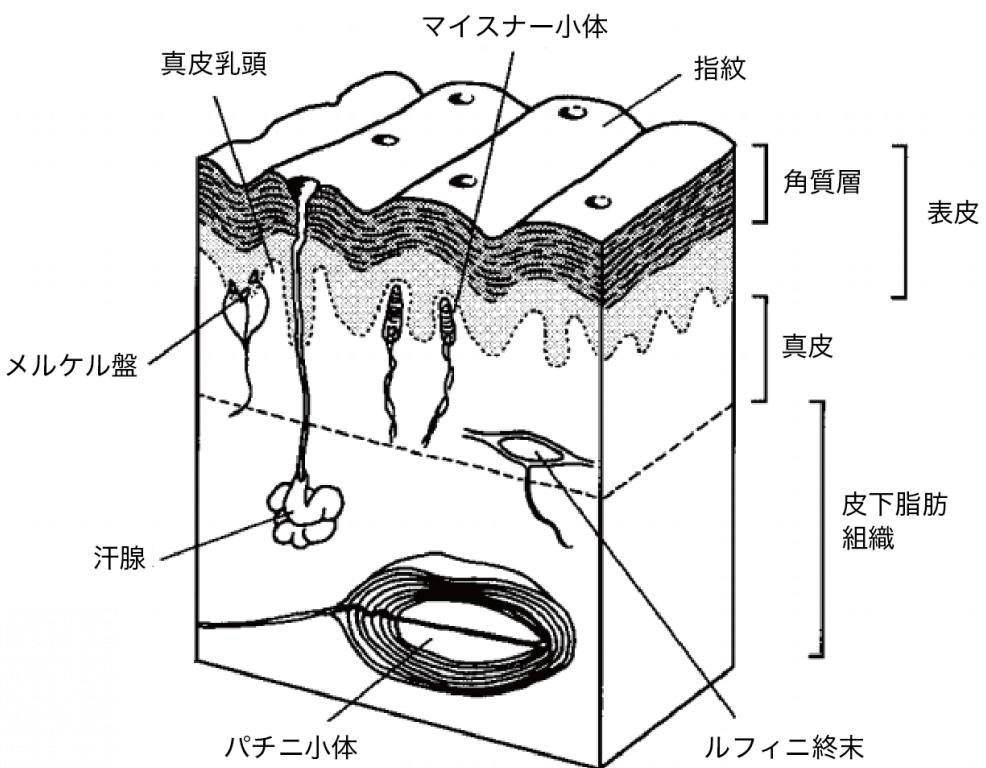


図 2.7 皮膚と機械受容器の構造 (前野 [40] を改変)

聴覚と触力覚の類似性

Von Békésy は、聴覚感覚器と触力覚の感覚器には構造上の類似点があると指摘した。彼は、ヒトの進化の過程で機械受容器がより敏感になった器官が基底膜になったという例から [41]、基底膜と機械受容器の信号変換メカニズムに類似点があり、特に刺激の強度知覚において似たような反応を見せる報告している [42]。

もちろん、それぞれの感覚受容器が知覚できる周波数領域が、触力覚の場合は 0.4 Hz から 1 kHz ほど [43] であり、聴覚の場合は 20 Hz から 20 kHz までとかなり異なっているため、この類推にも限界がある。しかしながら興味深いことに、幅は小さいながらも両者の周波数領域が重なりある部分があり (図 2.9)、この重なり合う部分の振動感度が同じような振る舞いを見せることが分かっている [44]。この現象が聴覚と触力覚の情報処理において、ごく初期の段階で情報の統合が行われるのでないかという根拠の一つとされている [45]。

また、聴覚と触力覚の同時性の特異性を説明する研究の一つに、Fujisaki and Nishida の研究

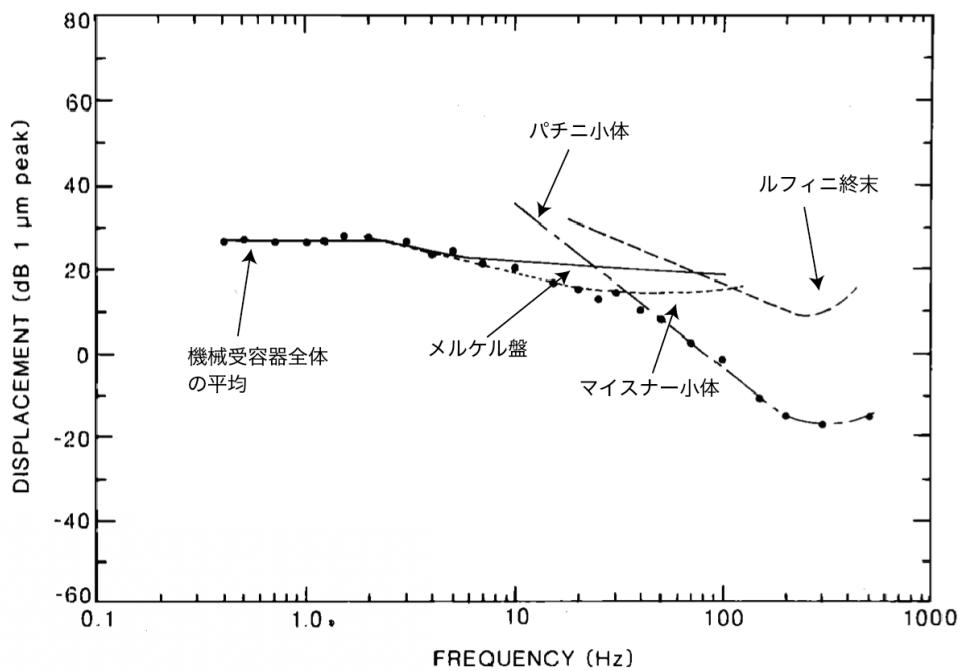


図 2.8 各機械受容器が担当する周波数領域 (Bolanowski *et al.*[39] を改変)

[46] がある。彼女らは同時性知覚の振る舞いが感覚の組み合わせによって変化するのかを同時性判断課題及び時間順序判断課題を用いて実験を行なった。その結果、視覚-聴覚・視覚-触力覚・聴覚-触力覚の同時性知覚において、聴覚-触力覚は視覚-聴覚・視覚-触力覚よりも時間分解能が高く、視覚-聴覚・視覚-触力覚の同期・非同期弁別閾が約 4 Hz となるのに対して、聴覚-触力覚の弁別閾が約 10 Hz と特異的に高くなることが明らかになった。

上記のように聴覚-触力覚マルチモーダル知覚は、構造上の類似性から同時性知覚に関する認知システムにおいてある程度共通の考え方を用いることができる可能性がある。また、時間統合感覚の分解能が高いことから時間同時知覚において他のマルチモーダル知覚よりも時間的なずれを感じやすいという特徴を持っている。

2.2.2 聴覚-触力覚マルチモーダルの実例

ここでは、現在一般に体験することができる聴覚-触力覚マルチモーダルの例を紹介したのち、聴覚-触力覚マルチモーダルの時間統合感覚に関する先行研究と問題点を述べる。

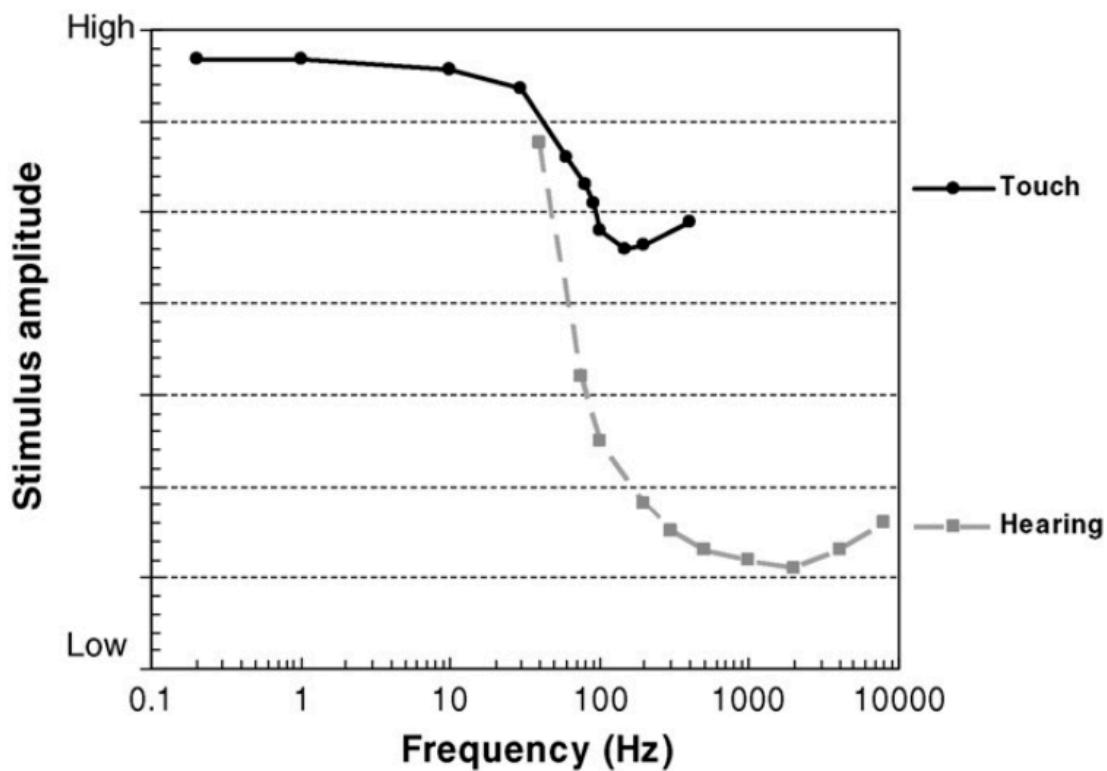


図 2.9 聴覚受容器と機械受容器の強度の知覚閾。およそ 20 Hz から 500 Hz まではそれぞれの知覚域が重なり、似たような振る舞いを見せている。(Soto-Faraco *et al.*[45] より引用)

映画やゲームといったデジタルコンテンツでは、体験の臨場感を向上させるといった目的で音と振動を組み合わせた体験表示がしばしば行われてきた。

1963 年に Morton L. Heilig により開発された SENSORAMA(図 2.10) は複数感覚を同時に表示するマルチモーダル体験の元祖とされている。これは街中をバイクに乗って走り回る一種の体感型ゲームになっており、このシステムでは、利用者が椅子に座ると目の前のディスプレイに立体ビデオ映像が映し出され、バイクで街中を移動するという状況の変化に応じて音響や椅子の振動、あるいはファンによる顔への風の表示、風景に合わせた匂いの表示などの仕組みが組み込まれていた [47]。

1974 年に公開されたユニバーサル・ピクチャーズのパニック映画、『大地震』(Earthquake) には、センサラウンド (Sensurround) という地震を疑似体験できる音響効果が搭載されていた(図 2.11)。このシステムでは低周波音を発生させるスピーカーが 16 個以上設置され、110–120 dB_{SPL} の音圧レベルを持った 5–40 Hz という低周波音が椅子や床だけでなく観客までも揺さぶった。このシステムは『大地震』だけでなく『ミッドウェイ』(1976 年)、『ジェット・ローラー・コースター』(1977 年)、劇場版『宇宙空母ギャラクティカ』(1978 年) などでも用いられた [48]。

1983 年パイオニア株式会社は、椅子に座って音楽を聴きながらその振動をも全身に伝えて心身

をリラックスさせることを目的としたシステム、ボディソニックを発表した(図2.12)[49]。このシステムは、ロケット工学で有名な糸川英夫による「楽器を演奏する人は耳で聴く音の他に、楽器を持つ手、身体を通して直接振動が伝わり、骨を通って聴覚系に伝播される音もある。音楽の中で聴く人に真の恍惚感を与えるのは、音波ではなく骨に直接伝わる振動、ボーンコンダクションである」というボーンコンダクション理論[50]をもとに、パイオニア株式会社創業者の松本望が考案したものである。このシステムは音楽聴取だけではなく、聴覚障がいの方が音楽を楽しむ際の道具として、また患者へのリラクゼーション効果を狙って歯科治療の際の手術椅子にも用いられた[51]。

1997年に任天堂より発売されたNINTENDO64用3Dシューティングゲーム『スターフォックス64』では、ゲーム内の行動に応じて振動が表示されるフォースフィードバックという機能が家庭用ゲームでは初めて採用された。このシステムは、コントローラに搭載された振動アタッチメント(図2.13)[52]がゲームの演出によって振動することで、プレイヤーのゲームへの熱中度が増すように意図されたものであった。このシステムは株式会社ソニー・コンピュータエンタテインメント(当時)のPlayStation用コントローラDUALSHOCK[53]など他社のゲーム機用コントローラでも採用されるようになった[54]。

最近では、Appleが2015年に発表したiPhone 6sにおいて、デバイス内にTaptic Engineと呼



図2.10 SENSORAMA[47]



図2.11 映画『大地震』のポスター
(Hauerslev[48]から引用)



図 2.12 金沢工業大学 ライブラリーセンター ポピュラー・ミュージック・コレクション (PMC) にあるパイオニア製の BODYSONIC SYSTEM MC-300。座面に振動子が取り付けあり、外部再生機から流れる音楽の低周波音が振動へと変換されて聴取者に呈示される。(筆者撮影)



図 2.13 NINTENDO64 の周辺機器として発売された振動パック (型番 NUS-013)[55]

ばれる振動子が搭載された [56]。ユーザーの行動に合わせて振動子が微かに震えることにより、画面をタップしたときに実際にクリックしたような感覚を得たり、物理的なボタンのないスマートフォンでも、実際にボタンを押したような感覚を得ることができる。この振動子は従来のフューチャーフォンや他社のスマートフォンなどに搭載されていた振動子よりも、繊細かつ振動周波数が広い。このため、前述のクリック感の演出の他にもゲームのユーザ体験を向上させるための手段としても頻繁に使用されている [57]。

このように、様々な分野のデジタルコンテンツにおいて日常的に音と振動を組み合わせた体験呈示が行われており、これらの表現は様々に形を変えながら、今日では一般的な体験演出手段として用いられている。

地震の再現やリラックス効果の演出などの音と振動を用いた表現は、両刺激の同時呈示による相乗効果で強い感情の誘起させることを目標に制作されている。しかしながら、音と振動の同時呈示が本当にそれらの感情を誘起させる効果があるのか、いくつかの研究で検証がされている。

福井らは、ボディソニックでのリラクセーション効果の検証のために、皮膚表面の微細振動からリラクセーション時に多く発現する α 帯域の出現率、末梢血流量の増減、ストレスホルモンとして分泌される β -endorphin の量をそれぞれ計測し、ボディソニックを体験する前後でそれぞれの値がどのように変化するのかを調査した。その結果、各実験において実験参加者がリラクセーションのために最適な選曲をし最適の音量と振動の強さを設定したときに、リラクセーションに関わる上記生体現象が見られることが分かった。[58, 59, 60]。

阪梨は、ボディソニックのような全身に振動を付与する装置を用いて、重低音感が強調されることでもたらされる臨場感の向上がどのような刺激条件の際に強く発生するかを調査した。その結果、振動付与から得られる低音感が、必ずしも聴取者にとって臨場感の向上に繋がるとは限らない可能性が示唆された。音楽聴取の際、振動の強度や発音タイミングを変化させ、印象と評価する実験を行ったところ、振動が付加された時に感じる不自然なリズム感が臨場感の低下に繋が

る可能性を得た。さらに、振動信号の時間軸波形において急峻なピークを示す箇所が音楽と同期していないと感じた時に、臨場感を低下させる効果があることを報告した [61]。

岡崎らは、モバイル環境における音楽体験の向上を目的とした聴触覚変換手法を提案し、その評価を行なった。ここでは、聴覚刺激を分周^{*3}させたものを触力覚刺激として用いる聴触覚変換手法であった。その結果、主に音楽刺激が高い周波数成分を持ち音波形をそのまま触力覚刺激として用いても振動が知覚できないような音源に対して、音楽波形を分周して振動呈示を行うことで、音源を知覚しやすくなるだけでなく、分周によって周波数が異なった音楽と振動の間に乖離感が生じず、心地よさや好み、楽しさといった主観的な評価が全体的に向上することが明らかになった [62]。

Lee *et al.* は、モバイル端末を使って音と振動が同時に呈示される体験において、聴覚刺激から触力覚刺激と関連性の高い音響特徴量を抽出し、それを基にリアルタイムに触力覚刺激を呈示するアルゴリズムを提案した。また、このアルゴリズムと従来の振動呈示アルゴリズムを比較して、コンテンツのユーザー体験がどのように変化するのかを調査した。この研究では、聴覚刺激で知覚される要素のうち、触覚での再現・知覚が可能な強度(聴覚刺激においては Loudness、触力覚刺激においては Intensity)と粗さ(Roughness)知覚に着目し、聴覚刺激からそれら特徴量を抽出し、それを触力覚刺激の特徴量へ変換した上で、触覚刺激をリアルタイム生成する手法を考案した。このアルゴリズムを用いて、モバイル端末で音と振動が提示される映画、ゲーム、音楽コンテンツを体験してもらい、その際の実験参加者の印象を調査した結果、全体的に一般的な触覚提示方法である聴覚刺激の信号波形を無編集で振動子に再生させた時よりも、強い衝撃が与えられると予想される特定のイベントが来た際に強い振動が呈示される本アルゴリズムの方が、好ましさや快適度、臨場感などのユーザ体験が高かったことを報告した [63]。

2.2.3 聴覚-触力覚マルチモーダルの問題点

前節では、聴覚-触力覚マルチモーダル刺激が呈示されるデバイスの歴史を紹介したのち、このような刺激呈示が目的としている臨場感の向上がどのような要素から成り立っているのか調査した先行研究を紹介した。取り上げた研究では、使用されたデバイスや環境は違えど、音と振動が呈示されるイベントや文脈に刺激が合っていたかが、強い感情の誘起に必要な要素であったと考えられる。

Altinsoy[64] は、聴覚-触力覚マルチモーダル刺激を使用したコンテンツの品質評価に関する報告において、体験の品質に関わる要素について説明している。図 2.14 では、聴覚-触力覚マルチモーダルを使って体験呈示を行う際に、どのような物理的要素と心理的な特徴が体験の品質に変わるかを図解している。ここで、図 2.14 の下部、Audio-Tactile VE(Virtual Environment)のセクションにおいて、物理的要素として Temporal attributes(時間的振る舞い)、Location(位

*3 周波数を $1/n$ にすること。

置)、Intensity(強さ)、Frequency(刺激の周波数)があることを挙げ、それに対応する心理的特徴として、Percieved Synchrony(知覚される同時性)、Localization(定位)、Loudness, Strength or Intensity(刺激の大きさ)、Pitch, Timbre(音程、音色または触感)が存在することを挙げた。このうち、どの項目が強い影響力を持っているのかという詳細な主従関係はさらなる研究が必要であるが、この中でも特に刺激の同時性と空間性が、それらが不一致であった場合の違和感が強く印象付けられることから、体験の品質のコントロールを行う上でこの2要素が重要であると述べられている[64]。

2.1.2で述べたように、我々のマルチモーダル刺激の同時性知覚では、刺激が時間統合窓内にある時に同時であると知覚され、その窓の大きさは様々な要因によって変化すると考えられている。Voorenでは、時間統合窓の大きさの変動に関わりのある要素として、呈示される刺激の空間性(Spatial separation)、刺激の複雑さ(Stimulus complexity)、刺激の周波数(Frequency)、予測情報(Predictive information)、一体性の仮定(Unity assumption)などを挙げている[32]。現在までに、いくつかの研究者がこれらの要素をコントロールした時の時間統合窓の変動について報告している。

例えば、Adelstein *et al.*はハンマーを叩くとそれに伴い音と振動のフィードバックが与えられる環境において、音刺激を1ms、50ms、200msと変化させた時の時間統合窓の変動について検討を行なった。その結果、刺激の長さを変化させても時間統合窓に顕著な違いを見せることはなかった。[65]。

また、Zampini *et al.*は聴覚刺激と触力覚刺激を同じ場所から呈示した場合と異なる場所から呈示した場合の2種類の刺激を用いて、時間統合窓がどのように変動するのか報告した。その結果、訓練された実験参加者とそうでない実験参加者で時間統合窓の大きさが顕著に変化したが、呈示場所の違いによって変化することはなかった[66]。

他にも刺激の信号包絡線(エンベロープ)を変化させた研究[67]など呈示刺激の要素と時間統合窓の変動について検討された例はあるが、いずれも実験環境や使用デバイスなどの違いにより、統一された結果が報告されることはなかった。そのため、臨場感などをより強く演出するために聴覚-触力覚マルチモーダル刺激の同時性を統制しようと考えても、どのように我々の時間統合窓が決定されているのか、そしてその窓内に収めるためには、どのような刺激のどの物理量を操作すれば良いのかは分からなかった。このような状況のため、現状の聴覚-触力覚マルチモーダルを用いたコンテンツ制作においては、最終的に製作者の経験によってそれぞれ調整される状況となっている。

2.3 本論文の射程

異なるモダリティの感覚情報処理において、時間感覚と空間感覚は相互依存しており、この関係は時空相待と呼ばれる(2.1.1参照)。それら感覚情報の統合にあたって、呈示タイミングや空間

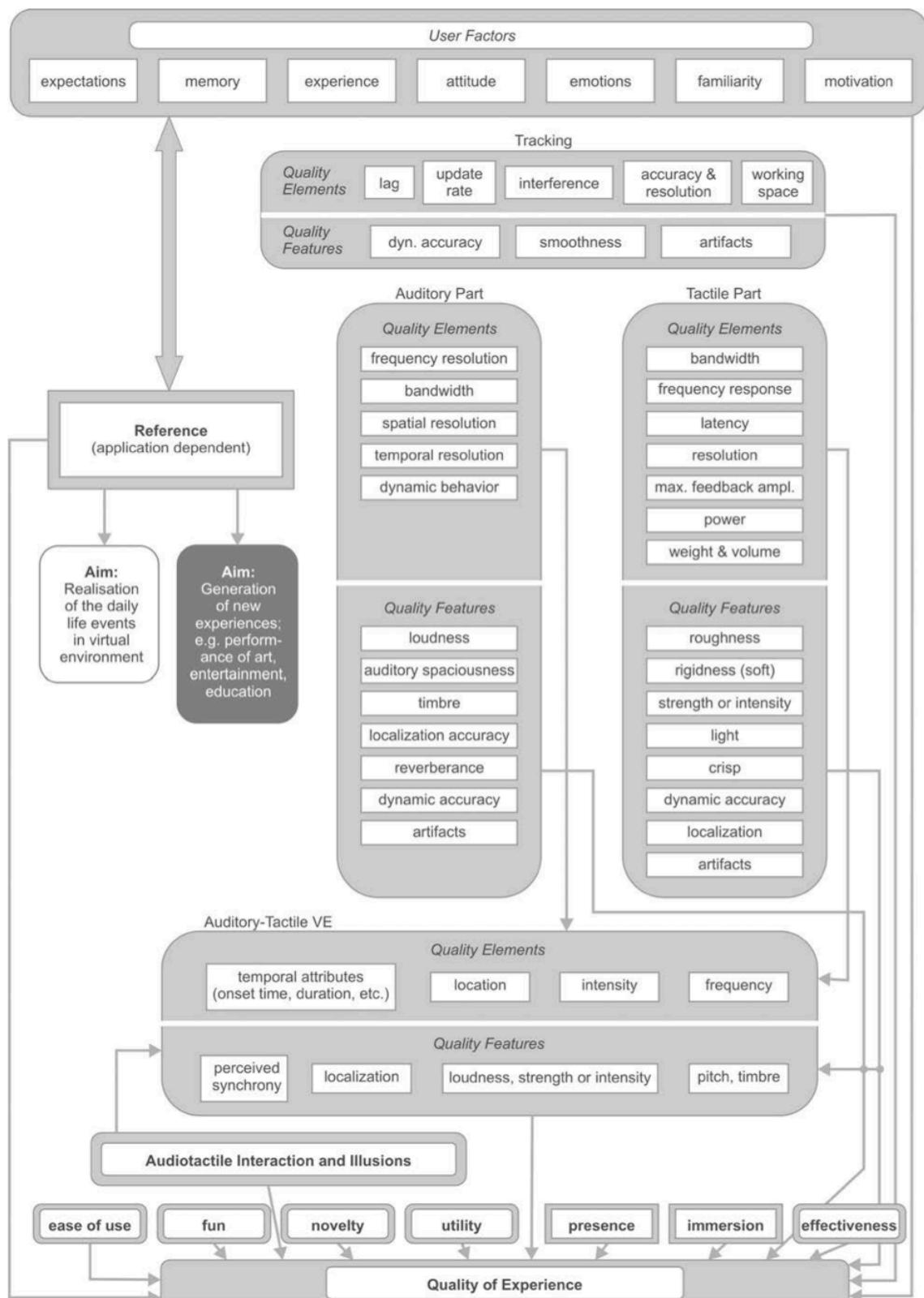


図 2.14 聴覚-触力覚マルチモーダル刺激の品質評価に関わる要素 (Altinsoy[64] より引用)

情報と言ったそれぞれの物理刺激のパラメータが少し異なっていても、ある一定の幅の時間窓内にあればそれらの感覚情報は統合され、時間的な同一性を得ることが知られている [32][68]。この時間窓のことを時間統合窓と呼び、これは同時性判断課題や時間順序判断課題を用いて観測することができる [4][16][33]。以前からこれら研究手法を用いて時間統合窓を調べると研究ごとにその長さが変動することが報告されてきたが、時間統合窓は刺激の特性や知覚者の経験や文脈などの要因によってその振る舞いを変化させることができた [32][66]。

これらの研究は主に視覚-聴覚間でのマルチモーダルの関係性において語られることが多かったが [35]、聴覚-触覚マルチモーダル刺激においても同じ枠組みで関係性を考察することができる。聴覚-触覚マルチモーダル刺激を使った表現は、以前からゲームや映画などで演出の一つとして用いられてきた [69][70]。ここでも視覚-聴覚間と同様に聴覚-触覚間のそれぞれの時間感覚の特性を考慮しながら体験を制作する必要がある (例えば [46])。しかしながら、状況によって変化する時間統合窓を考慮したマルチモーダル体験をどのように作れば良いか具体的に語られることはほとんどなかった。

そこで本論文では、聴覚刺激の周波数とノイズ量 (第 3 章)、両刺激の周波数間隔 (第 4 章)、そして聴覚刺激の強度 (第 5 章) をそれぞれ変化させた時の時間統合窓の変化を調査することで、時間統合窓の振る舞いを決定付ける要因の一端を解明することを目的とした。本論文で行なった全ての実験は、体験者がスマートフォンを持った状態で映像付きの音楽コンテンツや映画、ゲームを体験している際に音に合わせて振動が同時に呈示されるという状況を想定して行われ、実験デバイスはオーディオインターフェースを除き全て同じものを使用した。

またマルチモーダル感覚情報に関する精神物理学的実験においては、実験結果として得られたものが知覚に基づくものなのか、判断バイアスによるものなのかを厳しく区別する必要がある [16] (詳しくは 4.4.2)。しかしながら、時間統合窓を調査する同時性判断課題や時間順序判断課題と呼ばれる手法には、実験手続き上、実験参加者の判断バイアスを制御することが難しいという問題点があった [7]。そこで、本研究では判断バイアスを上記手法よりも制御することができると言われている [16] ベイズ統計学を用いた信号検出理論を用いることで、より堅牢な時間統合窓の調査方法を提案することも目的としている。

第3章

聴覚刺激の周波数とノイズ量を変化させた時の同時性知覚の変動

本節では、聴覚刺激の基本周波数及び聴覚刺激のノイズ量を変化させた時の同時性知覚について述べる。

3.1 背景

ゲームにおける銃声の表現や、映画での迫力のあるカーチェイス場面などで音と振動を同時に用いることで、単一の刺激提示よりも具体的な体験を呈示させるフィードバック方法がある。今回の実験ではこれらの体験実例から、体験の効果を制御する上で大きな影響を持っていると考えられる二つの特徴量に注目した。

1つは聴覚刺激の基本周波数（実験1）で、もう一つは聴覚刺激のノイズ量（実験2）である。基本周波数が低い音は迫力や莊厳な雰囲気という印象を強め[71]、しばしば触力覚刺激と一緒にこれらの印象を強めるために呈示されることがある[72]。また、ノイズ量の多い聴覚刺激は車のエンジン音などで見られ[73]、しばしばエンジン音のデザインの際、快-不快感などをコントロールするために、刺激のノイズ量をコントロールした刺激が用いられる[74]。今回は、この二つの特徴量が聴覚-触力覚における時間統合幅へどのような影響を与えるのかを調べるために以下の刺激を用意した。実験1では周波数の異なる正弦波で、実験2では中心周波数を4kHzとするバンドパスフィルター(BPF constant 0dB peak gain)[75]を適応することでノイズ量を変化させた音源である。

3.2 実験の概要

3.2.1 実験参加者

アメリカ、ニューヨーク州にあるロチェスター工科大学の学生及び教員 10 名 (男性 7 名女性 3 名、18 歳から 48 歳) が実験に参加し、そのうち 1 名は著者であった。実験時間はおよそ 50 分であった。なおすべての実験参加者は自己申告により、日常生活に支障が出るほどの聴覚障害及び手指の触力覚感覚障害がないことを確認した。また本研究はロチェスター工科大学倫理審査委員会 (RIT Institutional Review Board) の承認を受けて実施した (承認番号: HSRO #02121219)。実験に使用した実験参加同意書を付録 D に示す。

3.2.2 実験装置

実験参加者は簡易防音室で、実験用コンピュータ (MacBook pro Early 2015, 2.7 GHz Intel Core i5 with 8 GB of RAM, Apple Inc.) を前にし、ヘッドフォン (HD650, Sennheiser Electronic GmbH & Co. KG) を付け、振動子 (Haptuator Mark II, Tactile Labs Inc.)[76][77] が装着された実験デバイス (iPhone 8, Apple Inc.) を把持した状態で着席していた (図 3.1、図 3.2 参照)。触力覚刺激は、3D プリンタにて作られた専用のケースによって実験デバイスの背面に接着された振動子から発せられた。実験デバイス自体の大きさは $138.4 \times 67.3 \times 7.3$ mm で、振動子を付けた状態での重さは 178 g であった。

3.2.3 実験刺激

実験 1 で用いた正弦波は、250 Hz から 8 kHz までオクターブ間隔で分けた 6 種類を用意した。実験 2 では、ピンクノイズに適用したフィルターのバンド幅を変化させることで刺激のノイズ量を変化させた。このノイズ量を調整するためのパラメータとして、本実験ではスペクトル尖度という音響特徴量を用いた。スペクトル尖度は主に信号処理において目的信号以外の信号 (ノイズ) の量を計算するための指標として用いられることがある [78]。本実験ではこのスペクトル尖度の値が均等に分布されるように調整することで、6 種類のノイズ量が異なる刺激を作成した。本実験においてどのようなスペクトル尖度の値を使用すればよいかを決定するために、それぞれ異なるバンド幅のバンドパスフィルターを適用した 1,600 個のピンスノイズを生成した。バンド幅の値は 0.0001 オクターブから 15 オクターブまでで、それらの値を生成したピンクノイズに適用しノイズ量の異なる聴覚刺激を作成した。スペクトル尖度は MATLAB R2019a と MIRtoolbox 1.7.2[79] を用いて計算された。これによって図 3.3 に表すバンド幅とスペクトル尖度の関係図が作られ、この図を元にスペクトル尖度が 6 等分されるようなバンド幅の値が求められた。算出されたバンド幅の値は 0.0001, 0.00102, 0.0035, 0.167, 1, 15 オクターブで、これはスペクトル尖度

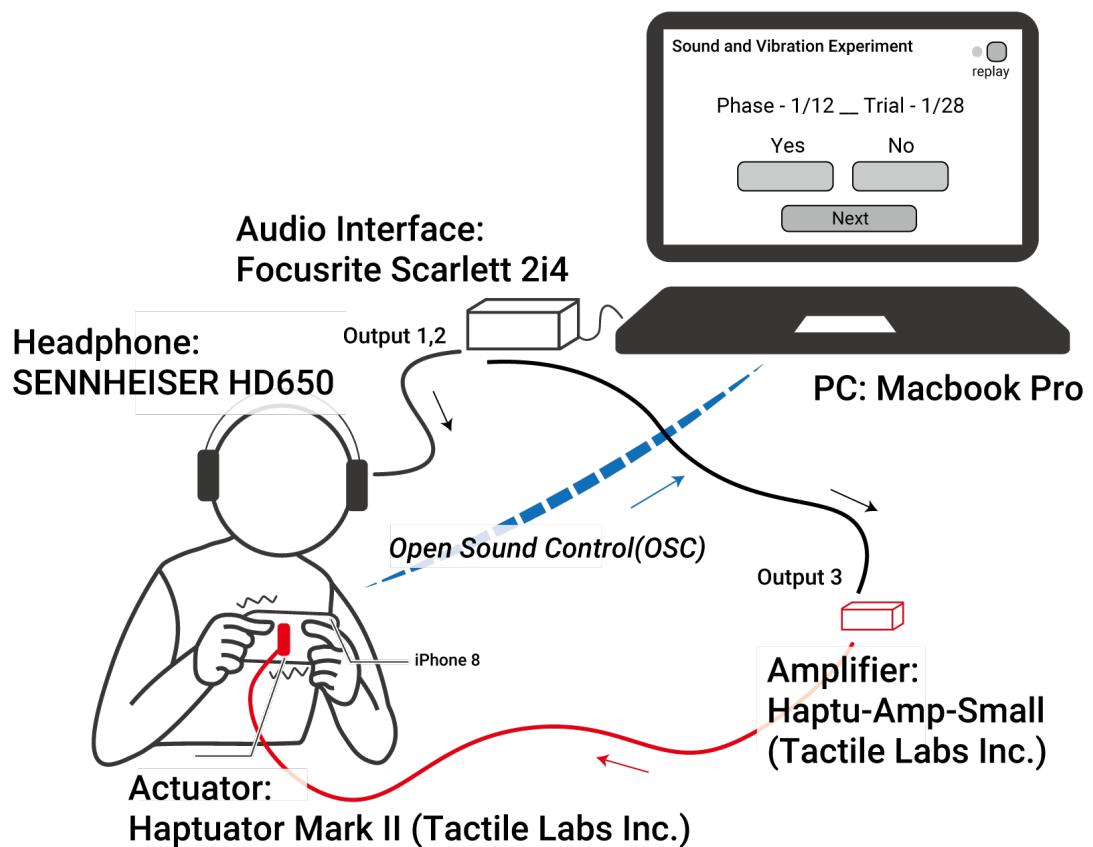


図 3.1 実験ダイアグラムと GUI



図 3.2 実験風景

における 2.1272, 5.1538, 9.4257, 12.9662, 16.2328, 20 に対応する。

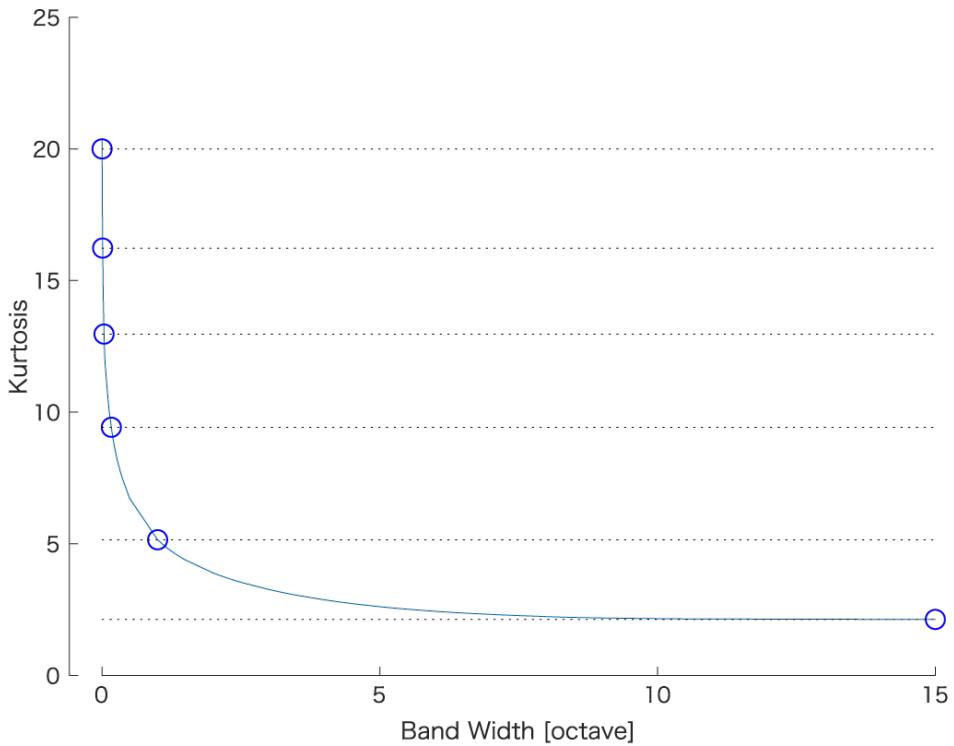


図 3.3 スペクトル尖度 (Kurtosis) とバンド幅の関係。青色の円は本実験にて使用されたパラメータ値を表す。

実験し使用した全ての音源は、人間の聴覚上等しく聞こえるように LKFS[80] を用いて -14 LKFS になるように調整された。聴覚刺激の音量は 1000 Hz を再生したとき疑似頭の聴取位置において 75 dB SPL であった。この実験で使用した触力覚刺激は 250 Hz の正弦波で、強度は 60 dB_{lv}[81] であった。触力覚刺激は全ての実験を通して周波数や強度が操作されることはなかった。聴覚刺激、触力覚刺激の長さは共に 50 ms だった。

聴覚刺激と触力覚刺激との間隔 (刺激開始時間差) は実験コンピュータ時間において -105 ms, -90 ms, -75 ms, -60 ms, -45 ms, -30 ms, -15 ms, 0 ms, 15 ms, 30 ms, 45 ms, 60 ms, 75 ms, 90 ms の合計 14 種類であった。刺激開始時間差の値が正の場合には聴覚刺激が触力覚刺激より遅れ、負の場合には聴覚刺激が触力覚刺激より早く開始することを意味する。

3.2.4 実験手続き

図 3.4 に実験の流れを示す。実験参加者は実験デバイスを両手で把持し、どちらかの親指で実験デバイス上に表示された GUI を操作することで回答を行った。各試行では 3.2.3 で述べた刺激が呈示され、その後実験参加者は聴覚刺激が先に呈示されたと感じたかそうでなかつたかを回答

した。刺激は GUI 上の “Next” ボタンを押した後 1–3 秒後に呈示された。実験参加者が聴覚刺激が先であると感じた場合は “Yes” ボタンを押し、触力覚刺激が先、または聴覚刺激と触力覚刺激が同時に再生されたと感じた場合は “No” ボタンを押した。実験刺激は “Replay” ボタンを押すことで何度も聴き直すことができた。実験参加者が 15 秒以内に回答が行えなかった場合は、その試行は無回答として記録された。実験は実験 1 が 6 種類、実験 2 が 6 種類の合計 12 種類あり、それぞれの種類で 14 種類の刺激開始時間差が 2 回繰り返され合計 24 回の刺激が呈示された。刺激開始時間差の値や刺激の種類は各実験においてそれぞれランダムで呈示された。実験終了後、実験参加者に対し 3 分ほどのアンケートを行い、年齢、性別、FLANDERS 利き手テスト [82] による利き手の調査を行った。

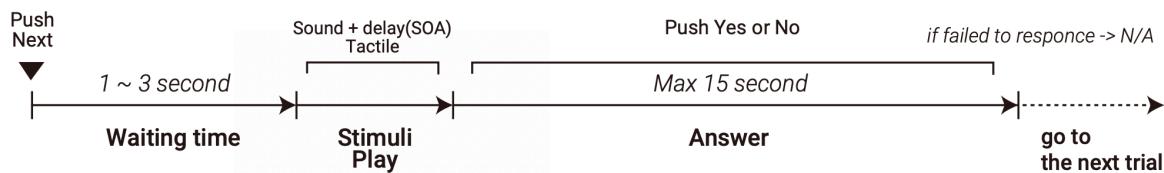


図 3.4 実験の流れ

3.3 結果

図 3.5 に分析方法の手順を示す。

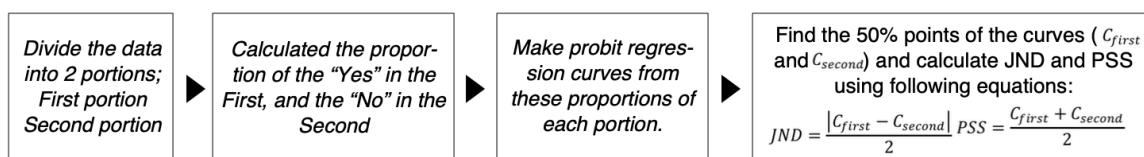


図 3.5 実験の分析方法

まず本実験において収集されたデータを 2 つのグループに分けた。3.2.4 で述べたように、本実験では 14 種類の刺激開始時間差の組み合わせが 2 回繰り返されていた。この繰り返しのうち、最初の 14 回を “First portion” とし、残りの 14 回を “Second portion” とした。その後、“First portion” での各刺激開始時間差における “Yes” と回答された割合と、“Second portion” の “No” の割合を算出した。このようにして得られた値は図 3.6 における赤色のアスタリスク及び青色の丸のようにプロットした。

各グループにおける回答率から正規累積分関数を用いて曲線当てはめを行い、算出された曲線を用いて丁度可知差異と主観的同時点を算出した。

この実験における丁度可知差異は、刺激の遅延値のうち、50 パーセントの割合で識別できる最小の刺激強度変化量を意味し、これは触力覚刺激に対する聴覚刺激の遅延許容度 (= 時間統合窓)

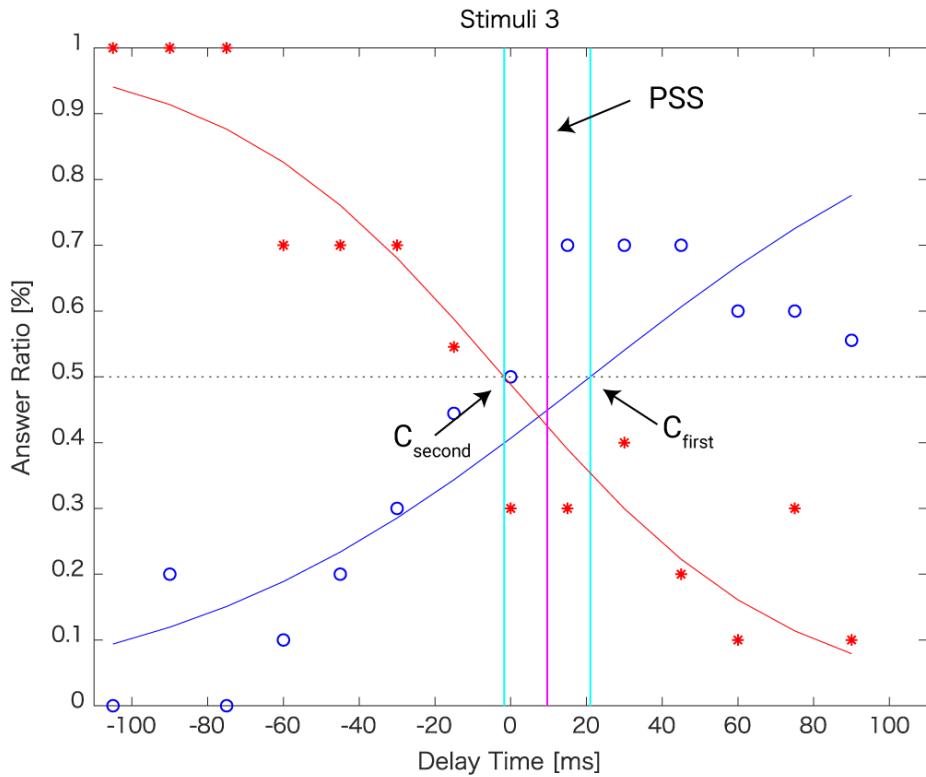


図 3.6 実験刺激が 1 kHz の正弦波のときの実験結果。赤色の左上から右下へ向かう線は “First portion” における回答率 (赤色のアスタリスク) を元に当てはめられた曲線で、青色の左下から右上に向かう線は “Second portion” における回答率 (青色の丸) から当てはめられた曲線である。これらの当てはめられた曲線と水色の縦線の交点は、 C_{first} と C_{second} を表し、水色の縦線の間のピンク色の線は主観的同時点 (PSS) を表している。

を表すことになる。一方主観的同時点は、主観的に聴覚刺激と触力覚刺激が同時に発せられたと判断される聴覚刺激の遅延量を示す。したがって、コンテンツ製作者が聴覚-触力覚刺激を用いた表現を作ろうとする際に、参考となる遅延時間の一つとなり得ると考えられる。

丁度可知差異及び主観的同時点は以下の式を用いて計算した。

$$\text{JND} = \frac{|C_{\text{first}} - C_{\text{second}}|}{2} \quad (3.1)$$

$$\text{PSS} = \frac{C_{\text{first}} + C_{\text{second}}}{2} \quad (3.2)$$

C_{first} は “First portion” における “Yes” の回答率が 50 パーセントである点を表し、 C_{second} は “Second portion” における “No” の回答率が 50 パーセントである点を表している。図 3.6 における水色の縦線が C_{first} 及び C_{second} に対応する。

図 3.7 は本実験の分析結果である。実験 1 の分析結果から、丁度可知差異及び主観的同時点の値は聴覚刺激の周波数の上昇と伴って上昇する傾向を見せることがわかった。一方、実験 2 にお

いては、バンド幅の値と丁度可知差異及び主観的同時点の値との関係性を導き出すことはできなかった。また、実験終了後に収集したアンケートから特異な結果を得ることはできなかった。

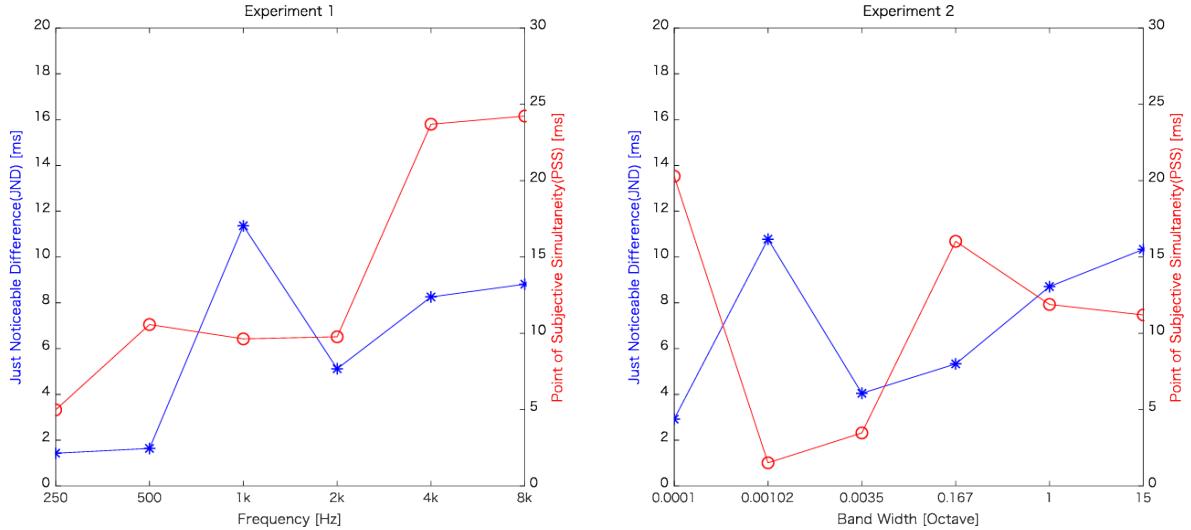


図 3.7 本実験で算出された丁度可知差異及び主観的同時点。青色のアスタリスクは丁度可知差異を表し、赤色の円は主観的同時点を表す。

3.4 考察

3.4.1 聴覚の処理過程と主観的同時点の関係

実験 1 の結果から、丁度可知差異及び主観的同時点の振る舞いは聴覚知覚の認知システムと同じ仕組みで説明できるのではないかという仮説を立てた。実験 1 では、聴覚刺激の周波数が低い場合、主観的同時点も低くなり、逆も然りであった。3.3 で述べたように、主観的同時点は聴覚刺激と触力覚刺激が同時に知覚される遅延値を示している。そのため、主観的同時点が低いということは聴覚刺激が低周波音であった場合、それが高周波である場合よりも触力覚刺激が聴覚刺激と混ざり合って聞こえうるという可能性が考えられる。

この現象を説明する例の一つに、両刺激の周波数が近いことに起因する協和感の誘起を挙げることができる。

岡崎龍太ら [83] は、聴覚と触覚の感覚構造が異なるにもかかわらず、聴覚と触力覚の周波数の知覚的な協和感を得ることができ、両感覚の周波数関係が近いほどその感覚が強まることを報告している。そのことから、触力覚刺激で用いた 250 Hz の正弦波に近い 250 Hz や 500 Hz の聴覚刺激は、その刺激が呈示された際に知覚的な協和感を覚えたことで両刺激が混ざり合ったと知覚され、相対的に主観的同時点が低い値をとったのではないかと考えられる。

また、聴覚の構造がマルチモーダル知覚における協和感に与える影響についても言及することができる。Parker[84] は、蝸牛内において高周波音は低周波音に比べ感度が高く、より早く知覚

されると報告している。このため、刺激開始時間差の値が同じであっても、聴覚刺激の周波数が高い場合蝸牛内で低周波音よりも早く音が処理されることで、結果的に脳の時間統合部分において聴覚刺激が触力覚刺激よりも早く到達していた可能性が考えられる。

3.4.2 聴覚刺激の周波数と丁度可知差異の関係

実験1の結果から、純音の周波数が低いほど丁度可知差異が低くなり、その逆も然りであった。従って、実験参加者は低周波音の刺激の方が高周波のものよりも刺激の遅延時間の違いをより敏感に知覚できたと考えられる。

このような周波数によって刺激の同時性の振る舞いが変化する現象は、マルチモーダル知覚においては報告された例はほとんどないが、両耳聴における同時性知覚の研究で報告された例がある。岡崎聰ら [85] は2つの聴覚刺激の周波数間隔が広いと、それに伴い丁度可知差異が大きくなることを報告した。

この例から、岡崎龍太ら [83] によるマルチモーダル刺激の協和感知覚のように、聴覚-触力覚間であっても両耳聴環境と似たような同時性に関わる振る舞いを見せるのであれば、岡崎聰ら [85] が報告した丁度可知差異に関する知覚現象も、マルチモーダル知覚の環境においても観測することができるかもしれない。

3.5 結論

本実験では、聴覚刺激の周波数及び聴覚刺激のノイズ量を変化させた時の丁度可知差異及び主観的同時点の変化について調査した。その結果、聴覚刺激の周波数が低い場合、丁度可知差異と主観的同時点が小さくなる傾向があることがわかった。これは、実験参加者が刺激の周波数関係によって両刺激が混ざり合ったと知覚された可能性が示唆された。また、聴覚-触力覚の同時性知覚は、両耳聴下での同時性知覚から説明できる可能性を得られた。一方、本実験においてはスペクトル尖度と丁度可知差異及び主観的同時点の関係性を見出すことはできなかった。

次に行なった実験では、本実験において示唆された聴覚-触力覚と両耳聴下の同時性知覚の類似点を検証するために、実際に岡崎聰ら [85][86] が行なった研究を聴覚-触力覚で再実験するような形で行なった。

第4章

聴覚-触力覚刺激の周波数間隔を変化させた時の時間統合窓の変動

3.4で述べたように、聴覚-触力覚の同時性知覚は両耳聴下での同時性知覚で見られる現象と似た傾向を用いる可能性が示唆された。このことから、本研究では両耳聴下での同時性知覚に見られる特異な時間統合窓の変動を報告した岡崎聰らの研究成果 [85][86] を、聴覚-触力覚条件下でも見られるのかどうかを検討した。

4.1 背景

岡崎聰ら [86] の研究では、周波数の異なる2つの純音をそれぞれ異なる耳に呈示される基準刺激と対象刺激として同時性判定課題を行った。実験刺激は持続時間が平均300 msの正弦波で、低い周波数を200 Hzと固定し、高い方の周波数を210 Hzから3413 Hzまで12段階で変化させた。そのうちのいくつかの刺激組み合わせは、人に調性的協和や音楽的協和の感覚を生じさせる特別な周波数距離、また調波関係と呼ばれる特別な周波数距離となるように調整されている。

実験の結果、時間統合窓^{*1}が「V字型」の挙動(図4.1)を示すことが分かった。これはすなわち、対象刺激の周波数がある一定の周波数になるまでは時間統合窓が急峻に低下し、その周波数を過ぎると時間統合窓が徐々に上昇する振る舞いを両耳聴下での同時性知覚で見せるということである。

また、同時性の範囲が調性的協和や音楽的協和、また調波関係に影響を受けないことも報告した[86]。調性的協和とは基礎的な感覚レベルでの協和感を意味し、音楽の文脈から切り離された、音そのものに対する協和感のことである[87]。また、音楽的協和とは完全5度や完全8度など、音楽理論によって2音の周波数比が単純な整数比となる周波数距離にあるものを指し、調波関係とは2音の周波数が整数倍の関係となる周波数距離のことを指す。

^{*1} 岡崎聰らの文献では、知覚的同時性範囲 (Point of Simultaneity range: PSR) と記載されている。

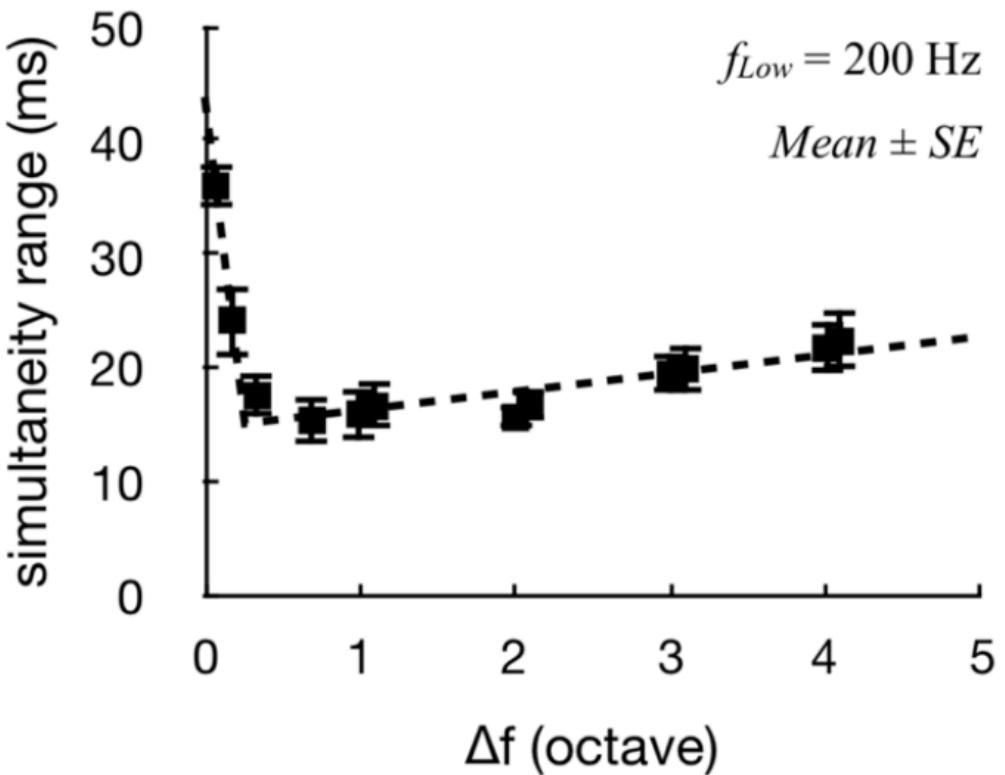


図 4.1 両耳間条件での同時性の範囲と周波数距離の関数(岡崎聰ら [86] より引用)

このように岡崎聰ら [86] での研究では特別な周波数距離に影響を受けないと述べる一方、聴覚-触力覚の条件下においてはこれに対となる報告がなされている。岡崎龍太ら [88] では聴覚-触力覚の条件下の場合、両刺激の調性的協和がその同時性知覚に影響を及ぼしうる可能性について報告している。そこでは、鋸歯状波を用いて聴覚-触力覚間で主観上「同じ」と感じる周波数の組み合わせが存在するかどうかを調査した。その結果、聴覚-触力覚間の周波数的な協和度は両刺激が倍音構造をなしているか否かで判断され、知覚構造が異なる聴覚-触力覚間においても、聴覚で感じる和音の協和度が知覚可能であると報告した。

仮に聴覚-触力覚の時間同期感覚も岡崎聰らの提唱する「V字型」を有するのであれば、聴覚-触力覚の時間同期感覚も「V字型」の傾向を見せる可能性を持つ。しかしながら、その場合では聴覚-触力覚の調性的協和判断の点において矛盾することとなる。

そこで本章では、マルチモーダル刺激が周波数的に分離している状態のときの、聴覚-触力覚同期に対する時間統合窓の変化を調べた。研究仮説は次の通りである。(1) 聴覚-触力覚間の時間統合窓は、岡崎らによって述べられた聴覚現象と同様に、時間統合窓の「V字形」特性の傾向を有し、また(2) 時間統合窓は、特別な周波数距離に存在する、調性的協和、音楽的協和、調波関係の影響を受けない特徴を持つ。

4.2 実験手法

実際の聴覚-触力覚マルチモーダルコンテンツで使用されている刺激は、銃撃音や自動車のエンジン音など多岐に渡り、それに付けられる振動も様々である。しかしながら、このような複雑な音を用いて実験を行った場合、音のどの周波数、あるいはどの周波数組み合わせが影響したか特定することが難しい。そのため本実験では、1つの周波数の正弦波からなる純音を用いて、聴覚に呈示される純音、触力覚に呈示される純音の知覚的な同時性の範囲を測定する。この聴覚と触力覚の2つの純音にはあらゆる周波数の組み合わせが可能であるが、この周波数組み合わせを系統的に表現するためには、2純音の周波数軸上の距離 Δf (単位: オクターブ) と、基準とする2純音のうち低い方の周波数(本実験では触力覚刺激)の2つを記述すれば良いため、以降ではそのように記述する。

4.2.1 実験参加者

東京芸術大学の学部生及び大学院生7名(男性5名女性2名、22歳から28歳)が実験に参加し、そのうち1名は著者であった。実験時間はおよそ90分であった。なおすべての実験参加者は自己申告により、日常生活に支障が出るほどの聴覚障害及び手指の触力覚感覚障害がないことを確認した。実験に使用した実験参加同意書を付録Dに示す。

4.2.2 実験装置

実験装置は3.2.2に記載されているものとおおよそ同じである(図4.2)。しかしながら、実験実施場所の都合によりオーディオインターフェイスを変更した。また、実験を行なった部屋は、簡易防音室から静かな教室に変更となっている。

4.2.3 実験刺激

実験刺激は岡崎聰ら[86]による実験を踏襲する形で作成された。刺激は周波数の異なる2つの純音で構成された(図4.3)。触力覚刺激は200Hzに固定され、聴覚刺激は210, 226, 253, 323, 400, 427, 800, 853, 1600, 1707, 3200, 3413Hzであった。聴覚刺激と触力覚刺激の周波数距離 Δf をオクターブで表すと、それぞれ0.07, 0.18, 0.34, 0.69, 1.00, 1.09, 2.00, 2.09, 3.00, 3.09, 4.00, 4.09オクターブであった。

このうち、 $\Delta f = 0.07, 0.18, 0.34, 0.69$ オクターブの4つの周波数距離は、調性的協和による時間同期範囲の影響を調べるために選択された。図4.4に示すように、調性的協和は周波数距離が0.25臨界帯域幅(単位:CB)に近づくにつれ低下し、0.25CBから1.2CBで上昇する[89]。本実験で用いられた上記4つの周波数距離は、調性的な協和度がそれぞれ0.5, 0, 0.5, 1となる周波

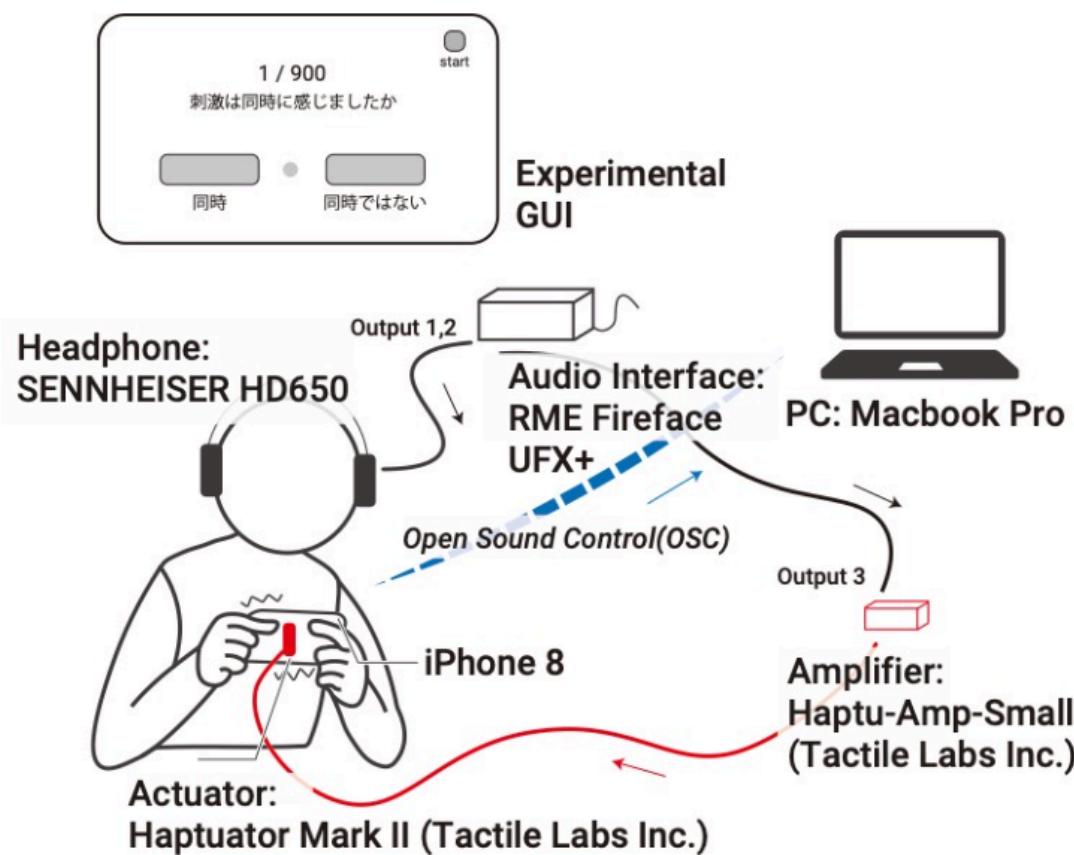


図 4.2 実験装置模式図及び実験 GUI

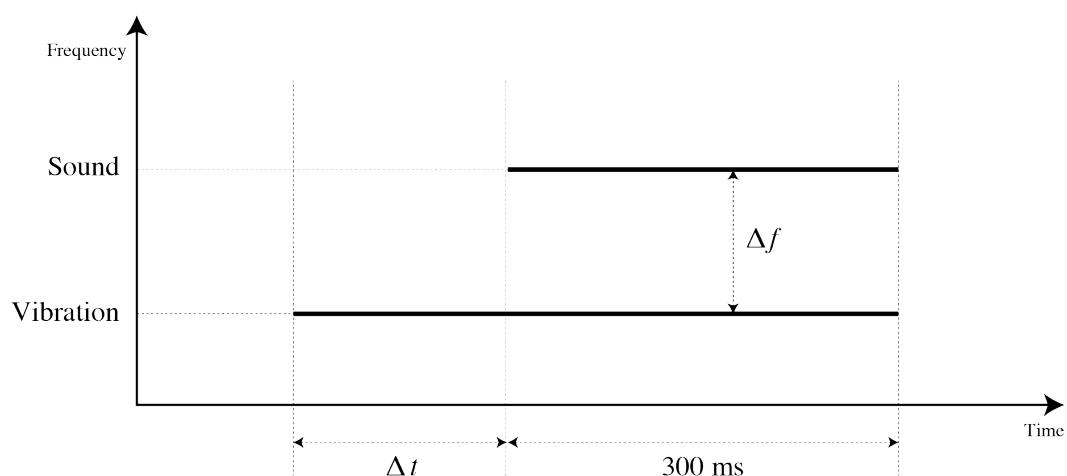


図 4.3 実験で用いた 2 純音刺激の模式図

数距離に対応していた。

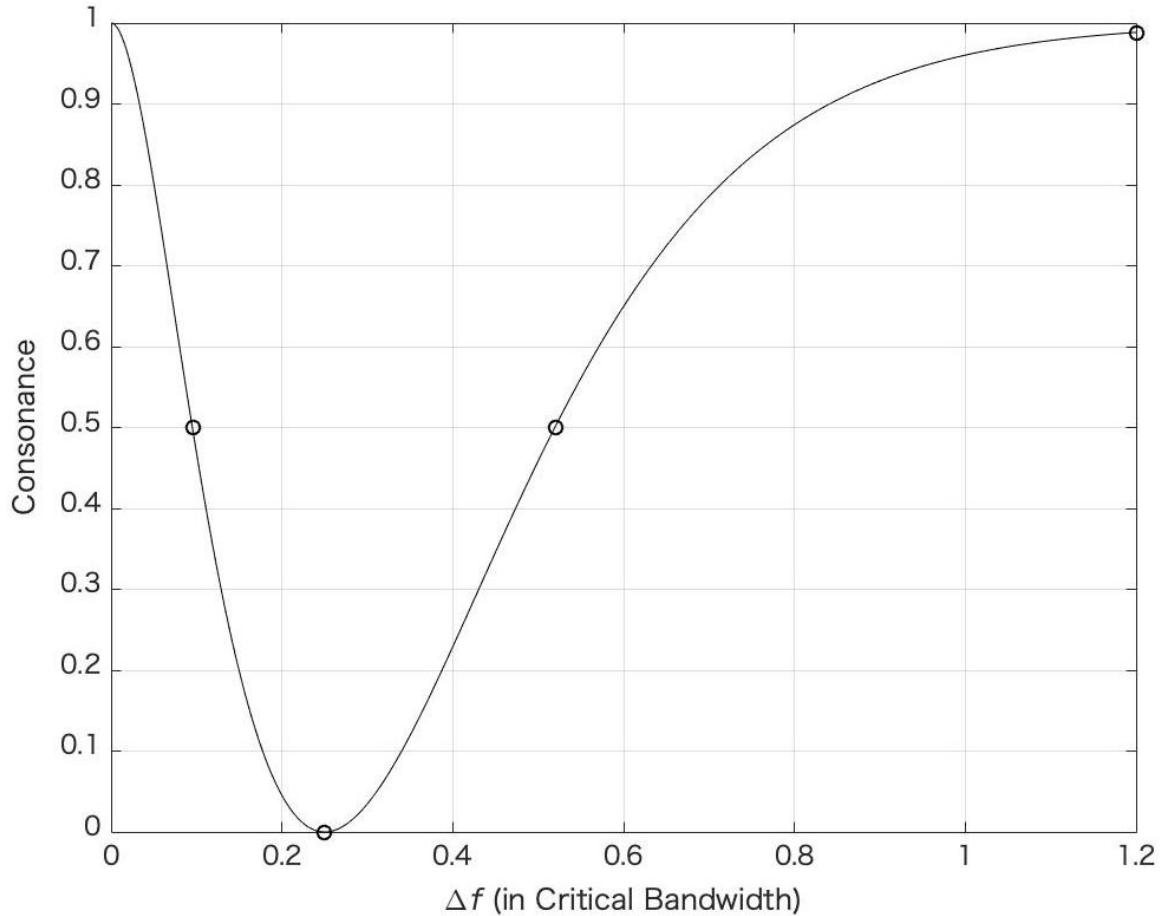


図 4.4 調性的協和曲線 [89] の近似関数 [90]。図中の丸は本実験で選択した 4 つの周波数距離を表す。

周波数距離 $\Delta f = 1.00, 2.00, 3.00, 4.00$ オクターブの 4 つの周波数距離は、音楽的協和かつ調波関係にあった。また、 $\Delta f = 1.09, 2.09, 3.09, 4.09$ オクターブの 4 つの周波数距離は音楽的に不協和かつ非調波関係にあった。

聴覚刺激と触力覚刺激は実験コンピュータ時間における同時または非同時に開始した。その値は $\Delta t = \pm 0, 2, 4, 7, 14, 27, 52, 100$ ms の 15 種類である。 Δt が正の場合には聴覚刺激が触力覚刺激より遅れ、 Δt が負の場合には聴覚刺激が触力覚刺激より早く開始することを意味する。2 音の終了は常に同時であった。これは、終了の時間差を手がかりに聴取者が同時性判断を行うことを避けるためであった。後続音の持続時間は 300 ms であり、先行音の持続時間は、300 ms + Δt (絶対値) であった。各音の開始と終了には 15 ms で無音となる傾斜がつけられ(コサインテーパー処理)、ゆるやかな立ち上がりあるいは立ち下がりとなった。こうした刺激作成上の工夫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の雑音が発生するのを抑えるために行われた [91]。

すべての刺激は数値解析ソフトウェア (MATLAB R2019a) によって作成され、また聴覚刺激は数値解析ソフトウェアを用いて、聴感上同じ音量になるよう予め処理を行った。この処理に用

いたラウドネス指標は LKFS[80] で、値は -14 LKFS であった。上記実験環境での聴覚刺激の音量は 1000 Hz を再生したとき疑似頭の聴取位置において 75 dB_{SPL} であった。

4.2.4 手続き

実験参加者は実験デバイスを両手で把持し、どちらかの親指で実験デバイス上に表示された GUI を操作することで回答を行った。各試行では、2 つの正弦波で構成される刺激 (図 4.3) が呈示され、その後、実験参加者は両刺激が同時に発せられたと感じたかそうでなかつたかを二者択一によって回答した。刺激が呈示された後、実験参加者は 4 秒以内に回答を行う必要があった。実験参加者は回答後、その回答に関するいかなるフィードバックも与えられなかつた。時間制限以内に回答が行えなかつた場合は、その施行は無回答として記録された。15 種の Δt における各 Δf 条件は、それぞれ 10 回再生され、すべての再生順序はランダムであった。実験参加者は 150 試行に一度、最低 1 分以上の休憩をおこなつた。また実験を行う前に、実験参加者はランダムに選ばれた 30 条件でフィードバックなしの練習を行つた。実験終了後、実験参加者に対し 3 分ほどのアンケートを行い、年齢、性別、FLANDERS 利き手テスト [82] による利き手の調査を行つた。

4.3 結果

全 12,600 試行データ (Δf 12 種 \times Δt 15 種 \times 10 回 \times 7 人) のうち、14 試行のデータは、回答時間超過による無回答により省かれ、分析に使用しなかつた。また、聴取者が実験の目的を予め知りすべての施行で非同時と判断すれば、 $\Delta t = 0$ 以外のすべての水準で非同時であったと答える最も正確な同時性判断ができるが、そのような判断をおこなつた聴取者はいなかつた。

まず、各周波数距離条件ごとに、聴覚-触力覚が同時に開始したと聞こえた判断の割合を求め、 x 軸を Δt 軸、 y 軸を判断の頻度としたグラフを条件ごとに作成した。このようにして得られたグラフを図 4.5 に示す。ここでの各パネルにおける値から線を結び、これを Δt 条件において刺激が知覚的に同時/非同時であると判断される境界とした。例えば、この線の内側にある領域は、周波数距離条件において同時と判断される頻度を表すことになる。また、得られた各条件ごとの判断の割合から分布の標準偏差を求め、それを時間統合窓の指標とした [34]。

以上のように求められた時間統合窓を図 4.6 に周波数距離の関数としてプロットした。図中の丸は周波数距離が調性的協和であるもの、プラスは音楽的協和かつ調波関係にあるもの、アスターiskは音楽的に不協和かつ非調波関係にあるものを示している。周波数距離よつて同時性の範囲に違いがあるかを調べるため分散分析を行つたが、有意差は認められなかつた ($p = 0.12$)。また実験終了後に収集したアンケートから特異な結果を得ることはできなかつた。

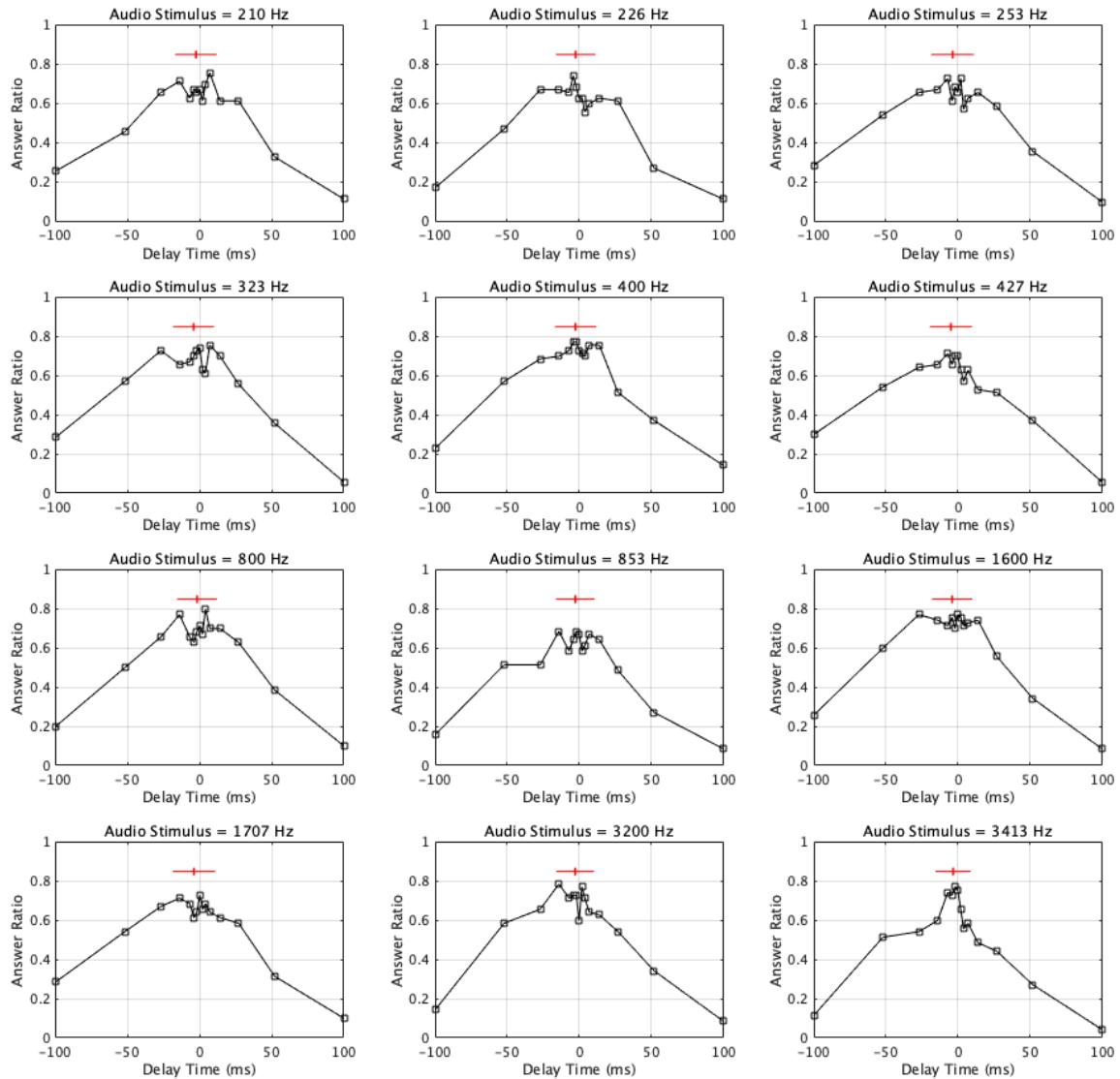


図 4.5 聴覚-触力覚が同時に開始して聞こえた判断の頻度分布。各パネルの縦軸はこれらの判断の割合を、横軸は Δt 水準を表している。各パネル上部の線はデータの平均及び標準偏差を表す。

4.4 考察

4.4.1 時間統合窓の「V字型」傾向

本実験では、聴覚-触力覚の 2 純音の時間統合窓が、岡崎聰らが提唱する時間統合窓の「V字型」の傾向及びそれらが特別な周波数距離に影響を及ぼすかどうかを調査した。その結果、岡崎聰らの結果(図 4.1)とは異なり、聴覚-触力覚の環境(図 4.6)では「V字型」の傾向を見ることはできなかった。これは両耳間と聴覚-触力覚間での刺激の処理過程の差異が理由の一つと考えられる。

岡崎聰らでは、聴覚における「V字型」の折れ点が聴覚末梢系にある臨界帯域幅によって決定

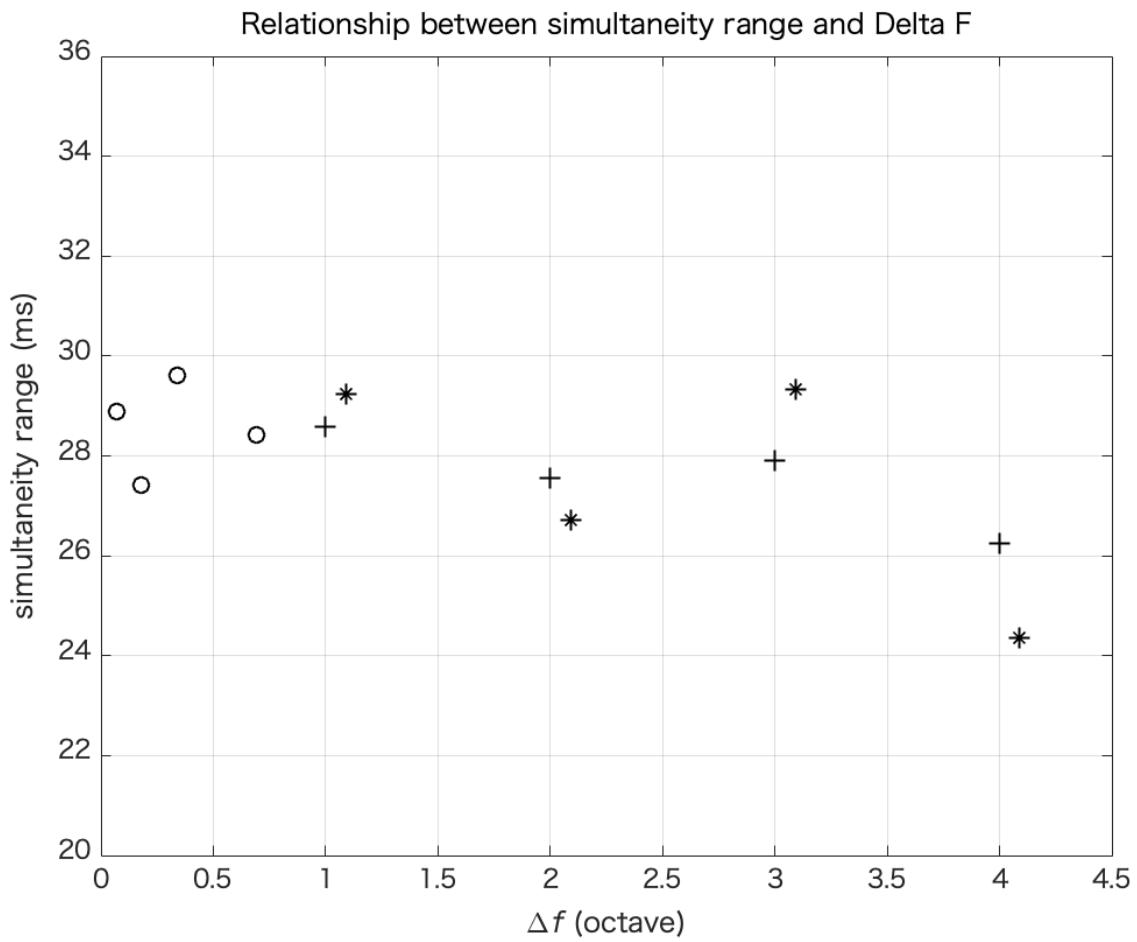


図 4.6 本実験において算出された時間統合窓と周波数距離の関係

されるのではないかと考察した [86]。2つの非同時に開始する純音が基底膜振動に変換された時、2音の周波数距離が近い場合には、基底膜上での振動の干渉が起こり2音の開始地点が同時となる。このことがV字型の左側での2音の周波数距離が近い場合における時間統合窓の幅の上昇の原因になっていると考察した [86]。

一方聴覚-触力覚の場合では、それぞれの刺激が別の感覚受容器で処理され、脳の高次の部分でそれらの信号が統合される。いくつかの研究で聴覚-触力覚の刺激の統合が比較的初期の段階において起こりうる可能性について議論されている [92][93] が、こうした異なる感覚情報がどのように統合され、刺激の同時性を形成しているのかは未だ研究段階だと見える。聴覚-触力覚環境におけるこのような処理過程から、感覚受容の初期段階での干渉がマルチモーダル感覚情報の統合過程上生じ得ないと言えるため、時間統合窓の「V字型」傾向は見られなかったと考えられる。

4.4.2 実験方法における判断バイアス

本実験の結果を線形回帰すると、時間統合窓が右下がりの振る舞いを見せることがわかった(図4.7左)。これは、周波数距離が大きいほど聴覚-触力覚の両刺激が同時であると知覚される許容幅が狭まる傾向にあると見て取れる。しかしながら、この結果は3.4.2で述べられた、聴覚刺激の周波数が低い時に時間統合窓が狭くなるという傾向(図4.7右)と反するものである。このような差異が生じた原因の一つとして、実験方法の違いによる判断バイアスが考えられる。

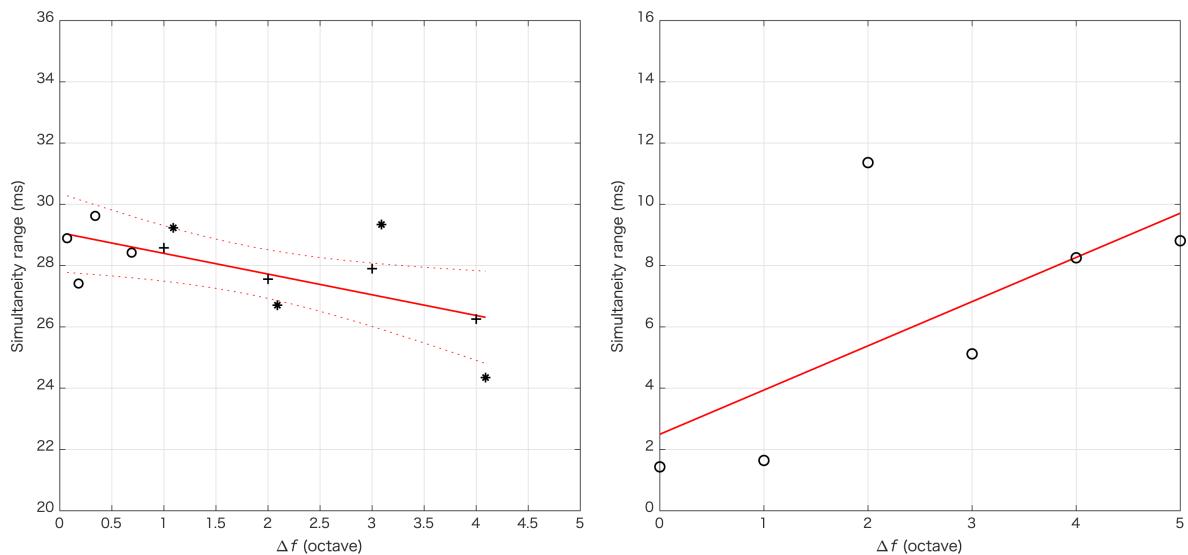


図4.7 本実験(左)と前回の実験(右)らの結果の比較。図中の直線は線形回帰線を表す。

本実験では、4.2.3で述べたように岡崎聰らによる実験を踏襲しており、実験手法も本実験上必要のない過程を除き同じものを用いた。そのため、今回は2つの刺激が同時と感じられたか否かを直接判断するように求める同時性判断課題を使用しているが、前回の実験では、2つの刺激の提示順序を判断するように求める時間順序判断課題が採用されていた(実験手法の詳細については2.1.2を参照)。同時性判断課題による実験は直接的に同時性知覚を問うことができる一方で、同時か否かの判断が極めて主観的で、判断基準のゆらぎが直接実験結果に反映されるため、知覚実験という観点からは判断バイアスが生じやすい方法と考えられている[4][16]。

対して時間順序判断課題は、2つの刺激によりそれぞれ独立に形成される心理連続体上の位置を単純に比較し、その差が正か負かの強制判断で結果が得られるため、同時性判断課題で見られるような判断基準の揺らぎによる判断バイアスの懸念が少ないとされている。その一方で、判断が記憶に基づくことになるため、認知バイアスを受けやすい面があるとされる[4][16]。

マルチモーダル感覚情報に基づく事象は高次の認知処理であると仮定され、同時性判断課題のような認知バイアスの懸念が少ない実験手法が用いられる。しかし4.4.1で述べたように、マルチモーダル事象知覚の背景となる、複数の感覚情報の統合と相互作用が初期の段階において生

じている例が少なくないことが明らかになってきているため、マルチモーダル感覚情報に関する精神物理学的実験においては、実験結果として得られたものが知覚に基づくものなのか、判断バイアスによるものなのかを厳しく区別する必要がある [16]。本研究における実験参加者の多くは、日常的に聴覚と触力覚の時間同期を注意深く判断する環境に属していなかったため、回答に悩んだ際の回答方法などの判断バイアスが大きく生じていた可能性がある。今後は近年の生理学的研究を参照し、感覚情報の統合が処理のどの段階で生じうるのかを推察した上で判断バイアスの影響を受けにくい測定法を選択する必要がある。

4.5 結論

本実験では、モバイルデバイスを手で把持する状況を想定し、純音を用いた聴覚刺激と触力覚刺激の周波数間隔を変化させた時の、両刺激における時間統合窓の変動について調査した。その結果、聴覚のみの同時性判断で見られる時間統合窓の「V字型」傾向は見られなかった。また、調性的協和及び調波関係といった特定の周波数距離に関わらず、周波数間隔の増加に伴い時間統合窓が若干減少する傾向が観測された。しかしながら、本実験結果は第3章の研究結果と反するものであり、その要因の一つとして実験時の回答手法における判断バイアスが大きく影響したと考えられる。そこで次回の実験では、判断バイアスの影響を考慮した新しい実験手法によって、時間統合窓の変動をさらに調査する。

第 5 章

聴覚刺激の強度を変化させた時の時間 統合窓の変動

4.4.2 で述べたように、同時性判断課題や時間順序判断課題などの心理物理学的手法による同時性知覚の実験は、実験参加者の判断バイアスに大きな影響を受けやすく、筆者が行った前述の 2 つの実験においても結果に大きなばらつきが生じた。そこで本章では、ベイズ推論による信号検出理論という実験手法を用いることで、従来の手法で生じていた問題を取り除くことを試みつつ、聴覚刺激の強度を変化させた時の時間統合窓の変動について検討する。

5.1 背景

2.1.2 で述べたように、複合感覚の刺激が統合されて一つの事象として知覚される時、その複数刺激が一つの事象として認知されうる時間的許容幅のことを時間統合窓と呼ぶ。この時間統合窓はさまざまな要因によって変化するが、いまだ要因と時間統合窓間の関係など、詳細な振る舞いまでは明らかにはっていない。そこで今回は、聴覚刺激の強度を変化させた時、時間統合窓に変化があるのかどうかを調べた。

一般に時間統合窓を求める方法として、同時性判断課題、もしくは時間順序判断課題がある [4][32]。しかしながら 4.4.1 で述べたようにマルチモーダル事象知覚の背景となる、複数の感覚情報の統合と相互作用が初期の段階において生じている例が少なくないことが明らかになってきているため、マルチモーダル感覚情報に関する精神物理学的実験においては、実験結果として得られたものが知覚に基づくものなのか、判断バイアスによるものなのかを厳しく区別する必要がある [16]。そこで、本実験ではこの問題を解決すべくベイズ推論による信号検出理論という実験手法を用いて、時間統合窓の振る舞いを調査することを提案する。

5.1.1 信号検出理論

信号検出理論は、1950 年代にレーダーシステムの通信工学的理論として考案された。そもそもはノイズに埋もれた信号の検出力をこの理論を用いて調べ、レーダーの性能評価をするために作られた。今日ではノイズと信号を、無視すべき刺激と知覚すべき刺激とにそれぞれみなし、心理物理的測定法に応用したものが一般に使用されている [94]。信号検出理論は同時性判断課題などに代表される閾値の測定という手法と比べると、被験者のバイアスとそれを取り除いた弁別力を同時に計測できる利点がある [95]。

信号検出理論用いて弁別力を求めるにはいくつかの仮定が必要である [94]。例えば、音圧とその物理量に対応したラウドネスのように、物理的な刺激は被験者の内部で、その強度に対応した 1 次元の心理量に変換される。そしてノイズだけの試行 (N) とノイズと信号が両方呈示される試行 (S+N) では、後者の方が物理エネルギーが高い分、得られる心理量の平均値も大きいとする。また同等の刺激であっても、被験者の内部で常に同じ心理量を与えられるかというとそうではなく誤差があり、全体としてはその誤差は正規分布を取るとする（正規性の仮定）。そしてノイズのみの正規分布も、ノイズ + 信号の正規分布も同じ分散を持つと仮定する（等分散性の仮定）。このような条件下、ノイズの中で信号の有無の判断を求める場合、信号の有無（信号がある場合 [S+N]、信号がない場合 [N]）と被験者の判断（信号があると判断した [s]、信号がないと判断した [n]）の組み合わせは以下の 4 通りになる。

信号 \ 反応	s	n
S+N	hit	miss
N	false alarm	correct rejection

1. hit : 信号があるときに、正しくその信号を検出した場合
2. miss : 信号があるにもかかわらず見落とした場合
3. false alarm : 信号がないにもかかわらず信号があると誤って反応した場合
4. correct rejection : 信号がないときに、「信号なし」と正しく反応した場合

図 5.1 信号の有無と被験者の判断の組み合わせ

これらの関係を図 5.2 に表す。横軸は一次元の心理量であり、ノイズ + 信号の正規分布の方が心理量が大きい分、右に寄っている。一方縦軸は実験参加者がその心理量に対して行う判断の頻度である。それぞれの正規分布の平均値が、物理量に対応した心理量であるので、その値を得る頻度が最も多く、平均値からずれるに従って頻度は減っていく^{*1}。図 5.2 の中央付近ある赤色の縦線

^{*1} 例えばカレー（ノイズ）に唐辛子入れたとき（ノイズ+信号）を想定する。心理量を辛さとすると、この心理量は 0 からプラス方向に離れた A という点を中心に分散 1 の正規分布を描く。この時の A、つまり正規分布の頂点は、実験参加者がただのカレーを食べて辛さが A であるという判断が一番多かったことを表している。また、もちろんカ

は被験者の判断の境目 (基準: k) である。

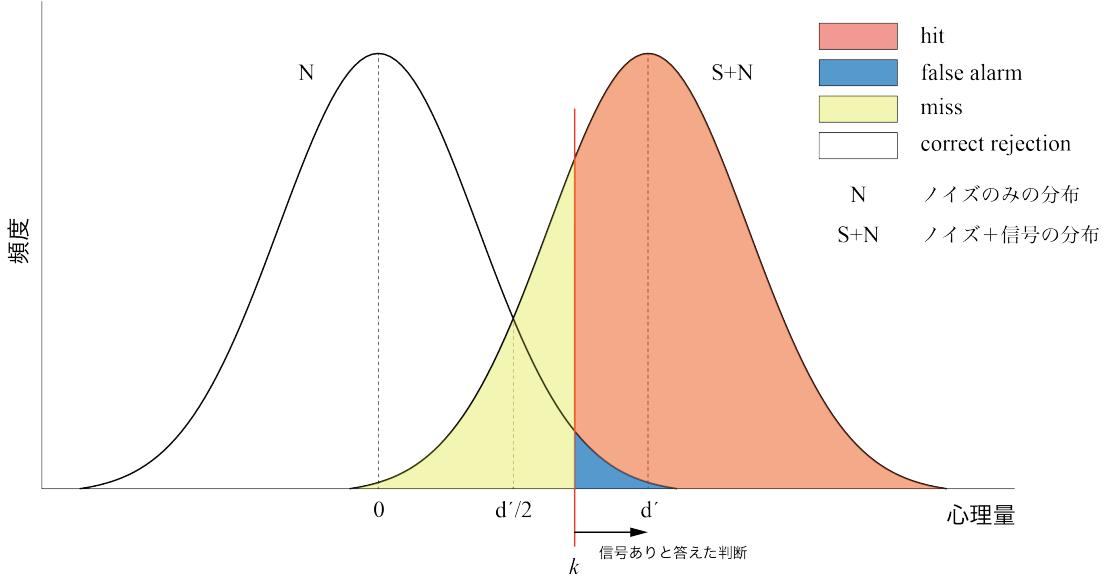


図 5.2 信号がある場合、信号がない場合に対する判断の分布 (難波 [95]、Lee *et al.*[96] を改変)

そうすると、ノイズ + 信号の正規分布のうち縦線よりも右側の部分の面積は hit の比率と等しくなり、一方 N の正規分布のうち縦線よりも右側の部分面積が、false alarm の比率となる。

これら 2 つの比率は実験によって計測される。ここで、二つの正規分布の平均値およびその間隔と基準の位置、また基準と正規分布の平均値との距離がそれぞれ算出することができる。信号検出理論では、ノイズ分布の平均 (M_0) とノイズ + 信号分布の平均 (M_1) との距離をノイズ分布の標準偏差 s で割った値

$$d' = \frac{M_1 - M_0}{s} \quad (5.1)$$

を被験者のその刺激についての弁別力の指標 (信号検出力、discriminability: d') と定義している。ここで、信号検出理論での分布は正規分布を仮定しているので、 $M_0 = 0$, $s = 1$ であり、 $d' = M_1$ となる。ノイズ分布に対するノイズ + 信号分布の位置は、刺激の強度と感覚系の特性に依存するため、 d' は実験参加者ごとに異なる判断基準の位置 k の影響を受けない。したがって、 d' は判断基準の位置とは全く独立な、信号検出の感度を示す指標となる。

さらに、判断基準の位置 k と $d'/2$ との差

$$c = k - \frac{d'}{2} \quad (5.2)$$

レーに唐辛子入れたとき (ノイズ+信号) は、カレー (ノイズ) だけの時と比べて辛さが大きくなるので、カレーに唐辛子入れたとき (ノイズ+信号) の正規分布はカレー (ノイズ) だけの正規分布より右に寄る。

によって被験者の反応の偏りである反応バイアス (bias: c) が求められる。図 5.2 からわかるように、 $d'/2$ はノイズ分布とノイズ + 信号分布が交わる点に一致する。仮に $k = d'/2$ であれば、 $c = 0$ であり、この時の状態は、hit の確率と correct rejection の確率が等しく、さらに false alarm と miss の確率も等しくなる。この意味で $c = 0$ ならば、被験者の反応に偏りがない状態(反応バイアスのない状態)を表すことになる。図 5.3 の左右の図を比較すると、判断基準の位置 k が小さな値となるほど、信号があると反応する確率が高いことを意味し、hit の確率が高くなる一方で、false alarm の確率も同時に高くなることがわかる。図 5.3 左のように $c > 0$ ならば $d'/2 < k$ の状態であり、信号がないと判断する傾向にあると解釈することができる。一方、図 5.3 右のように $c < 0$ ならば $d'/2 > k$ の状態にあり、信号があると判断する傾向が強いと判断できる。

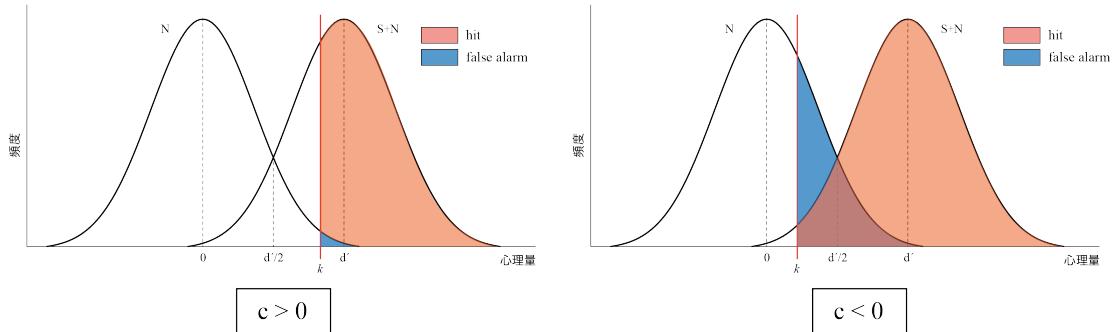


図 5.3 弁別力と反応バイアス (豊田 [97] を改変)

上述の通り、信号検出理論を用いる際には正規性と等分散性の仮定が満たされてなければならない。しかしながら、これらをの仮定を検証するためには多量のデータが必要になるため、多くの心理物理学的研究ではこの検証にコストがかかる場合が多い [94]。この問題を解決するために、ベイズ統計学を用いて信号検出理論を拡張した新しい理論が登場した。

5.1.2 ベイズ統計学とベイズ推定による信号検出モデル

頻度主義統計学とベイズ統計学

ベイズ統計学は、ネイマン-ピアソン理論による統計学(頻度主義統計学とも呼ばれる、現在広く一般に使われている統計学)とは異なる仮定から作られた主観主義による統計学である [98]。両者で使われる幾つかの用語は共通しており、互いに補い合った関係であるが [98]、それぞれ確率の捉え方に特徴的な違いを持っている。ここでは、両者の考え方を以下の例を使って説明する。

9 個の白球と 1 個の黒球が入った A のツボと 2 個の白球と 8 個の黒球が入った B のツボがあり、そのどちらかが目の前に置かれている状況を想定する。ここで、ツボから一個球を取り出し

たら黒球であった。このとき、目の前のツボが A のツボか B のツボのどちらであるかを推定したい。

頻度主義統計学では、仮説検定と呼ばれる手続きで以下のように推定する。まず、「目の前のツボが A のツボである」という仮説(帰無仮説)を立てる。そして、仮説が正しいなら小さな確率 α (有意水準)でしか観測されないような現象を設定する。この場合には、この現象は「黒球を取り出す」現象であり、この現象が現れる確率は $\alpha = 10\%$ となる。次に、この現象が観測されたかどうかを確かめる。もしも観測されたなら、「仮説が正しいならきわめて小さい確率でしか観測されないはずの出来事が、実際に観測された」という理由から、「仮説は正しくないだろう」として仮説を捨てる(棄却する)。この例の場合は、実際に黒球を観測したので、「目の前のツボが A のツボである」という帰無仮説が棄却され、「目の前のツボが B のツボである」という結果になる。

言うまでもなくこの推定の方法にはリスクがある。仮説が正しく、目にしてるのがAのツボであっても、10%の確率で黒の球を観測する。したがって、この方法論で結論を下す限り、10回に1回の割合で「本当はツボAなのに、間違ってツボBと結論してしまう」のである。したがって、有意水準 α を小さく設定すればするだけ、このような間違の確率を小さくすることができる。

次にベイズ統計学を使って推定すると以下のようになる。

まず、事前確率というのを設定する。これは黒球を観測する前の段階として、ツボAなのかツボBなのかについてデータが無いなりに何らかの確率(主観確率)を割り振ることである。もちろん全く情報がないので五分五分と割り振る(この過程は「理由不十分の原理」と呼ばれる)。次に黒球を取り出したという情報をを利用して、ツボAなのかツボBなのかについて確率のアップデート(ベイズ更新)を行う。ベイズ更新された確率は事後確率と呼ばれる。黒球を取り出したという情報によって、「ツボAである」の事後確率は9分の1、「ツボBである」の事後確率は9分の8と計算される^{*2}。

以上のように、ベイズ統計学では、頻度主義統計学とは異なり「目の前のツボが B のツボである」といった結論を出さない。可能性の比例関係を出すだけで、最終的には判断は利用者に委ねられる。

頻度主義統計学は仮説の正しさについての確率を表現せず、データの確率的な変動によって仮説を評価できる。一方、ベイズ統計学は仮説の正しさについて確率的に表現し、推定の結果の判断は行わない。ベイズ統計学のこの柔軟性により、頻度主義統計学では計算できなかった事象や有意水準の恣意性といった問題を回避することができると考えられている。[98, 99, 100]

ベイズ推論による信号検出理論

以上の例で挙げられたように、ベイズ統計学では全てのパラメータを確率分布によって記述する必要がある。Lee[101]によればベイズ理論を用いた信号検出理論のモデルは以下のようになる。

^{*2} 事後確率の導出は付録 A 参照

まず hit の度数を h 、false alarm の度数を f 、hit と miss の度数の合計 (信号がある場合の試行数) を n_S 、false alarm と correct rejection の度数の合計 (信号がない場合の試行数) を n_N とする。ノイズ + 信号分布における hit の確率を θ^h とすると、hit の度数 h は成功確率 θ^h 、試行回数 n_S の二項分布に従っていると見なすことができる。同様に、ノイズ分布における false alarm の確率を θ^f とし、false alarm の度数 f は成功確率 θ^f 、試行回数 n_N の二項分布に従っていると仮定する。

hit の確率 θ^h と false alarm の確率 θ^f は、モデルの母数 d と c の関数として標準正規分布の累積分布関数 $\Phi(\cdot)$ を用いてそれぞれ以下のように表される。

$$\theta^h = \Phi\left(\frac{1}{2}d - c\right) = \Phi(d - k) \quad (5.3)$$

$$\theta^f = \Phi\left(-\frac{1}{2}d - c\right) = \Phi(-k) \quad (5.4)$$

また、 c と d の事前分布には以下のような正規分布を仮定する [96]。

$$d \sim \text{Normal}\left(0, \sqrt{2}\right) \quad (5.5)$$

$$c \sim \text{Normal}\left(0, \frac{1}{\sqrt{2}}\right) \quad (5.6)$$

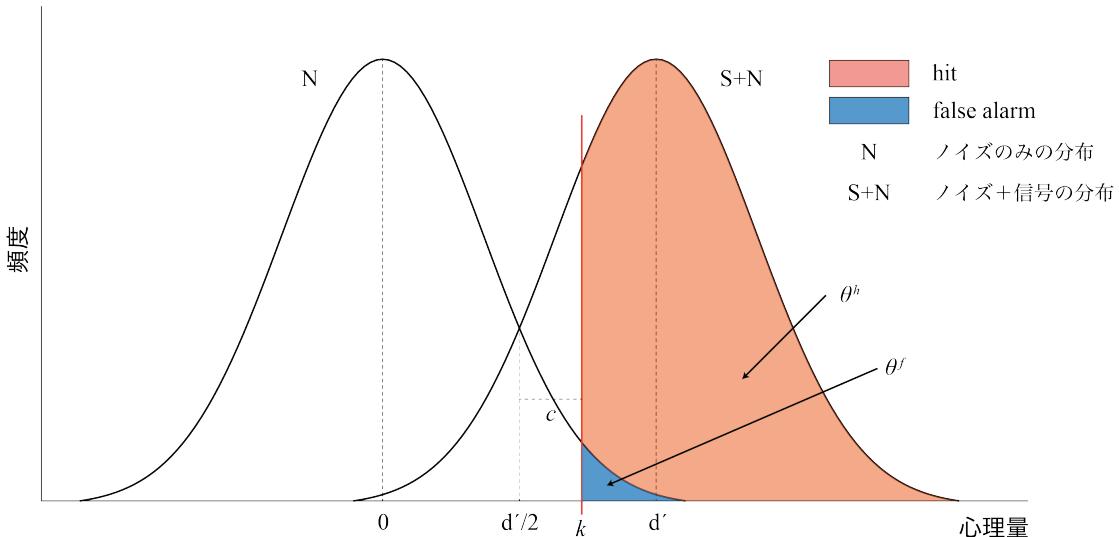


図 5.4 ベイズ理論による信号検出理論に必要なパラメータの図解 (Lee et al.[96] を改変)

これらをグラフィカルモデルに表すと図 5.5 のようになる。

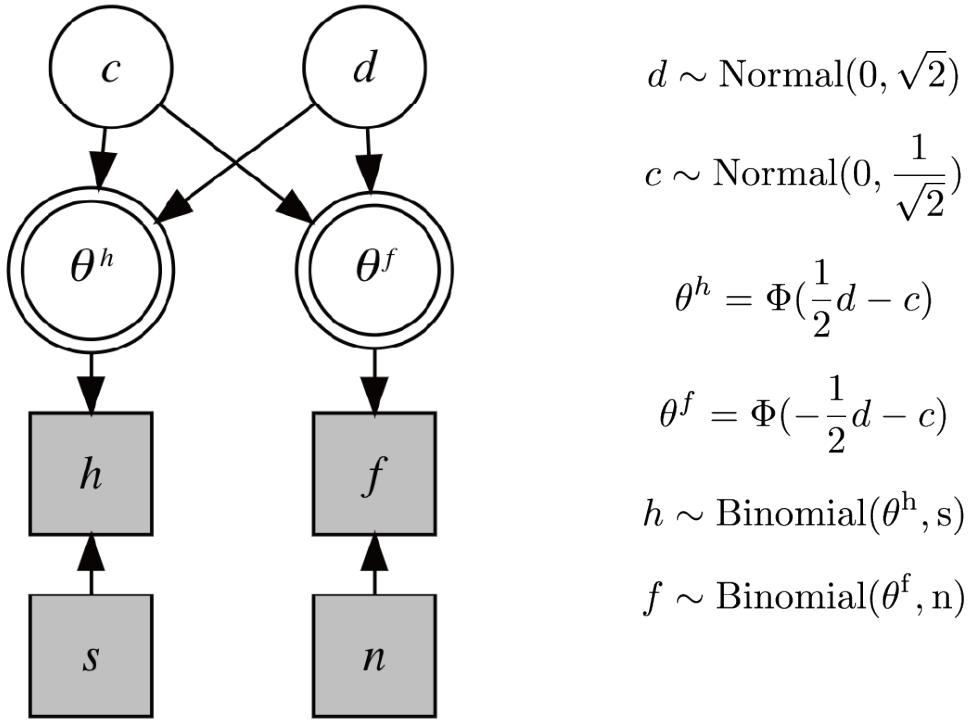


図 5.5 信号検出理論のグラフィカルモデル (豊田 [97]、Lee *et al.*[96] を改変)。円は連続値を表し、正方形は離散値を表す。また灰色は既知の値を示し、白色は未知の値を示す。二重の円は確定変数 (他の変数によって定義される変数) を意味する。

ベイズ推論による信号検出理論では、従来の信号検出理論で必要なデータ数よりも少ないデータ数で分析ができるという利点がある。従来の信号検出理論は、正規性の仮定を保証するために 300–500 ほどのデータを集める必要があるとされる [94]。しかしながら、ベイズ推論による信号検出理論では、事前分布で正規分布が仮定されているため、少ないデータ数でも結果の不確実性は高まるが計算を行うことが可能となる [101]。

また従来の信号検出理論では、例えば hit の度数が 100 で false alarm の度数が 0 の時のように、各パラメータの一部の割合が極端に高い場合ではエッジ補正という処理を行わなければ発散してしまい d' を計算することはできないが [102]、ベイズ推論による信号検出理論では、このようなケースであっても計算することが可能である [101]。

さらには階層ベイズモデルという手法を用いることで、実験参加者や項目の個人差もモデル化した高度な分析も可能となる。このような性質を持つため、ベイズ推論による信号検出理論は、4.4.2 で挙げられた判断バイアスの問題点を回避しつつ、従来の信号検出理論が持つ制約も克服で

きると考えられる。今回の実験では、このベイズ推論による信号検出理論を用いて弁別力 d' を求め、その値をもとに聴覚刺激の強度を変化させた時の時間統合窓の変動について調査した。

5.2 実験手法

5.2.1 実験参加者

東京芸術大学の学部生及び大学院生の 8 名 (男性 4 名女性 4 名、21 歳から 28 歳) が実験に參加した。実験時間はおおよそ 75 分であった。なおすべての実験参加者は自己申告により、日常生活に支障が出るほどの聴覚障害及び手指の触力覚感覚障害がないことを確認した。実験に使用した実験参加同意書を付録 D に示す。

5.2.2 実験装置

実験装置は 3.2.2 に記載されているものとおおよそ同じである (図 5.6 及び図 5.7)。しかしながら、実験実施場所の都合によりオーディオインターフェイスを変更した。また、実験を行なった部屋は、簡易防音室から静かな教室に変更となっており、実験中は部屋の電気が消された状態だった。さらに実験参加者は、呈示される聴覚刺激と触力覚刺激に集中するため使い捨てのアイマスクを着用した。

5.2.3 実験刺激

聴覚刺激は 1 kHz の正弦波をヘッドフォンに呈示し、触力覚刺激は 250 Hz の正弦波を振動子に呈示した。聴覚刺激の強度は 2 種類用意され、1 kHz の実験刺激を再生したとき疑似頭の聴取位置においてそれぞれ 80 dB_{SPL}, 68 dB_{SPL} になるよう調整された。聴覚刺激と触力覚刺激は実験コンピュータ時間における同時または非同時に開始した。刺激開始時間差の値は $\Delta t = \pm 0, 30, 45, 60, 75, 150$ ms の 11 種であった。刺激音の持続時間は 300 ms であった。刺激音の開始と終了には 15 ms のコサインテーパー処理がなされ、ゆるやかな立ち上がりあるいは立ち下がりとなった。こうした刺激作成上の工夫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の雑音が発生するのを抑えるために行われた [91]。

5.2.4 手続き

実験参加者は実験デバイスを水平方向に両手で把持し、親指で実験デバイス上に表示された GUI を操作することにより回答を行った。各試行では、2 つの正弦波で構成されるマルチモーダル刺激が前述の遅延を含んだ状態で呈示され、その後、実験参加者はその刺激の組み合わせが同時に発せられたか感じたかそうでなかつたかを回答した。刺激が呈示された後、実験参加者は 5 秒以



図 5.6 実験風景

内に回答を行う必要があった。刺激が呈示された 5 秒後、もう一度同条件の刺激が呈示され 5 秒間の回答時間が再度与えられた。2 度目の刺激呈示でも回答が行えなかった場合は、その施行は無回答として記録された。実験参加者は回答後、その回答についていかなるフィードバックも与えられなかった。11 種の Δt における各聴覚強度条件は、それぞれ 23 回再生され、すべての再生順序はランダムであった。実験参加者は 110 試行に一度、最低 1 分以上の休憩をおこなった。また実験を行う前に、実験参加者はランダムに選ばれた 10 条件でフィードバックなしの練習を行った。実験終了後、実験参加者に対し 3 分ほどのアンケートを行い、年齢、性別、FLANDERS 利き手テスト [82] による利き手の調査を行った。

5.3 結果

本実験では以下の手順によって時間統合窓を求めた。

1. 各条件の試行数、hit 及び false alarm の数を求める。
2. 求められた値から上記モデルを用いて各聴覚強度条件における Δt 条件の d' を求める。
3. 横軸を Δt 、縦軸を d' として実験参加者ごとにプロットする。

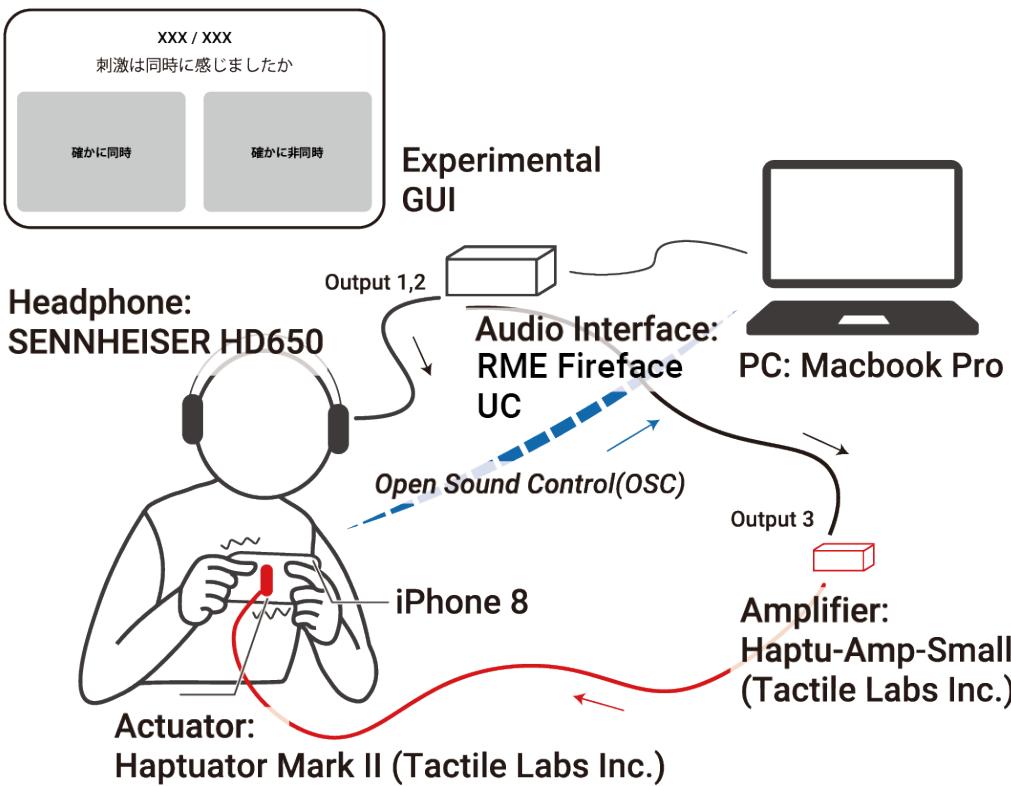


図 5.7 実験装置模式図及び実験 GUI

4. 得られた離散分布から平滑化スプラインを用いて回帰曲線を求める。
5. 得られた回帰曲線の平均を取り、各振動強度条件における平均 d' の推移を求める。
6. 得られた曲線が $d' = 1$ と交わる二点を求め、その距離を時間統合窓とする。

本実験では、全 4,048 試行データ (聴覚強度 2 種 \times Δt 11 種 \times 23 回 \times 8 人) が集められた。回答時間超過による無回答により省かれたデータはなかった。

各聴覚強度条件における $\Delta t = 0 \text{ ms}$ を信号検出理論におけるノイズ条件、 $\Delta t = \pm 30, 45, 60, 75, 150 \text{ ms}$ をノイズ + 信号条件とし、各条件の試行数、hit 及び false alarm の度数を求めた。その後、この値をもとに図 5.5 のモデルを用いて弁別力 d' を導出した。ベイズ推定には統計的推論のための確率的プログラミング言語 Stan 及び R 言語を用いた^{*3}。4 つのマルコフ連鎖それぞれにおいて、事後分布から 20000 回のサンプリングを行い、最初の 2000 回をウォームアップ期間として破棄し、合計 18000 個の母数の標本を用いて計算した。

^{*3} 本実験の分析に使用したコードは、Lee et al.[96] で記載されている JAGS コードを Martin Smira らによって Rstan 用に書き換えたものを使用した。コードは https://github.com/stan-dev/example-models/tree/master/Bayesian_Cognitive_Modeling から手に入れることができる (2020 年 12 月 3 日時点)。

ここでは例としてある一人の実験参加者の結果を紹介する^{*4}。図 5.8 は各聴覚強度条件における Δt 条件の d' をそれぞれプロットしたものである。

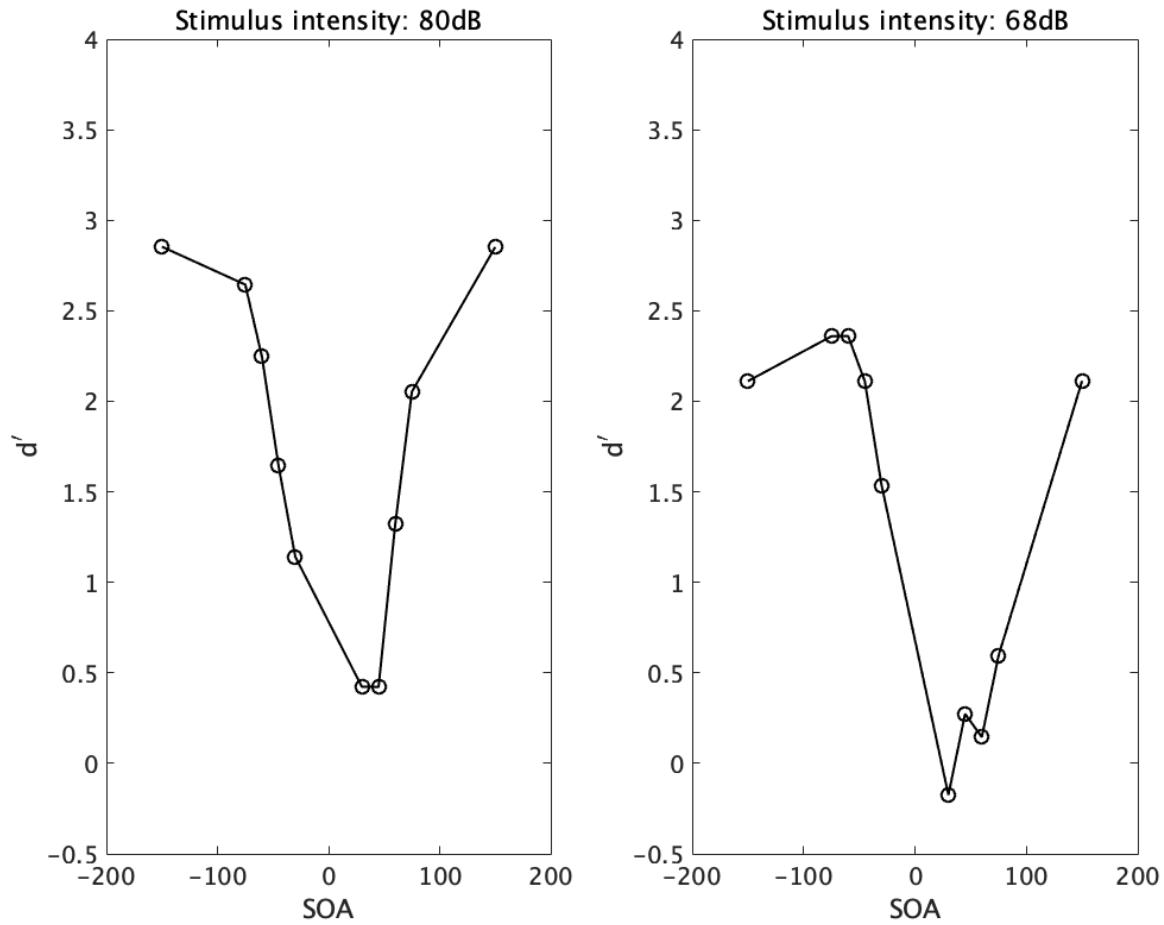


図 5.8 ある実験参加者の d' の推移

ここでは横軸は Δt 、縦軸は d' を表している。どの聴覚強度条件においても $\Delta t = 0 \text{ ms}$ 付近を頂点谷型の傾向を見せている。

次に、各実験参加者の離散分布から平滑化スプライン [103] を用い回帰曲線を求めた後、全実験参加者の回帰曲線を平均して各聴覚強度条件における平均 d' の推移を求めた。各曲線の残差平方和と決定係数は付録 C に表す。その後、得られた曲線が $d' = 1$ と交わる二点を求め、その幅を時間統合窓とした。その結果を図 5.9 に表す。

図 5.9 の赤色の太い曲線が平均 d' の推移を表し、縦軸の $d' = 1$ と赤色の太い曲線が交わる点が時間統合窓である。聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$ の場合の時間統合窓は 118 ms で、 $68 \text{ dB}_{\text{SPL}}$ の場合の時間統合窓は 126 ms であった。

^{*4} 全参加者の結果は付録 B に記載した。

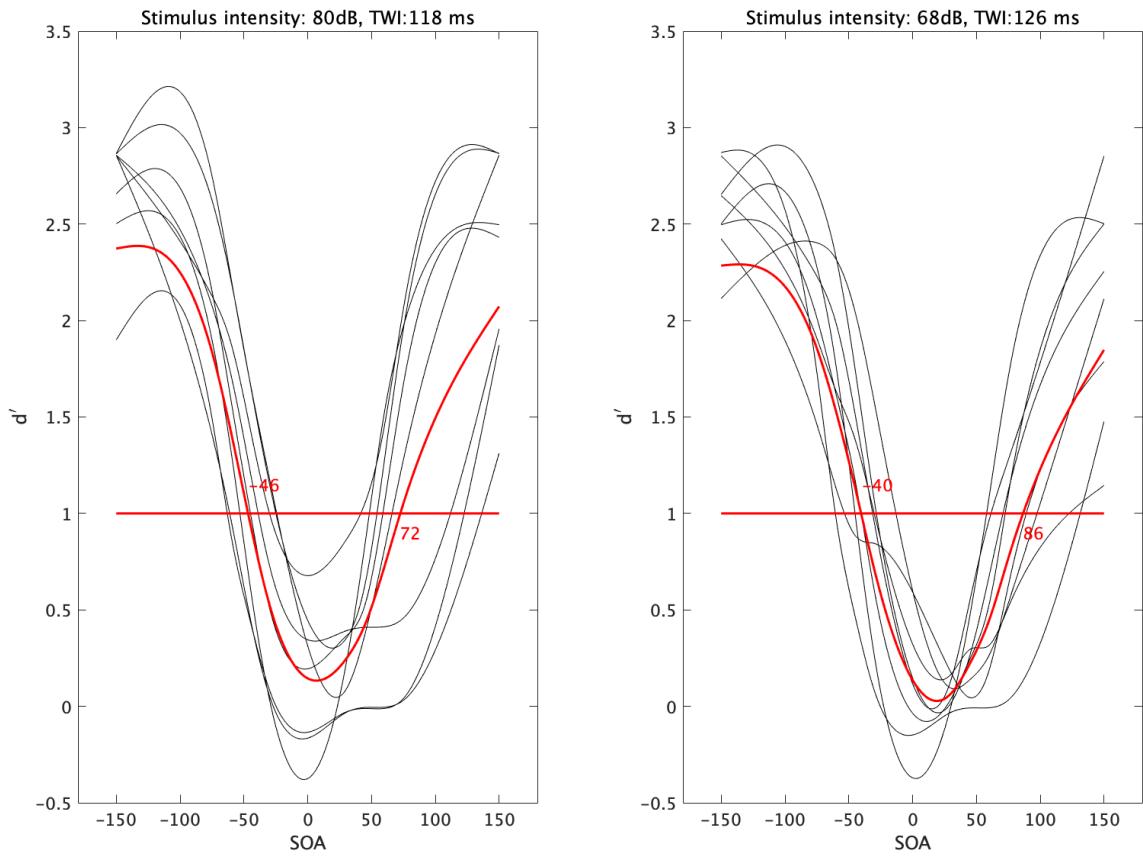


図 5.9 各聴覚強度条件における平均 d' の推移と時間統合窓

5.4 考察

5.4.1 同時性知覚実験における個人差

本実験では、ベイズ推論による信号検出理論を用いて弁別力 d' を求め、その値をもとに聴覚刺激の強度を変化させた時の時間統合窓の変動について調査した。その結果、聴覚強度条件が 80 dB_{SPL} の場合と 68 dB_{SPL} の場合では、80 dB_{SPL} の方が 8 ms 時間統合窓が狭くなるという結果が得られた。ここでは、得られたデータから個人の振る舞いに注目し、同時性知覚実験における個人差の傾向について述べる。

図 5.9 の二つのパネルを比較すると、68 dB_{SPL} の場合よりも 80 dB_{SPL} の時の方が各実験参加者間のばらつきが大きく、実験参加者のうち何名かは、刺激開始時間差が +100 ms を超えるまで d' がおおよそ 0–0.5 あたりで止まる振る舞いを見せていました。ここから、実験参加者内で特に聴覚強度条件が 80 dB_{SPL} の場合の時に同時性知覚の振る舞いにおいて 2 種類の判断の違いが生じていた可能性が考えられる。

図 5.10 は実験参加者の聴覚強度条件が 80 dB_{SPL} の場合の時間統合窓を、昇順に並び替えて表

したものである。

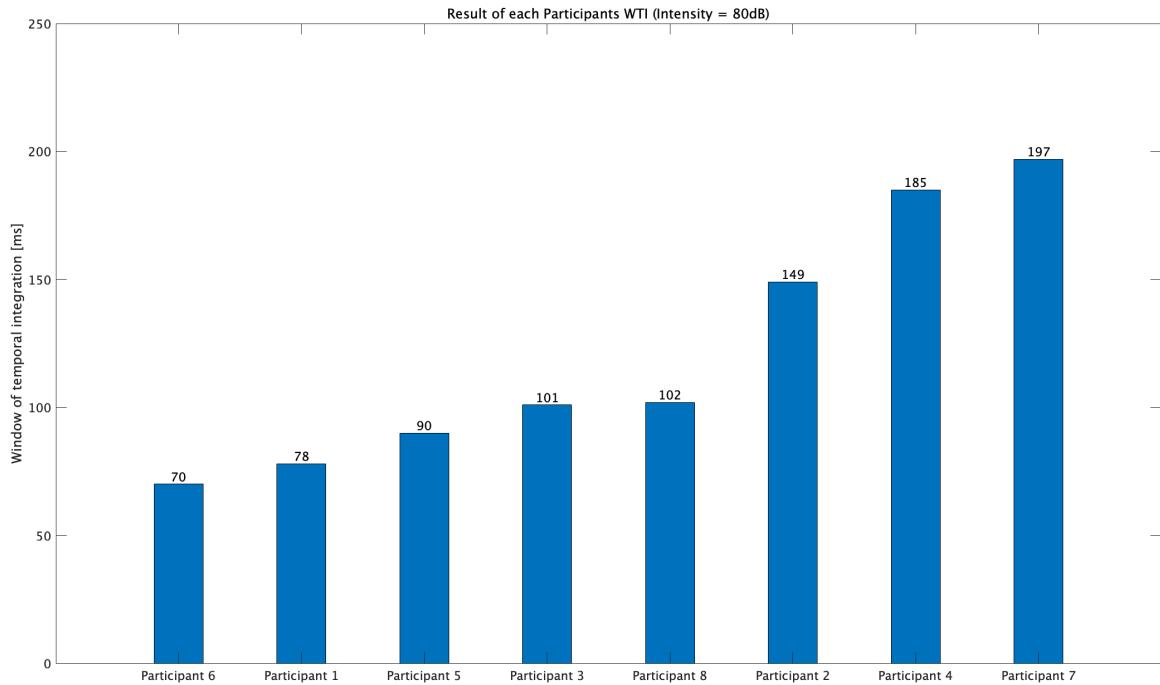


図 5.10 聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$ の時の実験参加者それぞれの時間統合窓。横軸は実験参加者で、縦軸は時間統合窓(単位: ms)である。図中の棒上部に記載されている数字は、実験参加者それぞれの時間統合窓の値である。

図 5.10 を見ると、本条件では実験者番号 2, 実験者番号 4, 実験者番号 7(実験者番号に関して、以降 Participant 2, Participant 4, Participant 7 のように表す)の時間統合窓はおおよそ $150 \text{ ms} - 200 \text{ ms}$ の間にあり、それ以外の実験参加者の時間統合窓よりも比較的大きいことが分かる。

そこで、Participant 2, Participant 4, Participant 7 の時間統合窓(以降、Group1 とする)とそれ以外の時間統合窓(以降、Group2 とする)それぞれの群の平均値に有意差があるのか有意水準 5% で Welch の t 検定を行ったところ、上記 2 群の平均点の差は有意であることがわかった ($t(2.7844) = -5.6429, p = 0.01342$)。

この結果をもとに、上記 2 群ごとに聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$ の場合の実験結果を分けた場合の実験結果を図 5.11 に表す。

図 5.11 から見て取れるように、Group1 は谷の広さが Group2 と比べ広く、谷が刺激開始時間差上でプラス方向に広がっている傾向を見せる群であったことが分かった。これは、聴覚刺激が触力覚刺激より後に呈示されていた場合、そうではない場合と比べて両者にズレがあると知覚するのが遅くなるという傾向が Group1 の群にあったと言える。Group2 や先行研究 [86][32] のように、同時性知覚の実験の際に刺激開始時間差上でプラスの方向に谷の頂点がシフトする現象は以前から報告されている。しかしながら、実験参加者によっては刺激開始時間差上でプラスの時

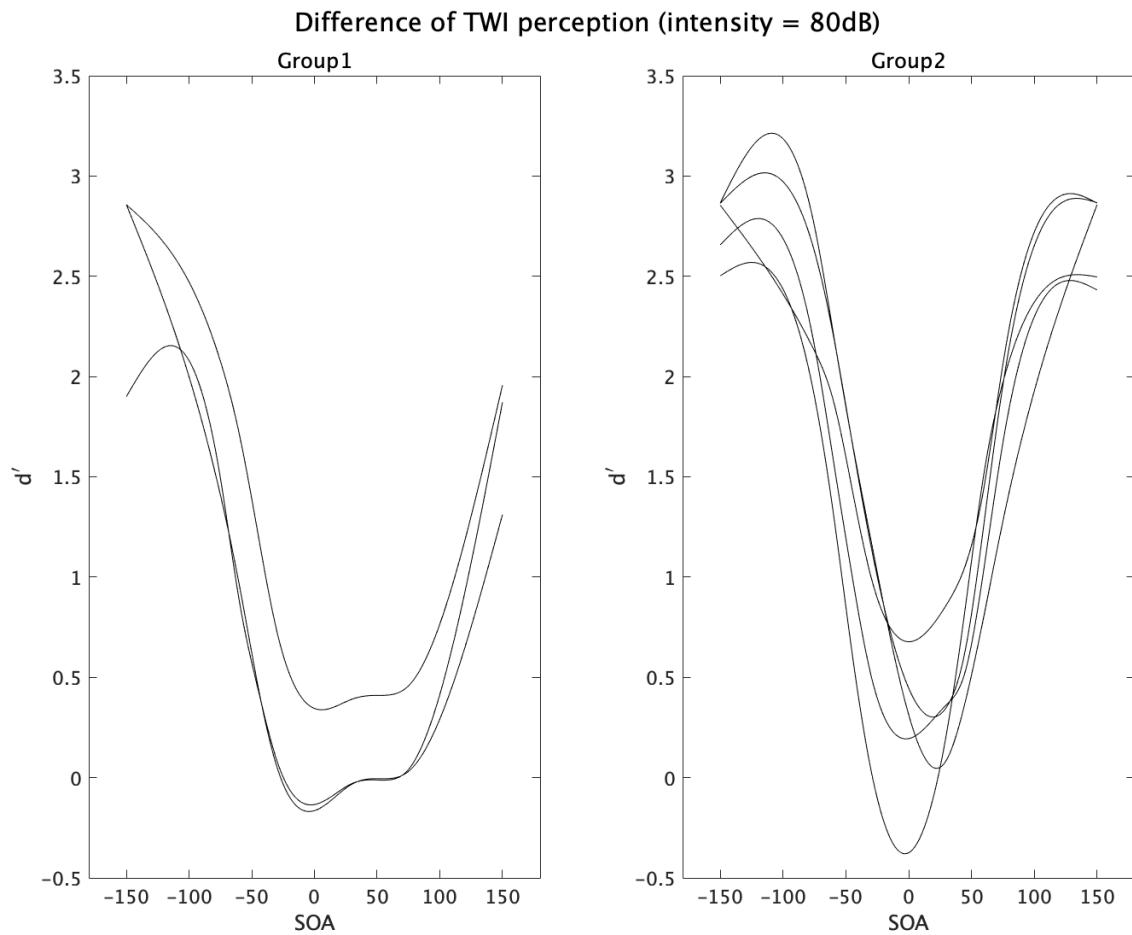


図 5.11 聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$ の場合を 2 群に分けた時の実験結果

の同時性知覚がマイナスと時と比べて鈍くなるという現象が見られたのは注目に値する。

本実験においては、実験後に収集したアンケートなどからこの分類を生じさせる要因を判明させることはできなかったが、この違いを特定するために時間知覚を問わない心理測定法（例えば同時タッピング課題 [104] や通過・反発錯覚を用いた課題 [46]）や生理学的アプローチを用いた測定方法 [105]などを実施し、人の同時性知覚の処理プロセスにおいて、どの段階でこの個人差が生じているのかを今後検討したい。

5.4.2 聴覚強度条件を変化させた時の時間統合窓の違い

本実験では、聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$ の場合の時間統合窓は 118 ms で $68 \text{ dB}_{\text{SPL}}$ の場合の時間統合窓は 126 ms と、 $80 \text{ dB}_{\text{SPL}}$ の方が時間統合窓が狭いことがわかった。

5.4.1 で述べたように、5.3 の結果は判断の違いによって 2 群が混在しているデータであるため、刺激開始時間差がプラスであってもマイナスであっても左右対象に判断が行えた群 (Group2) のみを抽出し、同じように時間統合窓の計算を行った。

図 5.12 は Group2 での時間統合窓の違いを表している。Group2 では聴覚強度条件が $80 \text{ dB}_{\text{SPL}}$

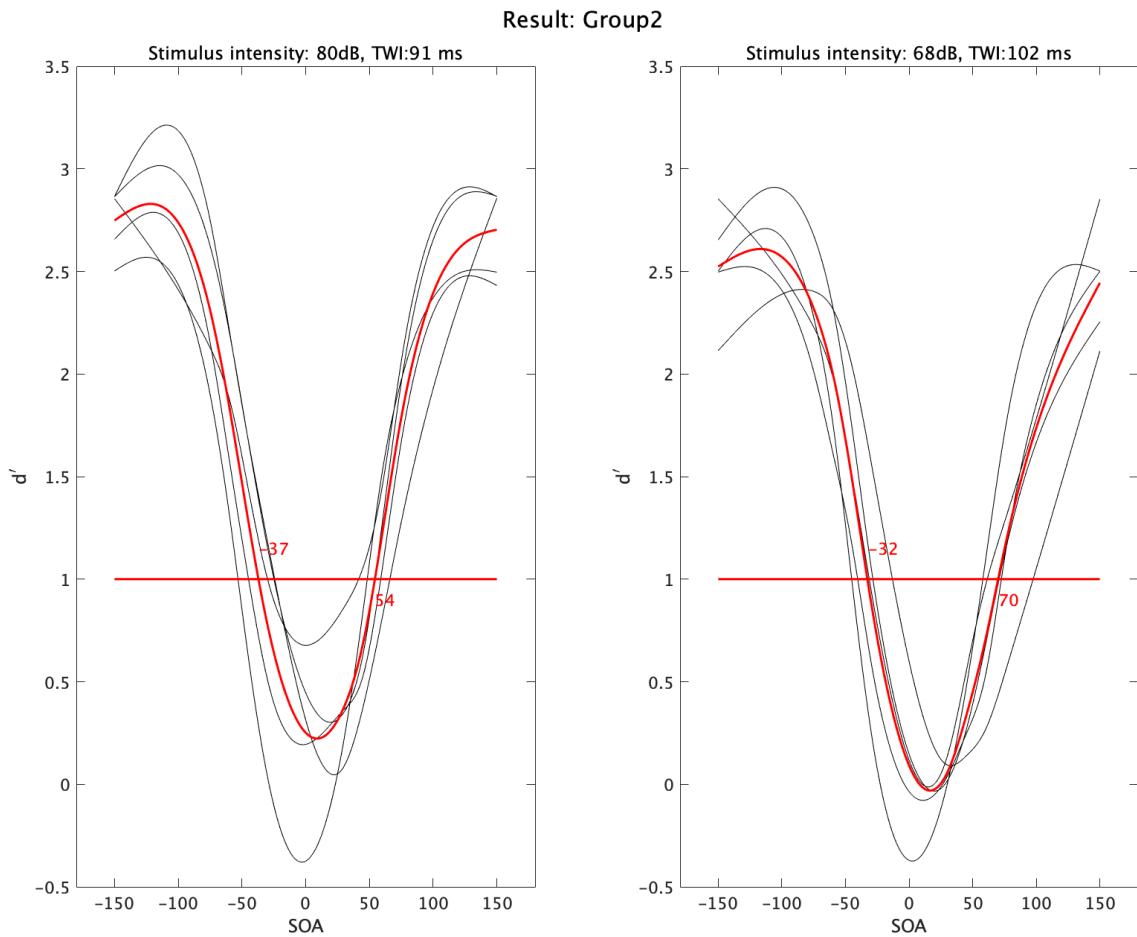


図 5.12 Group2 を抽出した時の実験結果

の場合の時間統合窓は 91 ms で、68 dB_{SPL} の場合の時間統合窓は 102 ms であった。そこで、Group2において聴覚強度条件により時間統合窓に違いが生じたのかを調べるために、各聴覚強度条件における Group2 での実験参加者の時間統合窓の平均を Welch の t 検定を用いて検定を行った結果、80 dB_{SPL} と 68 dB_{SPL} の時間統合窓の差は有意傾向であることがわかった。 $(t(6.484) = -2.0245, p = 0.086)$ 。

t 検定の結果や図 5.12 から見て取れるように、聴覚強度条件が 80 dB_{SPL} と 68 dB_{SPL} の間には同時性知覚において違いが生じていると考えられ、聴覚刺激の強度が小さい場合には両刺激のずれが分かりにくくなる傾向が存在する可能性がある。本実験では聴覚強度条件で時間統合窓に違いが生じるかどうかを判断することが目的であったため、両者の違いの要因などより詳細な振る舞いを検討することはできなかった。今後は聴覚刺激の強度の変化の度合いを増やし、この条件において同時性知覚がどのように変化するのかを検討したい。

5.5 結論

本実験では、モバイルデバイスを手で把持する状況を想定し、聴覚刺激の強度を変化させた時の時間統合窓の変動をペイズ推論による信号検出理論を用いて調査した。その結果、聴覚強度条件が 80 dB_{SPL} の場合の時間統合窓は 118 ms で、68 dB_{SPL} の場合の時間統合窓は 126 ms であり、聴覚強度条件の違いにより時間統合窓の振る舞いに差異が生じることがわかった。また聴覚-触力覚における同時性知覚判断において、聴覚刺激が先に呈示されていても後に呈示されていてもおよそ同じように両刺激のいずれの判別が行うことが可能な実験参加者グループと、聴覚刺激が触力覚刺激より後に呈示されていた場合、そうでない場合と比べ両者にずれがあると知覚するのが遅くなるという傾向を持つ実験参加者グループが存在することがわかった。今後は、聴覚強度条件を細かく設定し時間統合窓の振る舞いについてより詳細に検討しつつ、実験参加者グループの判断の違いが同時性知覚のどの段階で生じるのかを、心理物理学的手法だけでなく生理学的手法、身体性認知科学的手法を用いて検討したい。

第 6 章

考察

本章では、第 3 章から第 5 章までの結果を踏まえた全体的な考察を行った後、各実験で用いた実験手法を踏まえた発展的な考察を行う。

6.1 刺激の特徴量と時間統合窓の変動

6.1.1 時間統合窓の変動傾向について

本論文では、聴覚刺激の周波数とノイズ量（第 3 章）、両刺激の周波数間隔（第 4 章）、そして聴覚刺激の強度（第 5 章）をそれぞれ変化させた時の時間統合窓の変化を調査することで、時間統合窓の振る舞いを決定付ける要因の一端を解明することを目的とするものであった。それぞれの実験で得られた時間統合窓の大きさとその振る舞いの変化を図 6.1 に表す。

聴覚刺激の周波数を 6 段階で変化させた場合、周波数が高くなるに伴って時間統合窓が大きくなる傾向が見られた（図 6.1 の左上）。しかしながら、第 4 章で行った両刺激の周波数間隔を 12 段階変化させた実験（図 6.1 の左下）では、聴覚刺激の周波数を変化させているという条件は同じにもかかわらず、その傾向が相反するものであった。この理由として、4.4.2 で述べたように判断バイアスの含まれ方の違いが、この差異を生み出したと考えられる。この判断バイアスによる影響は 6.2 にて詳しく検討を行う。

また、聴覚刺激のノイズ量を 6 段階で変化させた場合では、時間統合窓に関して特異な傾向を見出すことができなかった（図 6.1 の右上）。その理由として、ノイズ量の変化で生じる物理的特徴量の差異ではない要因が刺激作成時の不手際により発生していた点が考えられる。第 3 章の実験では実験刺激の長さが 50 ms であり、その短さ故に目的周波数以外の雑音が発生するのを抑えるコサインテーパー処理 [91] が行われなかった。そのため、刺激が呈示された直後に全帯域を含むパルス上のノイズが瞬時に発生し、これが時間統合窓の判断に何らかの影響を及ぼしていたため、特異な傾向が見出せなかったと考えられる。

聴覚刺激の強度を 2 段階で変化させた場合では、聴覚刺激の強度が弱い時に時間統合窓が小さ

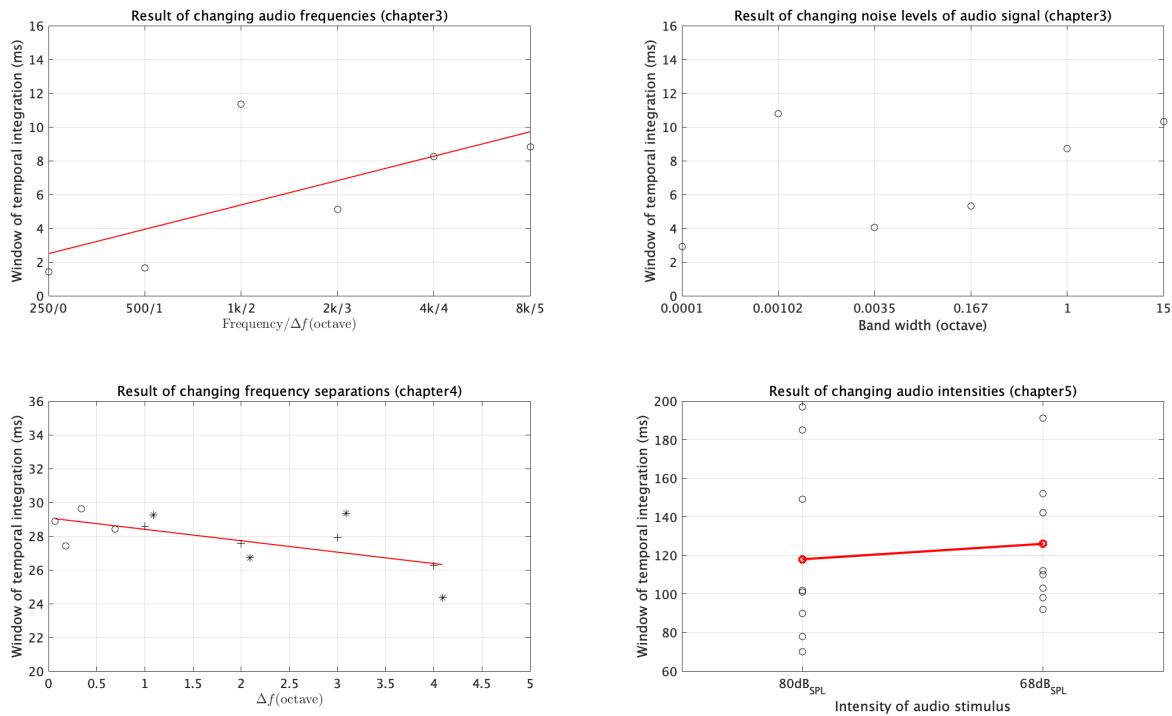


図 6.1 本論文で行った 3 つの実験の結果。上部 2 つのパネルは第 3 章で行った実験結果を表し、左下のパネルは第 4 章の実験結果、右下のパネルは第 5 章の実験結果を表している。

くなる傾向が見られた(図 6.1 の右下)。またこの実験では、実験参加者ごとに 2 種類の判断の違いが見られ、刺激の呈示順序にかかわらず両刺激のいずれ判断が同じであった群(Group2)のみを抽出した場合の時間統合窓は、80 dB_{SPL} の場合は 91 ms で、68 dB_{SPL} では 102 ms であった(5.4.1 参照)。この場合でも、前述のように聴覚刺激の強度が弱い時、時間統合窓が小さくなる傾向を持つことが分かった。

これら結果をまとめると、

1. 聴覚刺激の周波数を変化させた場合、2 つの実験の結果は相反するものであったが、時間統合窓の変動に何らかの影響を与える。
2. 聴覚刺激のノイズ量と時間統合窓の変動の関係は、本実験においては関係性を見出すことができなかった。
3. 聴覚刺激の強度の減少は、時間統合窓が小さくなる要因となりうる。

ということが分かった。

聴覚刺激の周波数に関しては、判断バイアスの含まれ方の違いの他にも、刺激の長さの違いや実験環境の違いなど、他にも時間統合窓に影響を及ぼしうる要因が存在する。今後の実験においては、実験手法以外の全ての条件を統制した上で検討を行うことにより、より詳細な関係を考察したい。

また、聴覚刺激のノイズ量に関しては、第 3 章での実験刺激の作成方法においていくつかの問

題が生じていたため、時間統合窓との正確な関係を記述することができなかった。ノイズ量の変化は 3.1 で述べたように、現在の聴覚-触力覚マルチモーダル表現において頻繁に用いられる要素であるため、今後の実験においては、実際のコンテンツで用いられている聴覚-触力覚刺激をより分析した上でさらに研究を行っていきたい。

聴覚刺激の強度に関して、第 5 章での実験手続き上、実験時間が実験参加者にとって負担になりうるため、強度の段階を 2 種類以上増やして検討することができなかった。今後は実験に用いたパラメータをより効率化することで、聴覚強度条件の段階をより増やして検討を行いたい。

6.1.2 各実験で得られた時間統合窓の大きさの違いについて

図 6.1 の各パネルの縦軸を見ると、各実験によって時間統合窓の大きさに違いがあることが分かる。第 3 章の実験(図 6.1 上部)では、時間統合窓の範囲はおおよそ 1 ms から 12 ms、第 4 章の実験(図 6.1 左下)では 24 ms から 30 ms の範囲、第 5 章の実験ではおおよそ 90 ms から 130 ms の範囲と、大きなばらつきがあることが分かる。

本来であれば、同じ事象を実験の対象とした場合、たとえ実験手法が異なっていても多くの場合似たような結果になるはずである [16] が、本研究ではそのばらつきが多分に顕著であったと見られる。

時間統合窓の大きさに関する研究について、1961 年に発表された Hirsh and Sherrick[68] による実験では、視覚-聴覚・視覚-触力覚・聴覚-触力覚での刺激を様々な刺激開始時間差で呈示し、時間順序判断課題によって時間統合窓が求められた。その結果、時間統合窓の大きさは 20 ms であり、感覚モダリティの組み合わせが異なっていても一定であったと報告された。

しかしながら、その後の研究により、時間統合窓の大きさは決まった幅ではなく、様々な要因で変化し個人差も大きいことが判明した。Occelli *et al.*[35] は、Hirsh and Sherrick 以降、2.2.1 で取り上げた Fujisaki and Nishida までに行われた、聴覚-触力覚間での時間統合窓の大きさに関する研究をまとめている(表 6.1)。

表 6.1 を見ると、実験刺激や手法が異なっていても、時間統合窓がおおよそ 20 ms から 80 ms の範囲にあり、この結果から考えると、第 4 章の実験はおおよそ先行研究に似た結果が得られ、第 3 章及び第 5 章では、先行研究の結果にそぐわない結果であったと言える。確かに第 3 章での実験は、実験手法の違いの他にも実験刺激の長さが 50 ms と短かったため、実験参加者が正確な同時性を判断を行えず、先行研究とは逸脱した結果が出たと考えられる。

一方第 5 章の実験に関しては、聴覚-触力覚間ではなく音声と発話画像を用いた時の時間統合窓に関する研究において、実験刺激が単純であった時と比べてより広くなることが報告されており、発話刺激を用いた時の時間統合窓は 389 ms、単純な刺激の時は 263 ms であった [106] とする研究もあり、一概に第 5 章の実験結果が先行研究の傾向から逸脱しているものということはできない。時間統合窓の大きさは、刺激の物理的な特性、特に刺激の立ち上がりの鋭さに強く依存し、刺激

表 6.1 聴覚-触力覚間の時間統合窓の大きさに関する諸研究 (Occelli *et al.*[35] より引用)

Author(s)	Stimuli and Design	Results
Hirsh and Sherrick (1961; Exp. 4)	Three sessions, one for each stimulus pairing (i.e., audiovisual, visuotactile, and audiotactile)	JND of around 20 ms for all stimulus combinations
Spence <i>et al.</i> (2003)	Exp. 1: Visuotactile pairs of stimuli, presented either from the same or from different positions in frontal space, uncrossed or crossed posture Exp. 2: Audiovisual pairs of stimuli, presented either from the same or from different positions in frontal space	Exp 1: Uncrossed, JND: 65 ms (Same) vs. 42 ms (Different); $p < .0005$ Crossed, JND: 50 ms (Same) vs. 58 ms (Different); <i>n.s.</i> Exp 2: JND: 53 ms (Same) vs. 42 (Different); $p < .05$ JND: 64 ms (Same) vs. 55 ms (Different); $p < .05$
Kitagawa <i>et al.</i> (2005; Exp. 1)	Audiotactile pairs of stimuli, presented either from the same or from different positions in rear space	Exp 1: JND: 75 ms (Same) vs. 82 ms (Different); <i>n.s.</i>
Zampini <i>et al.</i> (2005)	Audiotactile pairs of stimuli, presented either from the same or from different positions in frontal space	Exp 2: JND: 45 ms (Same) vs. 42 ms (Different); <i>n.s.</i> Exp 3: JND: 44 ms (Same) vs. 44 ms (Different); <i>n.s.</i>
Occelli <i>et al.</i> (2008)	Audiotactile pairs of stimuli, presented either from the same or from different positions in frontal space	Group X Relative Spatial Position, $p = .03$ Sighted, JND: 69 ms (Same) vs. 70 ms (Different); <i>n.s.</i> Blind, JND: 73 ms (Same) vs. 61 ms (Different); $p = .005$
Fujisaki and Nishida (2009; Exp. 4)	Four sessions (i.e., audiovisual, visuotactile, audiotactile, tactile)	JND: 36 ms (audiovisual), 29 ms (visuotactile), 25 ms (audiotactile), 17 ms (tactile); TT vs. AV: $p = 0.02$; TT vs. VT: $p = 0.02$, no other significant values

の強度が弱いなどといった刺激の開始が分かりにくく、刺激が呈示された場合は時間統合窓が小さくなるという報告 [7] にもあるように、第 5 章の実験は、何らかの要因により他の実験よりも刺激の開始が分かりにくかったため、時間統合窓の大きさがより多くなった可能性も考えられる。

また本研究の目標は、時間統合窓の大きさだけを解明するものではない。聴覚-触力覚刺激を使ったデジタルコンテンツ制作時の設計指標を作るために必要な、音と振動の同時/非同時感覚や非同時感覚から生じる違和感と、デジタルコンテンツ体験の質の関係性について考察することにある。本研究においては、刺激と時間統合窓の大きさだけに注視していたが、今後の研究では実験手法と時間統合窓の関係性だけに着目するのではなく、コンテンツ体験の質と時間統合窓の関係性について実験を行うなど、巨視的な捉え方で聴覚-触力覚の時間統合窓及び同時性知覚について考えていくべき。

6.2 新たな実験手法についての検討

6.2.1 心理物理学的実験手法と判断バイアス

4.4.2 にて、同時性判断課題による実験では判断基準のゆらぎが直接実験結果に反映されるため、知覚実験という観点からは判断バイアスが生じやすい方法と考えられている [16][4] と述べた。

これを克服するために、第5章では、判断バイアスに左右されない弁別力を求めることができる信号検出理論を用いて実験を行った。ここで述べられている判断バイアスはどのようなものであるのか。

判断バイアスとは、おもに実験手続き上の理由で判断に偏りが出てしまうことを指す[7]。第4章の実験を例にして考えてみる。

実験参加者はある刺激開始時間差を持った音と振動の刺激が呈示され、「音と振動が同時に呈示されたか」を回答した。この時、実験参加者が刺激開始時間差を正確に知覚できていれば、時間差が0または0に近い時に同時と回答し、それ以外の時は非同時であると回答するはずである。しかしながら、実験参加者が同時か非同時かが分からなかった場合はどうであろうか。この実験手法では、実験参加者は制限時間内に同時か非同時を分からなかった場合にもどちらかを選択しなければならない。そのため、同時か非同時かどうか曖昧だった場合でも「非同時」と判断しがちな実験参加者や、はっきり非同時であると確信したときだけ、「非同時」と答える実験参加者など、判断基準にばらつきが出る。この実験参加者による判断基準の違いを判断バイアスと呼ぶ。同時性判断課題においては二者択一が求められているため、実験担当者が判断基準の違いを実験後に知ることができず、判断バイアスを考慮できない状態で分析が行われてしまう。もしこの判断バイアスが実験結果に大きな影響を及ぼしていた場合、たとえ同じ実験を別日に行ったとしても、同じ結果が確かに得られるとは言えなくなる。

信号検出理論はこの判断バイアスの有無や大小を、同時性判断課題と同じ実験手続きで知ることができる。信号検出理論では、実験参加者の判断基準の違いを調べるために、わざと刺激の入っていない試行(ノイズ分布)を被験者に呈示し、刺激がないのにも関わらず「刺激がある」と応答する確率を求める。これをfalse alarmと呼ぶ(p.42の図5.1)。第4章の実験に当てはめると、刺激の入っていない試行とは刺激開始時間差が0である試行であり、これを呈示した時に「(音と振動は)同時ではない」と答えたfalse alarmの確率を求めることになる。同時か非同時かどうか曖昧な場合でも「非同時」と判断しがちな実験参加者はfalse alarmが大きく、はっきり非同時であると確信したときだけ「非同時」と答える実験参加者はfalse alarmが小さくなる。このデータから統計的処理を行うことで、被験者の判断基準の違いを除去することができる[107]。

この方法によって判断バイアスを適切に除去ができたかに思われるが、それでも実験担当者がコントロールできない判断バイアスが内在している。例えばノイズ分布でのfalse alarmの確率を求める際の判断である。false alarmの確率を求める際の実験手続きは二者択一であり、ここでの判断バイアスを確かめる術がない。そこで、実験参加者の回答の確信度を計測することにより、さらに実験参加者の判断基準の様子を検討することができる、第二種信号検出モデルという実験手法がある[108][109]。

これは、信号検出理論による刺激がある/ないの判断(第二種信号検出モデルにおいては第一種課題と呼ぶ)の後に、その判断の自信評定を二値(自信がある/自信がない)又は主観的確率(0%から100%)などで答えるものである(第二種課題)。すべての回答データを収集した後、第一種課

題での結果の正答率と第二種課題の回答の組み合わせで、p.42 の図 5.1 に似た表を作る（表 6.2）。

表 6.2 第二種課題の反応テーブル

	自信あり	自信なし
正答	Type II Hit	Type II Miss
誤答	Type II False Alarm	Type II Correct Rejection

なお、第一種課題の正答とは第一種課題での Hit と Correct Rejection のことを指し、第一種課題の誤答は第一種課題での False Alarm のことを指す。第二種課題の反応テーブルの各項目の意味は以下の通りである。

1. Type II Hit: 第一種課題で正答して、自信があった試行。
2. Type II Miss: 第一種課題で正答したが、自信がなかった試行。
3. Type II False Alarm: 第一種課題で誤答したが、自信があった試行。
4. Type II Correct Rejection: 第一種課題で誤答して、自信もなかった試行。

この分類をもとに通常の信号検出理論と同じ統計処理を行うことで [109]、実験参加者の確信度判断の弁別力を求めることができる。第二種信号検出モデルでも正規性の仮定が満たされない場合があるので [108]、5.1.2 で述べたベイズ統計学を用いてこの仮定を満たした Meta- d' という指標が提案されている [110]。

この確信度判断の弁別力を用いて、第 5 章で求めた d' に重み付けなどを行うことにより、より実験参加者の判断バイアスをコントロールした時間統合窓の計算を行うことができるかもしれない。しかしながら、第二種信号検出モデルの実験手続き上、それを行わない場合と比べ実験時間が倍になるため、実験の規模やコストを鑑みながら第二種信号検出モデルを使用した時の有用性を考える必要がある。

6.2.2 時間統合窓の大きさを求める心理物理学手法以外の手法

これまで、心理物理学の立場から時間統合窓の振る舞いについて論じてきた。しかしながら、2.1.1 や 2.1.2 でも触れたように、主観評価実験ではなく脳の神経活動の様子を観察することで、マルチモーダル刺激の同時知覚や時間統合窓の振る舞いについて検討することもできる。

我々にはそれぞれの感覚モダリティに対応した感覚器があるが、時間そのものを計測する感覚器や脳内機構は存在しない [16]。それにも関わらず我々は時間に関わる様々な認知や反応を起こす。このためには、時間経過を表象するためのタイムベースとなる発信器のような生体内時計が必要であるが、これらの処理はマルチモーダル刺激の同時性に関わる上丘だけでなく、小脳や大脳基底核などの皮質下や前頭葉野など様々な部分がともに担っていると考えられている [111]。

これら脳の神経活動の様子は、機能的磁気共鳴画像法^{*1}や脳波^{*2}から観測できる事象関連電位^{*3}などで観測することができる。このうち、機能的磁気共鳴画像法は機材が非常に高価で、かつ測定中は非常に大きな騒音が発生するため、聴覚に関わる脳活動の測定にはほとんど用いられない[4]。一方、脳波から観測できる事象関連電位は、感覚刺激や身体の運動などに対応して生じる脳の電気的活動を、電位の変動という形で記録することができる。事象関連電位の計測は、比較的安価で実験参加者の負担が少なく、実験機材にも依るがおおよそ 2 kHz の時間分解能を持つため、聴覚現象の脳活動の測定だけでなく、様々な感覚刺激の反応を計測するために用いられる[4]。

近年では事象関連電位を調べることでマルチモーダル刺激の同時知覚を検討する研究が見られるようになってきた。例えば、Zumer *et al.* は 200ms のノイズ音と顔の左側に付けられた振動子を実験刺激として、刺激開始時間差を 0 ms, ±20 ms, ±70 ms, ±500 ms の 7 段階に変化させた時の同時性の判断を、刺激が呈示されているときだけボタンを押すことを求める Redundant target paradigm[4] を用いて行った。また別日に同じ刺激を用いて脳波を測定する実験も行った。脳波測定の際、実験参加者はいかなる反応もせず、刺激に対して受動的に聴取を行った。その結果、脳波は刺激開始時間差の違いによって異なる振る舞いを見せることが分かった。刺激開始時間差が ±20 ms 以下の場合は、脳の後部において、110 ms と 400 ms 付近に事象関連電位が観測され、刺激開始時間差が ±70 ms の時には P200 と呼ばれる特有の事象関連電位が観測された[105]。

脳波によって心理活動の振る舞いを計測する際の根本的な問題として、脳波や事象関連電位のデータが本当にその心理活動によるものなのかを判断することが難しいという点がある[112]。しかしながら、Zumer *et al.* による研究のように心理物理学的手法と組み合わせながら脳波の測定を行うなど、実験参加者の活動と脳波の振る舞いを結びつけることができる要素を、他の学問の手法を領域横断的に用い検討することによって解決できる可能性がある。従来は、機能的磁気共鳴画像法だけでなく脳波の測定であっても、非常に高価な実験機材と安定して測定するための環境が必要であったが、最近では、OpenBCI による 3D プリンタで出力が可能な脳波ヘッドセット[113] など、安価かつ安定性の高い脳波測定装置が手に入るようになった。今後の実験では、この認知神経科学的手法と心理物理学的実験手法を組み合わせた実験パラダイムを検討したい。

^{*1} Functional magnetic resonance imaging: fMRI

^{*2} Electroencephalogram: EEG

^{*3} Event-related potential: ERP

第 7 章

結論

本研究は、聴覚-触力覚刺激での時間統合窓の振る舞いを決定付ける要因の一端を解明するためには、両感覚刺激の特徴量を変化させた時の時間統合窓の変動を主観評価実験によって調査し、その結果を分析及び考察したものである。

第 3 章では、聴覚刺激の周波数とノイズ量をそれぞれ変化させた時の時間統合窓の変動を時間順序判断課題によって調査した。本実験では聴覚刺激として 250 Hz から 8 kHz までオクターブ間隔で分けた 6 種類の正弦波と、スペクトル尖度の値が調整された 6 種類のノイズ量が異なる刺激を用いた。実験参加者は振動子が接着されたスマートフォンを把持した状態で、聴覚刺激が先に呈示されたかどうかを回答した。その結果、聴覚刺激の周波数が低い場合、時間統合窓が小さくなる傾向があることが分かった。これにより、実験参加者が刺激の周波数関係によって両刺激が混ざり合ったと知覚された可能性が示唆された。また、聴覚-触力覚の同時性知覚は、両耳聴下での同時性知覚から説明できる可能性を得られた。一方、本実験においてはスペクトル尖度と時間統合窓の関係性を見出すことはできなかった。

第 4 章では、両刺激の周波数間隔を変化させた時の時間統合窓の変動を同時性判断課題によって調査した。本実験では、触力覚刺激を 200 Hz と固定し、聴覚刺激を 210 Hz から 3413 Hz まで段階的に変化させることで、両刺激の周波数がある間隔を持って分離している状態を作成した。実験参加者は第 3 章と同じ実験デバイスを用いて、両刺激が同時に発せられたと感じたかそうでなかつたかを二者択一によって回答した。その結果、周波数間隔の増加に伴い時間統合窓が若干小さくなる傾向が観測された。しかしながら、本実験結果は第 3 章の研究結果と反するものであり、その要因の一つとして実験時の回答手法における判断バイアスが大きく影響したと考えられた。

第 5 章では、聴覚刺激の強度を変化させた時の時間統合窓の変動をベイズ推論による信号検出理論によって調査した。本実験では、聴覚刺激は 1 kHz の正弦波を 80 dB_{SPL} と 68 dB_{SPL} になるよう調整されたものを用いた。実験参加者は第 3 章と同じ実験デバイスを用いて、両刺激が同時に発せられたと感じたかそうでなかつたかを二者択一によって回答した。その結果、聴覚強度条件が 80 dB_{SPL} の場合の時間統合窓は 118 ms で、68 dB_{SPL} の場合の時間統合窓は 126 ms であ

り、聴覚強度条件の違いにより時間統合窓の振る舞いに差異が生じることがわかった。また聴覚-触力覚における同時性知覚判断において、聴覚刺激が先に呈示されていても後に呈示されていてもおおよそ同じように両刺激のずれの判別が行うことが可能な実験参加者グループと、聴覚刺激が触力覚刺激より後に呈示されていた場合、そうでない場合と比べ両者にずれがあると知覚するのが遅くなるという傾向を持つ実験参加者グループが存在することがわかった。

これらの結果により、聴覚刺激の周波数と強度を変化させた際に時間統合窓がそれぞれ変動することが分かった。しかしながら、特徴量と時間統合窓の大きさの具体的な関係までを調査することはできなかった。その理由の一つとして、従来の同時性知覚の実験方法では、実験参加者の判断バイアスをコントロールすることが難しく、たとえ同じ実験環境、実験刺激条件を用いたとしても、結果が大きく異なることがあるからである。本論文においても、第3章での実験の一部と第4章で実験手法を変化させて実験を行った結果、時間統合窓の振る舞いが正反対のものになっていた。

また、同時性知覚において、特に聴覚刺激が触力覚刺激より後に呈示されていた場合に判断が特異に変化する群がある可能性が示唆された。一般に同時性知覚はある点を中心に対照となるとされていたが、それは分析の過程で正規確率密度関数を用いて曲線回帰を行うからである。しかしながら、本実験ではこの関数を用いない方法で分析したことにより、実験参加者群によって判断の対称性が見られない例があることを発見したことは注目に値する。

本研究では、ベイズ推論による信号検出理論を用いて、従来手法よりも判断バイアスをコントロールすることを考えた。しかしながら、この手法を用いても信号検出理論では両刺激が同時にどうかの二者択一を行うため、分析によって実験参加者の判断基準が取り除かれた弁別力が求められるとはいえ、いまだこの判断によるバイアスが残った状態であることは否めない。そのため、この判断の確信度を聞く第二種信号検出モデルや、脳波測定で得られる事象関連電位を調査する方法など、より判断バイアスをコントロールすることができる手法を、実験コストとの兼ね合いを考えながらさらに検討する必要がある。

ここまで、本実験で明らかになったことと、そこから言えることについて述べた。しかし、時間統合窓の振る舞いを決定付ける要因や、刺激特徴量との関係性が完全に明らかになったとは言えない。第5章や第6章で述べたように、いまだ時間統合窓の振る舞いを的確に観察することのできる実験手法は確立されておらず、これにより実験環境や実験刺激によって各研究結果にはらつきがある状況である。今後の研究では、この問題を解決する新たな実験手法を、心理物理学だけではなく、認知神経科学、さらには人工知能といった分野と連携しながら学際的に検討することが必要であろう。今後の研究では、先に述べた実験手法の問題を解決しつつ、デジタルコンテンツの体験を向上させる、ヒトの認知システムに基づいた聴覚-触力覚刺激呈示設計について議論を深められることを期待する。

参考文献

- [1] 本多明生, 神田敬幸, 柴田寛, 浅井暢子, 寺本涉, 坂本修一, 岩谷幸雄, 行場次朗, 鈴木陽一. 視聴覚コンテンツの臨場感と迫真性の規定因. 日本バーチャルリアリティ学会論文誌, Vol. 18, No. 1, pp. 93–101, 2013.
- [2] 総務省. 五感情報通信技術に関する調査研究会 報告書, 2000. https://www.soumu.go.jp/main_sosiki/joho_tsusin/policyreports/chousa/gokan/pdf/060922_2.pdf (2020年12月2日アクセス).
- [3] 日本音響学会. 音響キーワードブック. コロナ社, 2016.
- [4] 樋田浩一. 身体運動に対する聴覚フィードバックの認知メカニズムに関する研究. PhD thesis, 明治大学, 2017.
- [5] Khalafalla O Bushara, Jordan Grafman, and Mark Hallett. Neural correlates of auditory–visual stimulus onset asynchrony detection. *Journal of Neuroscience*, Vol. 21, No. 1, pp. 300–304, 2001.
- [6] M. Alex Meredith and Barry E. Stein. Descending efferents from the superior colliculus relay integrated multisensory information. *Science*, Vol. 227, No. 4687, pp. 657–659, 1985.
- [7] 岩宮眞一郎, 北川智利, 積山薫, 金基弘, 高木創, 笠松広司. 視聴覚融合の科学. 音響サイエンスシリーズ / 日本音響学会編, No. 11. コロナ社, 2014.
- [8] Rose Marie Pangborn and Brenda Hansen. The influence of color on discrimination of sweetness and sourness in pear-nectar. *The American Journal of Psychology*, Vol. 76, No. 2, pp. 315–317, 1963.
- [9] Joseph A Maga. Influence of color on taste thresholds. *Chemical Senses*, Vol. 1, No. 1, pp. 115–119, 1974.
- [10] G. J. Thomas. Experimental study of the influence of vision on sound localization. *Journal of experimental psychology*, Vol. 28, No. 2, pp. 163–177, 02 1941.
- [11] Charles E Jack and Willard R Thurlow. Effects of degree of visual association and angle of displacement on the “ventriloquism” effect. *Perceptual and motor skills*, Vol. 37, No. 3, pp. 967–979, 1973.

- [12] 大山正, 今井省吾, 和氣典二, 菊地正. 新編感覚・知覚心理学ハンドブック. 誠信書房, 1994.
- [13] 黒住幸一, 赤井田卓郎, 岡田清孝. テレビの映像と音声の相対時間差に関する検討. テレビジョン学会技術報告, Vol. 20, No. 50, pp. 37–42, 1996.
- [14] Rec. ITU-R BT.1359-1. Relative timing of sound and vision for broadcasting. Standard, International Telecommunication Union Radiocommunication Sector, Geneva, CH, 1998.
- [15] 一般社団法人日本音響学会. 音のなんでもコーナー Q and A (122), 2020 年 11 月 23 日アクセス. <https://acoustics.jp/qanda/answer/122.html>.
- [16] 難波精一郎, 茅阪直行, 桑野園子, H. Fastl, 菅野禎盛, 三浦雅展, 入交英雄, 鈴木陽一. 音と時間. 音響サイエンスシリーズ / 日本音響学会編, No. 13. コロナ社, 2015.
- [17] Harry Helson. The tau effect—an example of psychological relativity. *Science*, Vol. 71, No. 1847, pp. 536–537, 1930.
- [18] J. Cohen, C. E. M. Hansel, and J. Sylvester. A new phenomenon in time judgment. *Nature*, Vol. 172, pp. 901–901, 1953.
- [19] Edmund Husserl, 立松弘孝. 内的時間意識の現象学. みすず書房, 1967.
- [20] 佐藤文隆. 宇宙論への招待: プリンキピアとビッグバン. 岩波新書, 新赤版 6. 岩波書店, 1988.
- [21] 藤永保 (編). 新版 心理学事典. 平凡社, 1981.
- [22] H. Woodrow. Time perception. In S. S. Stevens, editor, *Handbook of experimental psychology*, pp. 1224–1236. John Wiley, 1953.
- [23] James Jerome Gibson, 古崎敬, 古崎愛子, 辻敬一郎, 村瀬晃. 生態学的視覚論: ヒトの知覚世界を探る. サイエンス社, 1985.
- [24] 松田文子, 調枝孝治, 甲村和三, 神宮英夫, 山崎勝之, 平伸二. 心理的時間: その広くて深いなぞ. 北大路書房, 1996.
- [25] Alan Johnston and Shin'ya Nishida. Time perception: Brain time or event time? *Current Biology*, Vol. 11, No. 11, pp. R427 – R430, 2001.
- [26] 藤崎和香. 同時性の知覚. 池田まさみ, 石口彰 (編), 認知心理学演習 日常生活と認知行動—応用・実践編—, pp. 39–68. オーム社, 2012.
- [27] J. L. Schnapf, T. W. Kraft, and D. A. Baylor. Spectral sensitivity of human cone photoreceptors. *Nature*, Vol. 325, No. 6103, pp. 439–441, 1987.
- [28] D. P. Corey and A. J. Hudspeth. Response latency of vertebrate hair cells. *Biophysical journal*, Vol. 26, No. 3, p. 499, 1979.
- [29] Mizobuchi Keiko, Kuwabara Satoshi, Toma Shinobu, Nakajima Yoshio, Ogawara Kazue, and Hattori Takamichi. Single unit responses of human cutaneous mechanoreceptors to air-puff stimulation. *Clinical Neurophysiology*, Vol. 111, No. 9, pp. 1577–1581, 2000.

- [30] Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, Steven Siegelbaum, A. James Hudspeth, Sarah Mack, 金澤一郎, 宮下保司, 岡野栄之, 和田圭司, 加藤総夫, 入來篤史, 藤田一郎, 伊佐正, 定藤規弘, 大隅典子, 笠井清登. カンデル神経科学. メディカル・サイエンス・インターナショナル, 2014.
- [31] Jennifer Garcia and Pbroks13. File:cartesian theater.svg, 2020年11月28日アクセス.
https://commons.wikimedia.org/wiki/File:Cartesian_Theater.svg.
- [32] Jean Vroomen and Mirjam Keetels. Perception of intersensory synchrony: a tutorial review. *Attention, Perception, & Psychophysics*, Vol. 72, No. 4, pp. 871–884, 2010.
- [33] L. R. Harris, V. Harrar, P. Jaekl, and A. Kopinska. Mechanisms of simultaneity constancy. In Romi Nijhawan and Beena Khurana, editors, *Space and time in perception and action*, pp. 232–253. Cambridge University Press, 2010.
- [34] Rebecca Winter, Vanessa Harrar, Marta Gozdzik, and Laurence R Harris. The relative timing of active and passive touch. *Brain Research*, Vol. 1242, pp. 54–58, 2008.
- [35] Valeria Occelli, Charles Spence, and Massimiliano Zampini. Audiotactile interactions in temporal perception. *Psychonomic bulletin & review*, Vol. 18, No. 3, pp. 429–454, 2011.
- [36] 岩宮真一郎. よくわかる最新音響の基本と仕組み：音の科学、技術、文化を基礎から学ぶ：「音」の世界を包括的に解説. How-nual 図解入門. 秀和システム, 第2版, 2014.
- [37] 平原達也. 音を聴く聴覚の仕組み (小特集: 誌上ビギナーズセミナー). 日本音響学会誌, Vol. 66, No. 9, pp. 458–465, 2010.
- [38] 日本聴覚医学会. 聴覚検査の実際 第4版. 南山堂, 2017.
- [39] Stanley J Bolanowski Jr., George A Gescheider, Ronald T Verrillo, and Christin M Checkosky. Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical society of America*, Vol. 84, No. 5, pp. 1680–1694, 1988.
- [40] 前野隆司. ヒト指腹部と触覚受容器の構造と機能. 日本ロボット学会誌, Vol. 18, No. 6, pp. 772–775, 2000.
- [41] John C Montgomery and John A MacDonald. Sensory tuning of lateral line receptors in antarctic fish to the movements of planktonic prey. *Science*, Vol. 235, No. 4785, pp. 195–196, 1987.
- [42] Georg Von Bekesy. Similarities between hearing and skin sensations. *Psychological review*, Vol. 66, No. 1, p. 1, 1959.
- [43] Ronald T. Verrillo. Vibration sensation in humans. *Music Perception: An Interdisciplinary Journal*, Vol. 9, No. 3, pp. 281–302, 1992.
- [44] George A Gescheider. Some comparisons between touch and hearing. *IEEE Transactions on Man-Machine Systems*, Vol. 11, No. 1, pp. 28–35, 1970.

- [45] Salvador Soto-Faraco and Gustavo Deco. Multisensory contributions to the perception of vibrotactile events. *Behavioural brain research*, Vol. 196, No. 2, pp. 145–154, 2009.
- [46] Fujisaki Waka and Nishida Shinya. Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental brain research*, Vol. 198, pp. 245–59, 07 2009.
- [47] 日本バーチャルリアリティ学会, 館暲, 佐藤誠, 広瀬通孝. バーチャルリアリティ学. 工業調査会, 2010.
- [48] Thomas Hauerslev. . . . in sensurround “earth-shattering”—a new dimension to the motion picture, 2011. https://www.in70mm.com/news/2011/in_sensurround/english/index.htm (2020年12月2日アクセス).
- [49] パイオニア株式会社. 身体で聴こう音楽会とボディソニックについて, 2020年11月21日アクセス. <https://jpn.pioneer/ja/corp/sustainability/karadadekikou/about/>.
- [50] 糸川英夫. ボーンコンダクションと音楽療法. 小松明, 佐々木久夫(編), 音楽療法最前線, pp. 155–183. 人間と歴史社, 1994.
- [51] 矢島卓郎. 体感音響装置による音楽療法の現状と展望—重症心身障害児者への更なる適用を目指して—. 目白大学総合科学研究, Vol. 14, pp. 67–79, 2018.
- [52] 任天堂株式会社. Nintendo64 振動パック, 2020年11月21日アクセス. <https://www.nintendo.co.jp/n01/n64/option/sindop.html>.
- [53] MARU-CHANG. Scph, 2020年11月22日アクセス. <https://maru-chang.com/hard/scph/index.php/>.
- [54] Red Bull UK. 操作方法を変えたビデオゲーム, 2014. <https://www.redbull.com/jp-ja/the-games-that-shaped-the-way-we-play> (2020年12月2日アクセス).
- [55] Evan-Amos. File:n64-rumble-pak.jpg, 2020年11月28日アクセス. <https://commons.wikimedia.org/wiki/File:N64-Rumble-Pak.jpg>.
- [56] 柏尾南壮. iPhone 6s/6s Plus の新機能を支える日本企業 分解で見えたその存在感, 2015. <https://www.itmedia.co.jp/mobile/articles/1510/21/news096.html> (2020年12月2日アクセス).
- [57] Apple Inc. Haptics - user interaction - ios - human interface guidelines, 2020年11月21日アクセス. <https://developer.apple.com/design/human-interface-guidelines/ios/user-interaction/haptics/>.
- [58] 福井美香, 南前恵子, 三瓶まり. 体感音響システムによるリラクセーション効果の検討(1): Microvibration による分析. 鳥取大学医療技術短期大学部紀要, Vol. 24, pp. 9–16, 1995.
- [59] 三瓶まり, 福井美香, 南前恵子. 体感音響システムによるリラクセーション効果の検討(2): 末梢血流量による分析. 鳥取大学医療技術短期大学部紀要, Vol. 24, pp. 17–22, 1995.
- [60] 南前恵子, 三瓶まり, 福井美香. 体感音響システムによるリラクセーション効果の検討(3):

β -endorphin 量変動による分析. 鳥取大学医療技術短期大学部紀要, Vol. 24, pp. 23–29, 1995.

- [61] 阪梨英樹. 音楽聴取時の全身振動と聴覚情報のマルチモーダル知覚. 電子情報通信学会技術研究報告. EA, 応用音響, Vol. 112, No. 76, pp. 1–6, 2012.
- [62] 岡崎龍太, 栗林英範, 梶本裕之. 分周刺激を用いた聴触覚変換手法による音楽体験向上. 日本バーチャルリアリティ学会論文誌, Vol. 21, No. 2, pp. 335–343, 2016.
- [63] Jaebong Lee and Seungmoon Choi. Real-time perception-level translation from audio signals to vibrotactile effects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2567–2576, 2013.
- [64] M Ercan Altinsoy. The quality of auditory-tactile virtual environments. *Journal of the Audio Engineering Society*, Vol. 60, No. 1/2, pp. 38–46, 2012.
- [65] Bernard D Adelstein, Durand R Begault, Mark R Anderson, and Elizabeth M Wenzel. Sensitivity to haptic-audio asynchrony. In *Proceedings of the 5th international conference on Multimodal interfaces*, pp. 73–76, 2003.
- [66] Massimiliano Zampini, Timothy Brown, David I Shore, Angelo Maravita, Brigitte Röder, and Charles Spence. Audiotactile temporal order judgments. *Acta psychologica*, Vol. 118, No. 3, pp. 277–291, 2005.
- [67] Alfonso Balandra and Hasegawa Shoichi. Sensitivity to haptic-audio envelope asynchrony. *arXiv preprint arXiv:1906.11571*, 2019.
- [68] Ira J. Hirsh and Carl E. Sherrick Jr. Perceived order in different sense modalities. *Journal of experimental psychology*, Vol. 62, No. 5, pp. 423–432, 1961.
- [69] Mauricio Orozco, Juan Silva, Abdulmotaleb El Saddik, and Emil Petriu. The role of haptics in games. *Haptics rendering and applications*, pp. 217–234, 2012.
- [70] Dean Chang. Haptics: gaming's new sensation. *Computer*, Vol. 35, No. 8, pp. 84–86, 2002.
- [71] 土肥哲也, 赤松友成, 新井伸夫, 井上保雄, 入江尚子, 時田保夫, 中右介, 町田信夫, 山極伊知郎. 低周波音: 低い音の知られざる世界, 音響サイエンスシリーズ / 日本音響学会編, 第16巻. コロナ社, 2017.
- [72] P. Lemmens, F. Crompvoets, D. Brokken, J. van den Eerenbeemd, and G. de Vries. A body-conforming tactile jacket to enrich movie viewing. *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 7–12, 2009.
- [73] 前田修. エンジン音リアルタイム合成技術の開発. ヤマハ発動機技報, 2001.
- [74] 星野博之. 自動車車室内騒音の音質評価 (小特集: 機械騒音の音質改善). 日本音響学会誌, Vol. 53, No. 6, pp. 462–465, 1997.

- [75] Robert Bristow-Johnson. Cookbook formulae for audio eq biquad filter coefficients, 2020 年 12 月 2 日アクセス. <https://webaudio.github.io/Audio-EQ-Cookbook/audio-eq-cookbook.html>.
- [76] Tactile Labs Inc. Haptuator mark II, 2012. <http://tactilelabs.com/products/haptics/haptuator-mark-ii-v2/> (2020 年 12 月 2 日アクセス).
- [77] Hsin-Yun Yao and Vincent Hayward. Design and analysis of a recoil-type vibrotactile transducer. *The Journal of the Acoustical Society of America*, Vol. 128, No. 2, pp. 619–627, 2010.
- [78] Uemura Yoshihisa, Takahashi Yu, Saruwatari Hiroshi, Shikano Kiyohiro, and Kondo Kazunobu. Automatic optimization scheme of spectral subtraction based on musical noise assessment via higher-order statistics. *IEEE International Workshop Acoustic Signal Enhancement*, 2008.
- [79] Olivier Lartillot. *MIRtoolbox 1.7.2 User's Manual*.
- [80] International Telecommunication Union. *Rec ITU-R BS.1770-4, Algorithms to measure audio programme loudness and true-peak audio level*. International Telecommunication Union, 2015.
- [81] 日本工業規格. JIS C 1510 振動レベル計 Vibration level meters. 日本工業規格, 1995.
- [82] 大久保街亜, 鈴木玄, Nicholls Michael E. R. 日本語版 FLANDERS 利き手テスト —信頼性と妥当性の検討—. 心理学研究, Vol. 85, No. 5, pp. 474–481, 2014.
- [83] Okazaki Ryuta, Hachisu Taku, Sato Michi, Fukushima Shogo, Vincent Hayward, and Kajimoto Hiroyuki. Judged consonance of tactile and auditory frequencies. In *2013 World Haptics Conference*, pp. 663–666, 2013.
- [84] Ellen M Parker. Auditory constraints on the perception of voice-onset time: The influence of lower tone frequency on judgments of tone-onset simultaneity. *The Journal of the Acoustical Society of America*, Vol. 83, No. 4, pp. 1597–1607, 1988.
- [85] Okazaki Satoshi and Ichikawa Makoto. Perceptual fusion and simultaneity for auditory stimuli. In *Proceedings of ICMPC-APSCOM 2014 Joint Conference*, pp. 316–319, 08 2014.
- [86] 岡崎聰. 2 純音の知覚的同時性の範囲が従う法則の記述及び聴覚基礎過程の解明. PhD thesis, 千葉大学, 2018.
- [87] 山本由紀子, 仁科エミ, 大西仁. 協和感研究の動向と課題. 認知科学, Vol. 22, No. 2, pp. 282–296, 2015.
- [88] 岡崎龍太, 蜂須拓, 佐藤未知, 福嶋政期, 梶本裕之. 触覚-聴覚間における周波数的協和性(触覚と力覚(2)). ロボティクス・メカトロニクス講演会講演概要集 2013, pp. 2A2–A14. 一般社団法人日本機械学会, 2013.

- [89] Reinier Plomp and Willem Johannes Maria Levelt. Tonal consonance and critical bandwidth. *The journal of the Acoustical Society of America*, Vol. 38, No. 4, pp. 548–560, 1965.
- [90] John MacCallum, Jeremy Hunt, and Aaron Einbond. Timbre as a psychoacoustic parameter for harmonic analysis and composition. In *International Computer Music Association*, 2005.
- [91] William Morris Hartmann and Dan C Sartor. Turning on a tone. *The Journal of the Acoustical Society of America*, Vol. 90, No. 2, pp. 866–873, 1991.
- [92] Lexi E Crommett, Alexis Pérez-Bellido, and Jeffrey M Yau. Auditory adaptation improves tactile frequency perception. *Journal of neurophysiology*, Vol. 117, No. 3, pp. 1352–1362, 2017.
- [93] Nadia Bolognini, Costanza Papagno, Daniela Moroni, and Angelo Maravita. Tactile temporal processing in the auditory cortex. *Journal of cognitive neuroscience*, Vol. 22, No. 6, pp. 1201–1211, 2010.
- [94] 石田翼. 信号検出理論について —数式と実践. <http://www5e.biglobe.ne.jp/~tbs-i/psy/tsd/tsdrev.pdf>, 2001.
- [95] 難波精一郎, 桑野園子. 音の評価のための心理学的測定法. 音響テクノロジーシリーズ / 日本音響学会編, No. 4. コロナ社, 1998.
- [96] Michael David Lee, Eric-Jan Wagenmakers, 井関龍太, 岡田謙介. ベイズ統計で実践モーリング: 認知モデルのトレーニング. 北大路書房, 2017.
- [97] 豊田秀樹. 実践ベイズモーリング: 解析技法と認知モデル. 朝倉書店, 2017.
- [98] 松原望. ベイズ統計学概説 : フィッシャーからベイズへ. 培風館, 2010.
- [99] 清水裕士. 心理学におけるベイズ統計モデリング. 心理学評論, Vol. 61, No. 1, pp. 22–41, 2018.
- [100] 小島寛之. 完全独習ベイズ統計学入門. ダイヤモンド社, 2015.
- [101] Michael David Lee. Bayessdt: Software for bayesian inference with signal detection theory. *Behavior Research Methods*, Vol. 40, No. 2, pp. 450–456, 2008.
- [102] Joan G Snodgrass and June Corwin. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of experimental psychology: General*, Vol. 117, No. 1, p. 34, 1988.
- [103] J. M. Barry. *SPLINS and SMOOTH: two FORTRAN smoothing routines*. Australian Atomic Energy Commission, 1973.
- [104] Sugano Yoshimori, Mirjam Keetels, and Jean Vroomen. Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Experimental Brain Research*, Vol. 201, No. 3, pp. 393–399, 2010.

- [105] Johanna M Zumer, Thomas P White, and Uta Noppeney. The neural mechanisms of audiotactile binding depend on asynchrony. *European Journal of Neuroscience*, 2020.
- [106] Norman F Dixon and Lydia Spitz. The detection of auditory visual desynchrony. *Perception*, Vol. 9, No. 6, pp. 719–721, 1980.
- [107] 中野靖久. 心理物理測定法. *Vision*, Vol. 7, No. 1, pp. 17–27, 1995.
- [108] Susan J Galvin, John V Podd, Vit Drga, and John Whitmore. Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic bulletin & review*, Vol. 10, No. 4, pp. 843–876, 2003.
- [109] 草薙邦広. 外国語教育研究における第二種信号検出モデル: 基本の理解とベイジアンモデリング. *広島外国語教育研究*, Vol. 21, pp. 169–185, 2018.
- [110] Brian Maniscalco and Hakwan Lau. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and cognition*, Vol. 21, No. 1, pp. 422–430, 2012.
- [111] Marc Wittmann. The inner experience of time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 364, No. 1525, pp. 1955–1967, 2009.
- [112] 入戸野宏. 事象関連電位(ERP)と認知活動: 工学心理学での利用を例に. *日本行動科学学会誌*, Vol. 42, pp. 25–36, 2003.
- [113] OpenBCI. Cyton biosensing board, 2020年11月28日アクセス.
<https://shop.openbci.com/collections/frontpage/products/cyton-biosensing-board-8-channel?variant=38958638542/>.
- [114] Stan Development Team. Stan reference manual: 15.3 notation for samples, chains, and draws, 2020年11月29日アクセス. https://mc-stan.org/docs/2_18/reference-manual/notation-for-samples-chains-and-draws.html.
- [115] John K. Kruschke, 前田和寛, 小杉考司, 井関龍太, 井上和哉, 鬼田崇作, 紀ノ定保礼, 国里愛彦, 坂本次郎, 桦取恵太, 高田菜美, 竹林由武, 徳岡大, 難波修史, 西田若葉, 平川真, 福屋いづみ, 武藤杏里, 山根嵩史, 横山仁史. ベイズ統計モデリング: R, JAGS, Stanによるチュートリアル. 共立出版, 2017.
- [116] 株式会社社会情報サービス. 統計WEB 統計の時間 Step1. 基礎編 27. 回帰分析 27-4. 決定係数と重相関係数, 2020年11月30日アクセス. <https://bellcurve.jp/statistics/course/9706.html>.

付録 A

事後確率の導出

まず、全事象 Ω を考える。さらに全事象をいくつかの既知情報 B_1, \dots, B_n に分割する。 B_1, \dots, B_n は互いに排反で、かつ全てを合わせると全事象 Ω になるとすると。

これらの事象が成り立つ確率 $P(B_1), \dots, P(B_n)$ を事前確率と呼ぶ。事前確率は予め分かっているものとする。

次に新事象 A が発生したとする。すると、 A は必ず $P(B_1), \dots, P(B_n)$ のどれかと交わる。この時、新事象 A が起きた時の事象 B が起きる条件付き確率 $P(B_n|A)$ を事後確率と呼び、次の式から求められる。

$$P(B_n|A) = \frac{P(A \cap B_n)}{P(A)} \quad (\text{A.1})$$

ここで乗法定理 $P(A \cap B_n) = P(B_n) \cdot P(A|B_n)$ を A.1 に代入する。

$$P(B_n|A) = \frac{P(A \cap B_n)}{P(A)} = \frac{P(B_n) \cdot P(A|B_n)}{P(A)} \quad (\text{A.2})$$

これをベイズの定理という。なお、 $P(A|B_n)$ は尤度と呼ばれる。

$P(A)$ は周辺尤度と呼ばれ、 $P(A) = P(A \cap B_1) + \dots + P(A \cap B_n)$ と書くことができる。これを A.2 に代入すると、

$$P(B_n|A) = \frac{P(A \cap B_n)}{P(A)} = \frac{P(B_n) \cdot P(A|B_n)}{P(A \cap B_1) + \dots + P(A \cap B_n)} \quad (\text{A.3})$$

となる。A.3 にも乗法定理 $P(A \cap B_n) = P(B_n) \cdot P(A|B_n)$ を適用すると、次の式が導かれる。

$$P(B_n|A) = \frac{P(B_n) \cdot P(A|B_n)}{P(A \cap B_1) + \dots + P(A \cap B_n)} = \frac{P(B_n) \cdot P(A|B_n)}{\sum_{i=1}^k P(B_i) \cdot P(A|B_i)} \quad (\text{A.4})$$

本文中では、9 個の白球と 1 個の黒球が入った A のツボと 2 個の白球と 8 個の黒球が入った B のツボがあり、そのどちらかが目の前に置かれている状況を想定し、ツボから一個球を取り出し

たら黒球であった時の、目の前のツボが A のツボか B のツボのどちらであるかを推定する時のベイズ統計学的解釈が例として挙げられていた。

ここで、黒球が取り出されたという事象を事象 A、球を A のツボから取り出す事象を事象 B_1 、B のツボから取り出す事象を事象 B_2 とする。理由不十分の原理により、全ツボのうちどれか一つが選ばれる確率は全て等しいと考えられるので、事前確率は以下のようになる。

$$P(B_1) = P(B_2) = \frac{1}{2} \quad (\text{A.5})$$

また、黒球が取り出される確率は以下のようになる。

A のツボの場合

$$P(A|B_1) = \frac{1}{10} \quad (\text{A.6})$$

B のツボの場合

$$P(A|B_2) = \frac{4}{5} \quad (\text{A.7})$$

例えば、A のツボから黒球が取り出される時の確率を求めるすると、ベイズの定理により、

$$\begin{aligned} P(B_1|A) &= \frac{P(B_1) \cdot P(A|B_1)}{P(A)} \\ &= \frac{P(B_1) \cdot P(A|B_1)}{\sum_{i=1}^2 P(B_i) \cdot P(A|B_i)} \\ &= \frac{P(B_1) \cdot P(A|B_1)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2)} \\ &= \frac{\frac{1}{10} \cdot \frac{1}{2}}{\frac{1}{10} \cdot \frac{1}{2} + \frac{4}{5} \cdot \frac{1}{2}} \\ &= \frac{1}{9} \end{aligned} \quad (\text{A.8})$$

となり、ツボから一個球を取り出したら黒球であった時の、目の前のツボが A のツボである確率は $\frac{1}{9}$ となる。

同じように、B のツボから黒球が取り出される時の確率は、

$$\begin{aligned}
P(B_2|A) &= \frac{P(B_2) \cdot P(A|B_2)}{P(A)} \\
&= \frac{P(B_2) \cdot P(A|B_2)}{\sum_{i=1}^2 P(B_i) \cdot P(A|B_i)} \\
&= \frac{P(B_2) \cdot P(A|B_2)}{P(B_1) \cdot P(A|B_1) + P(B_2) \cdot P(A|B_2)} \\
&= \frac{\frac{4}{5} \cdot \frac{1}{2}}{\frac{1}{10} \cdot \frac{1}{2} + \frac{4}{5} \cdot \frac{1}{2}} \\
&= \frac{8}{9}
\end{aligned} \tag{A.9}$$

となる。

付録 B

各個人の実験結果

ここでは、第 5 章の実験において集められた全 8 名の実験参加者の結果を表す。

各節にある図は、実験参加者の d' の推移を表している。左のパネルは聴覚強度条件が 80 dB_{SPL}、右のパネルは 68 dB_{SPL} の時の結果である。各パネルの横軸は刺激開始時間差 (SOA) を表し、縦軸は d' である。

各節にある表は、ベイズ推定で得られた、各刺激開始時間差における事後分布のパラメータである。mean は事後分布の平均値、sd は事後分布の分散、50 % は事後分布の中央値をそれぞれ表す。また、2.5 % と 97.5 % はそれぞれ事後分布の 2.5 % 分位点と 97.5 % 分位点を表し、この 2 点は 95 % 確信区間を意味する。Rhat^{*1}はマルコフ連鎖モンテカルロ法の収束判断の指標の一つであり、Rhat ≤ 1.1 のときに収束したと判断される [115]。

Participant 1

Participant 1 の d' の推移及び事後分布パラメータを図 B.1 と表 B.1、表 B.2 に表す。

^{*1} Gelman and Rubin potential scale reduction statistic の統計量 (\hat{R})。詳しくは、Stan Reference Manual の 15.3.1 Potential Scale Reduction[114] を参照。

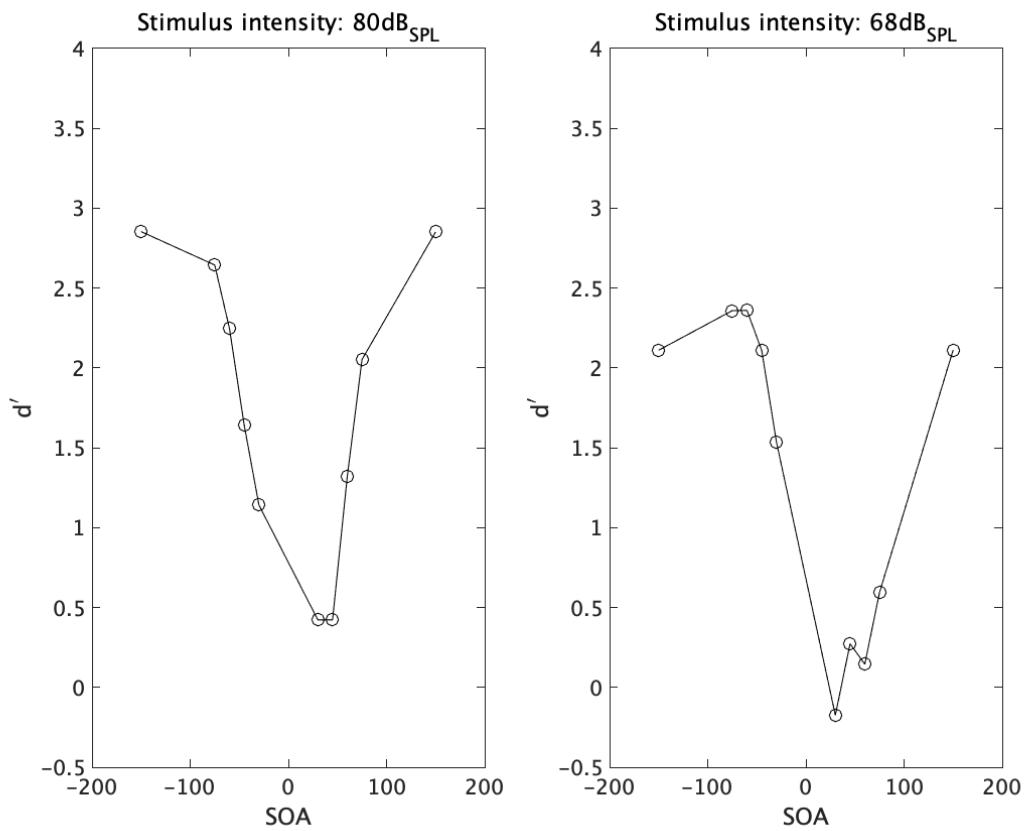


図 B.1 Participant 1 の d' の推移

表 B.1 Participant 1 80 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.854	0.492	2.840	1.928	3.864	1.000
-75 ms	2.644	0.469	2.632	1.758	3.603	1.000
-60 ms	2.247	0.454	2.232	1.399	3.178	1.000
-45 ms	1.642	0.474	1.625	0.762	2.630	1.000
-30 ms	1.142	0.504	1.120	0.212	2.191	1.000
30 ms	0.423	0.594	0.405	-0.698	1.637	1.000
45 ms	0.423	0.594	0.405	-0.696	1.641	1.000
60 ms	1.321	0.489	1.301	0.418	2.348	1.000
75 ms	2.053	0.457	2.035	1.200	2.994	1.000
150 ms	2.852	0.491	2.839	1.926	3.855	1.000

表 B.2 Participant 1 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.111	0.418	2.103	1.317	2.951	1.000
-75 ms	2.359	0.458	2.345	1.496	3.296	1.000
-60 ms	2.360	0.459	2.346	1.509	3.301	1.000
-45 ms	2.109	0.421	2.100	1.309	2.953	1.000
-30 ms	1.536	0.378	1.531	0.806	2.289	1.000
30 ms	-0.174	0.428	-0.173	-1.019	0.668	1.000
45 ms	0.273	0.390	0.272	-0.486	1.045	1.000
60 ms	0.145	0.399	0.144	-0.637	0.928	1.000
75 ms	0.594	0.376	0.592	-0.136	1.343	1.000
150 ms	2.113	0.424	2.105	1.312	2.966	1.000

Participant 2

Participant 2 の d' の推移及び事後分布パラメータを図 B.2 と表 B.3、表 B.4 に表す。

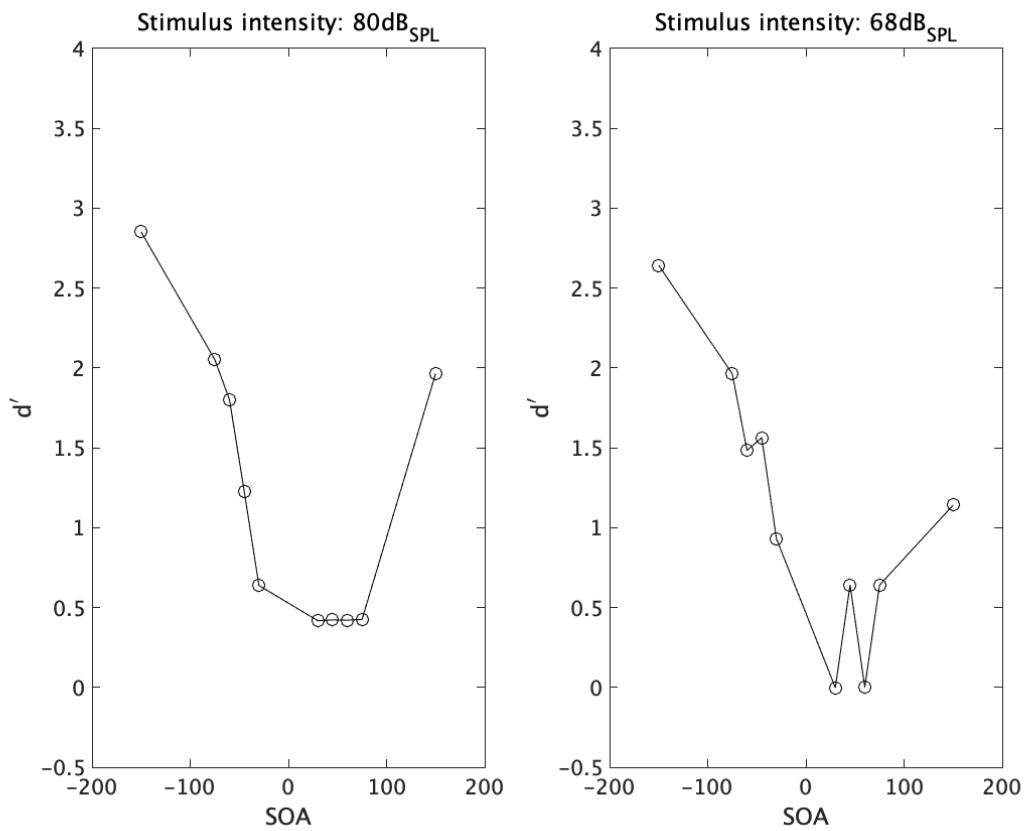


図 B.2 Participant 2 の d' の推移

表 B.3 Participant 2 80 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.852	0.490	2.836	1.930	3.862	1.000
-75 ms	2.052	0.459	2.037	1.196	3.000	1.000
-60 ms	1.802	0.464	1.785	0.944	2.765	1.000
-45 ms	1.229	0.494	1.206	0.318	2.259	1.000
-30 ms	0.638	0.559	0.621	-0.411	1.782	1.000
30 ms	0.417	0.592	0.404	-0.711	1.632	1.000
45 ms	0.422	0.593	0.406	-0.707	1.623	1.000
60 ms	0.421	0.595	0.406	-0.701	1.638	1.000
75 ms	0.425	0.594	0.408	-0.710	1.644	1.000
150 ms	1.964	0.460	1.947	1.112	2.920	1.000

表 B.4 Participant 2 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.643	0.471	2.631	1.751	3.608	1.000
-75 ms	1.964	0.459	1.946	1.111	2.913	1.000
-60 ms	1.483	0.480	1.464	0.597	2.486	1.000
-45 ms	1.563	0.473	1.545	0.686	2.551	1.000
-30 ms	0.930	0.524	0.908	-0.037	2.028	1.000
30 ms	-0.002	0.685	-0.006	-1.340	1.350	1.000
45 ms	0.642	0.561	0.624	-0.409	1.796	1.000
60 ms	0.001	0.685	-0.003	-1.337	1.357	1.000
75 ms	0.640	0.560	0.622	-0.403	1.792	1.000
150 ms	1.141	0.503	1.120	0.218	2.186	1.000

Participant 3

Participant 3 の d' の推移及び事後分布パラメータを図 B.1 と表 B.5、表 B.6 に表す。

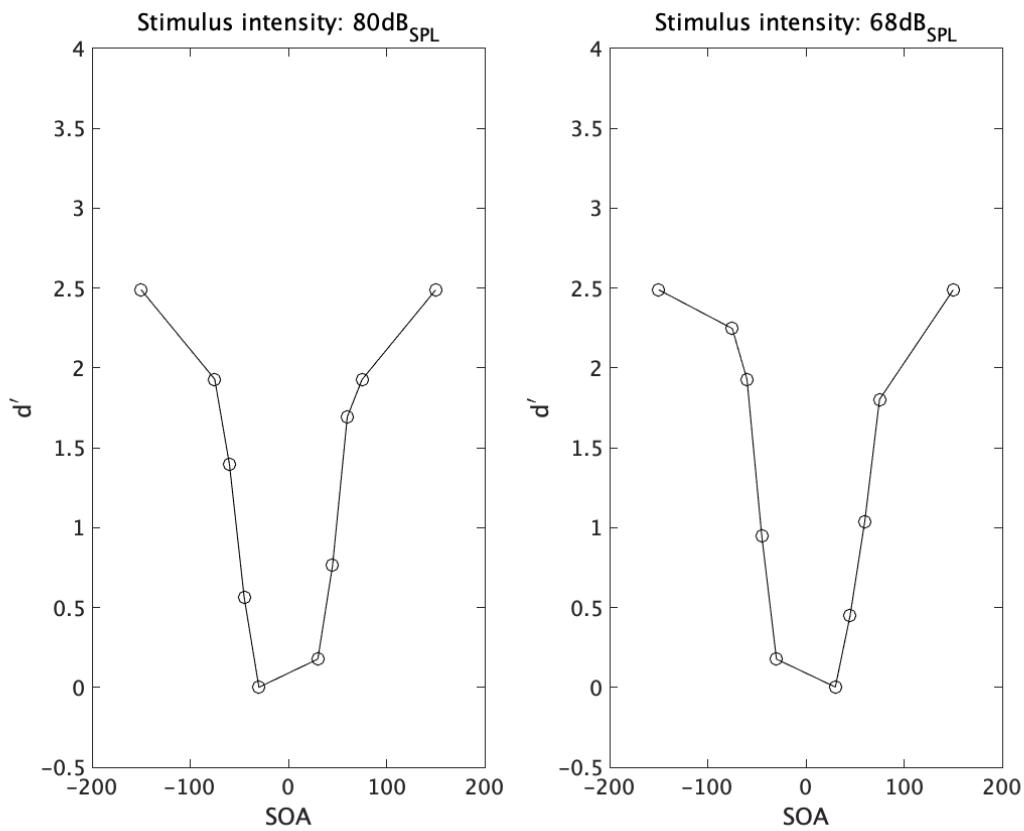


図 B.3 Participant 3 の d' の推移

表 B.5 Participant 3 80 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.490	0.457	2.477	1.635	3.420	1.000
-75 ms	1.929	0.403	1.924	1.155	2.735	1.000
-60 ms	1.395	0.386	1.389	0.654	2.167	1.000
-45 ms	0.563	0.404	0.559	-0.220	1.369	1.000
-30 ms	0.001	0.448	0.001	-0.875	0.880	1.000
30 ms	0.177	0.430	0.175	-0.655	1.023	1.000
45 ms	0.763	0.393	0.757	0.011	1.558	1.000
60 ms	1.693	0.392	1.686	0.940	2.475	1.000
75 ms	1.927	0.400	1.922	1.161	2.726	1.000
150 ms	2.488	0.459	2.477	1.633	3.429	1.000

表 B.6 Participant 3 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.490	0.463	2.478	1.614	3.441	1.000
-75 ms	2.247	0.427	2.237	1.436	3.113	1.000
-60 ms	1.929	0.401	1.921	1.163	2.731	1.000
-45 ms	0.951	0.391	0.946	0.197	1.735	1.000
-30 ms	0.177	0.426	0.175	-0.658	1.021	1.000
30 ms	0.002	0.445	0.003	-0.874	0.870	1.000
45 ms	0.448	0.407	0.442	-0.335	1.260	1.000
60 ms	1.040	0.385	1.035	0.300	1.808	1.000
75 ms	1.803	0.398	1.795	1.039	2.599	1.000
150 ms	2.488	0.460	2.475	1.624	3.425	1.000

Participant 4

Participant 4 の d' の推移及び事後分布パラメータを図 B.1 と表 B.7、表 B.8 に表す。

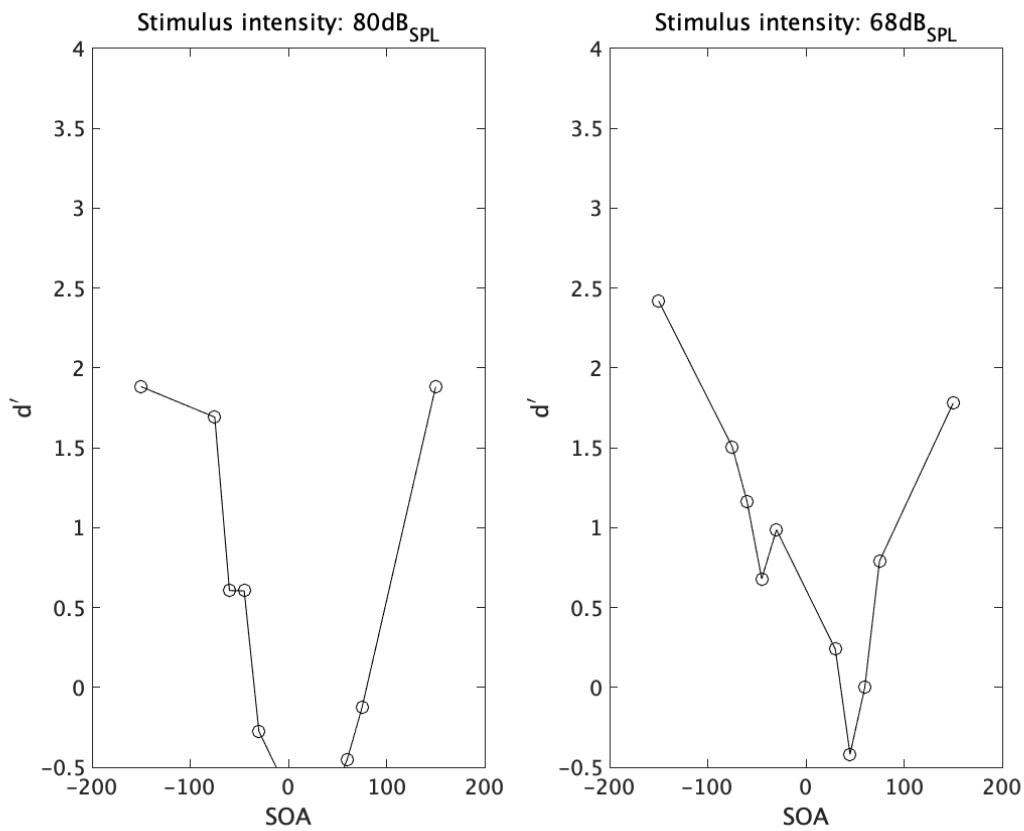


図 B.4 Participant 4 の d' の推移

表 B.7 Participant 4 80 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	1.883	0.411	1.874	1.100	2.711	1.000
-75 ms	1.692	0.392	1.687	0.938	2.474	1.000
-60 ms	0.606	0.354	0.606	-0.078	1.303	1.000
-45 ms	0.605	0.352	0.602	-0.076	1.302	1.000
-30 ms	-0.276	0.393	-0.271	-1.055	0.492	1.000
30 ms	-1.038	0.511	-1.015	-2.098	-0.094	1.000
45 ms	-0.680	0.437	-0.672	-1.562	0.152	1.000
60 ms	-0.450	0.409	-0.447	-1.268	0.338	1.000
75 ms	-0.127	0.380	-0.126	-0.877	0.618	1.000
150 ms	1.881	0.416	1.873	1.093	2.717	1.000

表 B.8 Participant 4 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.419	0.443	2.408	1.577	3.318	1.000
-75 ms	1.506	0.409	1.497	0.727	2.331	1.000
-60 ms	1.165	0.418	1.155	0.372	2.016	1.000
-45 ms	0.680	0.440	0.669	-0.161	1.569	1.000
-30 ms	0.987	0.421	0.980	0.187	1.834	1.000
30 ms	0.239	0.479	0.233	-0.692	1.191	1.000
45 ms	-0.419	0.596	-0.402	-1.635	0.710	1.000
60 ms	-0.001	0.508	-0.000	-0.998	1.001	1.000
75 ms	0.791	0.433	0.782	-0.036	1.665	1.000
150 ms	1.779	0.411	1.770	0.999	2.611	1.000

Participant 5

Participant 5 の d' の推移及び事後分布パラメータを図 B.1 と表 B.9、表 B.10 に表す。

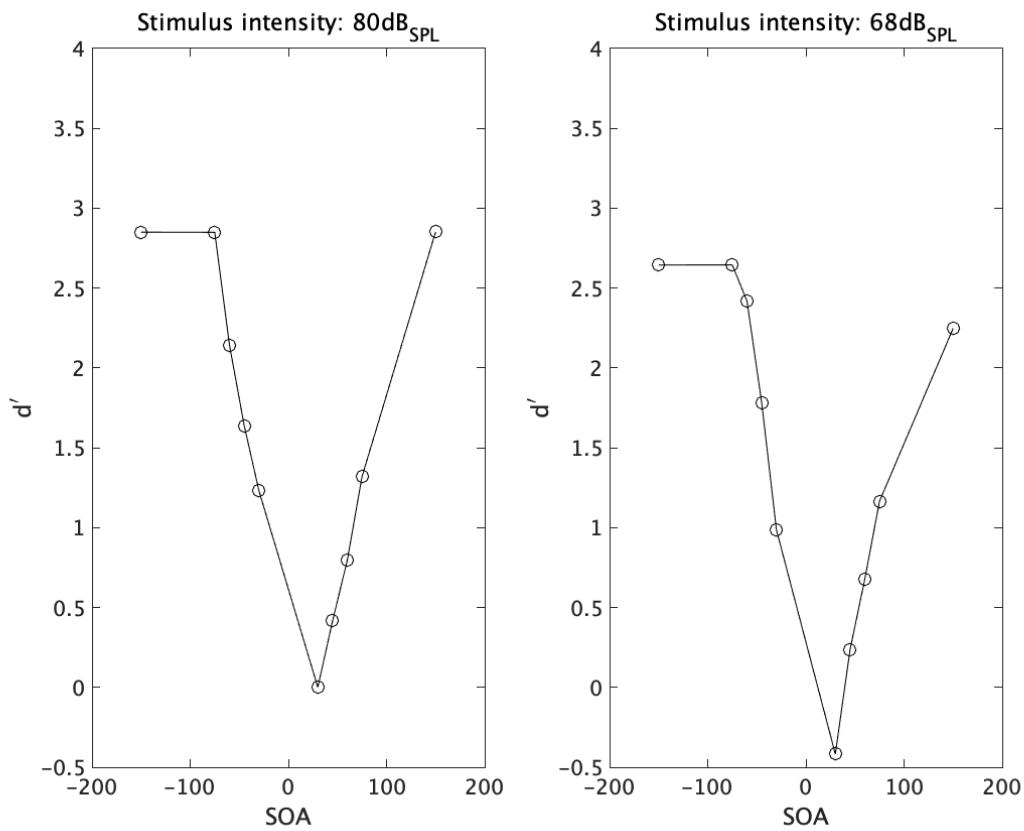


図 B.5 Participant 5 の d' の推移

表 B.9 Participant 5 80 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.850	0.489	2.837	1.930	3.840	1.000
-75 ms	2.849	0.493	2.834	1.927	3.868	1.000
-60 ms	2.144	0.458	2.129	1.291	3.086	1.000
-45 ms	1.639	0.470	1.622	0.766	2.607	1.000
-30 ms	1.231	0.497	1.211	0.319	2.268	1.000
30 ms	0.001	0.693	-0.002	-1.364	1.360	1.000
45 ms	0.416	0.597	0.401	-0.710	1.633	1.000
60 ms	0.795	0.537	0.773	-0.199	1.912	1.000
75 ms	1.319	0.488	1.300	0.414	2.331	1.000
150 ms	2.852	0.489	2.838	1.934	3.848	1.000

表 B.10 Participant 5 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.645	0.470	2.631	1.756	3.606	1.000
-75 ms	2.645	0.470	2.632	1.757	3.608	1.000
-60 ms	2.419	0.446	2.410	1.569	3.316	1.000
-45 ms	1.780	0.410	1.775	0.994	2.605	1.000
-30 ms	0.986	0.421	0.980	0.183	1.835	1.000
30 ms	-0.418	0.593	-0.402	-1.626	0.709	1.000
45 ms	0.236	0.472	0.230	-0.677	1.181	1.000
60 ms	0.678	0.436	0.671	-0.159	1.557	1.000
75 ms	1.165	0.417	1.158	0.375	2.010	1.000
150 ms	2.248	0.425	2.241	1.436	3.108	1.000

Participant 6

Participant 6 の d' の推移及び事後分布パラメータを図 B.1 と表 B.11、表 B.12 に表す。

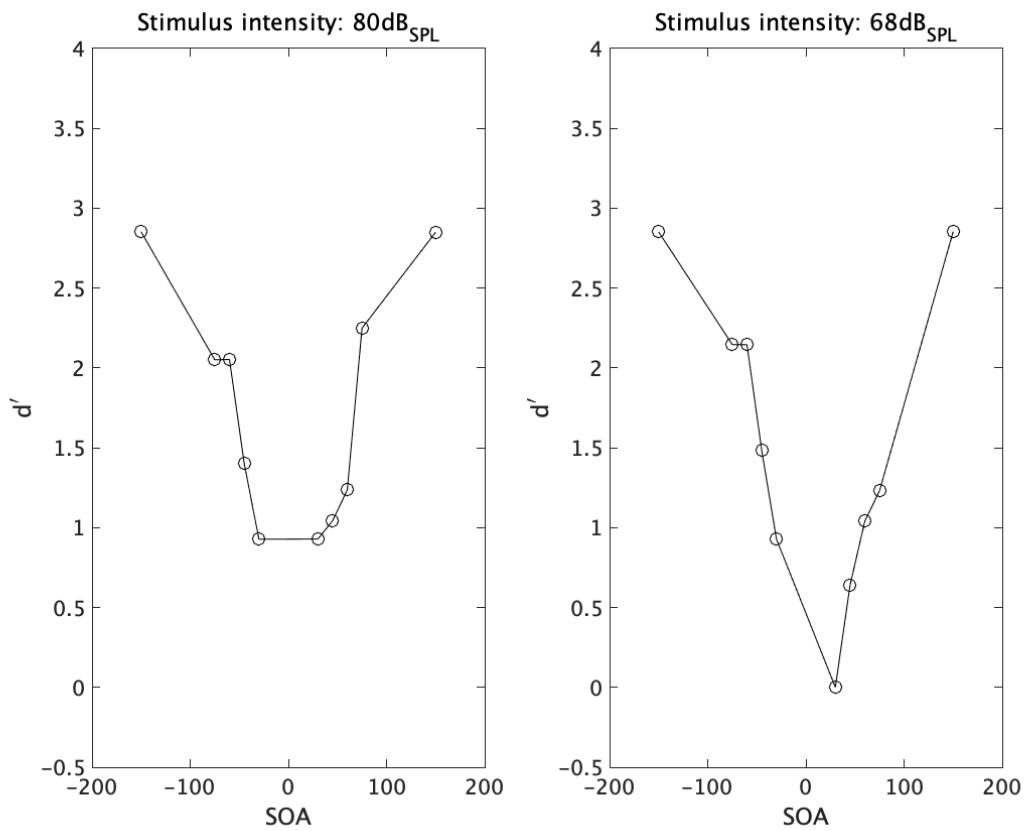


図 B.6 Participant 6 の d' の推移

表 B.11 Participant 6 80 dB_{SPL} の時のペイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.853	0.492	1.928	2.837	3.867	1.000
-75 ms	2.051	0.454	1.206	2.035	2.989	1.000
-60 ms	2.053	0.456	1.207	2.035	2.995	1.000
-45 ms	1.403	0.484	0.507	1.383	2.408	1.000
-30 ms	0.928	0.524	-0.042	0.908	2.018	1.000
30 ms	0.929	0.523	-0.035	0.905	2.017	1.000
45 ms	1.043	0.513	0.096	1.019	2.114	1.000
60 ms	1.236	0.496	0.321	1.215	2.275	1.000
75 ms	2.249	0.459	1.383	2.236	3.192	1.000
150 ms	2.850	0.495	1.926	2.834	3.867	1.000

表 B.12 Participant 6 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.853	0.492	2.837	1.928	3.867	1.000
-75 ms	2.051	0.454	2.035	1.206	2.989	1.000
-60 ms	2.053	0.456	2.035	1.207	2.995	1.000
-45 ms	1.403	0.484	1.383	0.507	2.408	1.000
-30 ms	0.928	0.524	0.908	-0.042	2.018	1.000
30 ms	0.929	0.523	0.905	-0.035	2.017	1.000
45 ms	1.043	0.513	1.019	0.096	2.114	1.000
60 ms	1.236	0.496	1.215	0.321	2.275	1.000
75 ms	2.249	0.459	2.236	1.383	3.192	1.000
150 ms	2.850	0.495	2.834	1.926	3.867	1.000

Participant 7

Participant 7 の d' の推移及び事後分布パラメータを図 B.1 と表 B.13、表 B.14 に表す。

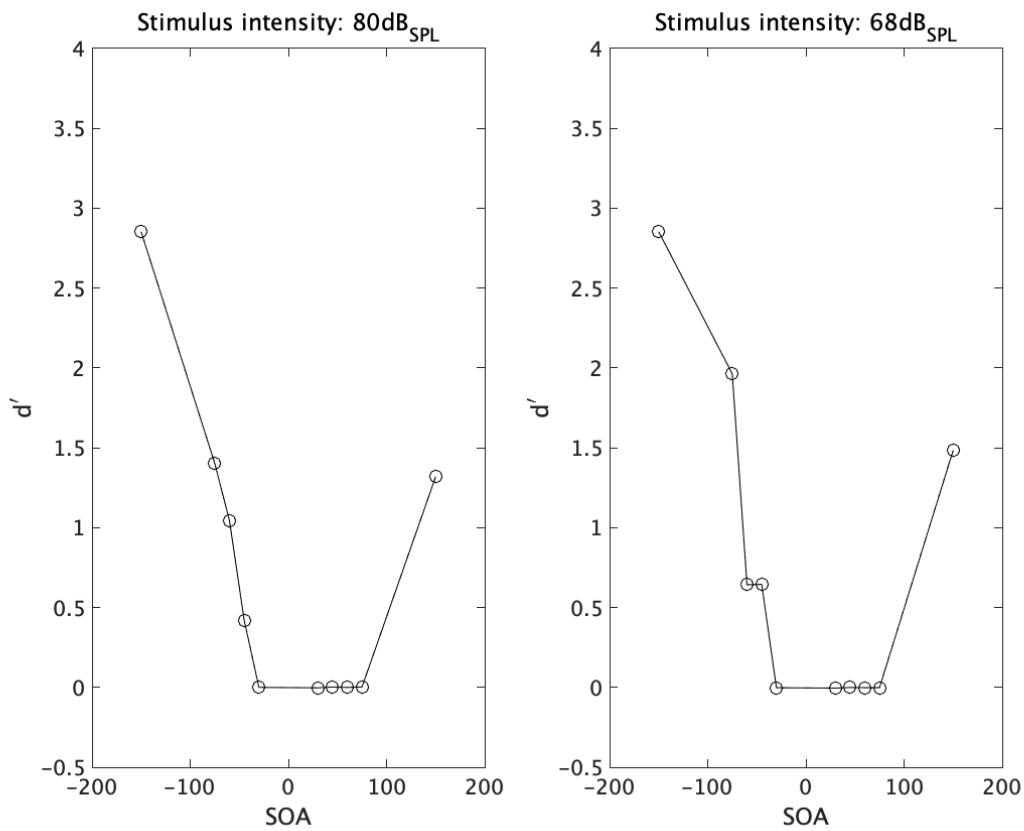


図 B.7 Participant 7 の d' の推移

表 B.13 Participant 7 80 dB_{SPL} の時のペイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.853	0.491	2.839	1.934	3.856	1.000
-75 ms	1.402	0.482	1.384	0.512	2.407	1.000
-60 ms	1.043	0.514	1.021	0.093	2.112	1.000
-45 ms	0.420	0.595	0.404	-0.713	1.634	1.000
-30 ms	0.000	0.688	0.000	-1.354	1.355	1.000
30 ms	-0.003	0.694	-0.001	-1.370	1.370	1.000
45 ms	0.002	0.690	0.000	-1.352	1.353	1.000
60 ms	-0.000	0.691	-0.003	-1.360	1.359	1.000
75 ms	0.002	0.687	-0.001	-1.354	1.359	1.000
150 ms	1.318	0.487	1.300	0.417	2.326	1.000

表 B.14 Participant 7 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.854	0.490	2.839	1.934	3.859	1.000
-75 ms	1.963	0.457	1.947	1.110	2.908	1.000
-60 ms	0.643	0.557	0.627	-0.397	1.785	1.000
-45 ms	0.645	0.557	0.628	-0.395	1.792	1.000
-30 ms	-0.002	0.689	-0.003	-1.363	1.360	1.000
30 ms	-0.004	0.691	-0.004	-1.371	1.353	1.000
45 ms	-0.001	0.692	0.003	-1.369	1.355	1.000
60 ms	-0.002	0.684	-0.003	-1.354	1.345	1.000
75 ms	-0.002	0.690	-0.006	-1.359	1.355	1.000
150 ms	1.484	0.477	1.465	0.601	2.469	1.000

Participant 8

Participant 8 の d' の推移及び事後分布パラメータを図 B.1 と表 B.15、表 B.16 に表す。

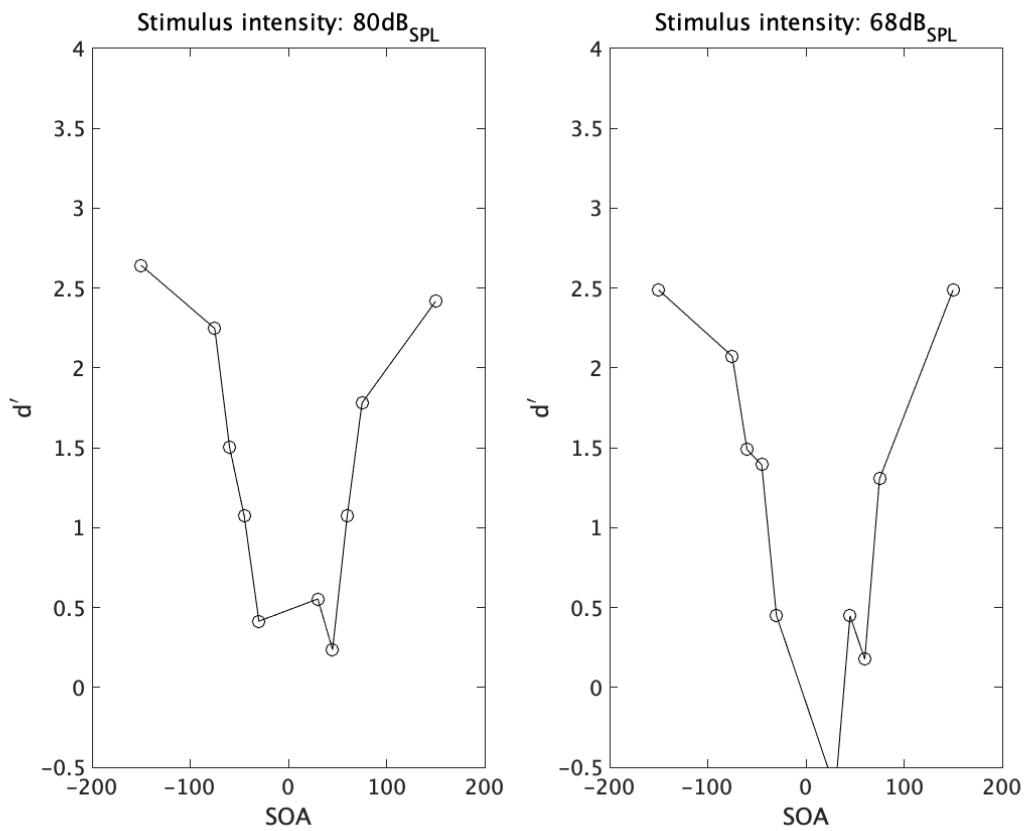


図 B.8 Participant 8 の d' の推移

表 B.15 Participant 8 80 dB_{SPL} の時のペイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.643	0.469	2.628	1.767	3.602	1.000
-75 ms	2.247	0.427	2.238	1.434	3.117	1.000
-60 ms	1.507	0.411	1.497	0.727	2.338	1.000
-45 ms	1.074	0.419	1.064	0.276	1.920	1.000
-30 ms	0.415	0.458	0.408	-0.464	1.337	1.000
30 ms	0.554	0.450	0.545	-0.306	1.463	1.000
45 ms	0.236	0.477	0.231	-0.687	1.196	1.000
60 ms	1.076	0.420	1.069	0.280	1.925	1.000
75 ms	1.780	0.411	1.772	0.998	2.605	1.000
150 ms	2.417	0.439	2.407	1.585	3.303	1.000

表 B.16 Participant 8 68 dB_{SPL} の時のベイズ推定結果

	mean	sd	50 %	2.5 %	97.5 %	Rhat
-150 ms	2.488	0.461	2.475	1.612	3.435	1.000
-75 ms	2.073	0.409	2.065	1.294	2.901	1.000
-60 ms	1.489	0.389	1.482	0.747	2.270	1.000
-45 ms	1.393	0.385	1.390	0.649	2.161	1.000
-30 ms	0.450	0.409	0.445	-0.345	1.260	1.000
30 ms	-0.640	0.558	-0.624	-1.789	0.410	1.000
45 ms	0.448	0.410	0.443	-0.341	1.265	1.000
60 ms	0.177	0.423	0.173	-0.646	1.014	1.000
75 ms	1.306	0.386	1.299	0.564	2.077	1.000
150 ms	2.490	0.461	2.474	1.627	3.437	1.000

付録 C

平滑化スプラインによる回帰曲線の残差平方和と決定係数

ここでは、各実験参加者の離散分布から平滑化スプラインを用い回帰曲線を求めた時の、回帰式の当てはまりの良さの指標である残差平方和と決定係数を表す。残差平方和 (SSE) とは回帰式と実際の値のズレの和を表し、値が 0 に近いほど回帰式の確率的誤差成分が小さく、近似が予測に有効であることを示す。

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (C.1)$$

ここで、 y_i は i 番目の変数の値、 \hat{y}_i は y_i の予測値である。

決定係数は一般に R^2 で示され、近似がデータの変動をどの程度適切に説明できるかを測定する指標である。0 から 1 までの値を取り、1 に近いほどその回帰式によって分散の大部分が説明できることを示す。例えば、決定係数の値が 0.8234 の場合、平均に関してデータの全変動の 82.34% を近似が説明できることを意味する。

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (C.2)$$

ここで、 y_i は i 番目の変数の値、 \bar{y} は全変数の平均値、 \hat{y}_i は y_i の予測値である。ここから、決定係数 R^2 は説明変数が目的変数をどれくらい説明しているか、つまり「回帰変動が全変動に対してどれだけ多いか=残差変動が全変動に対してどれだけ少ないか」を表す指標であることが分かる^{*1}。

^{*1} 詳しい解説は [116] 参照

表 C.1 聴覚強度条件が 80 dB_{SPL} の時の回
帰曲線の残差平方和と決定係数

	SSE	R²
Participant 1	2.256	0.706
Participant 2	0.528	0.926
Participant 3	2.377	0.685
Participant 4	2.081	0.808
Participant 5	2.111	0.785
Participant 6	1.010	0.802
Participant 7	0.279	0.966
Participant 8	1.685	0.749

表 C.2 聴覚強度条件が 68 dB_{SPL} の時の回
帰曲線の残差平方和と決定係数

	SSE	R²
Participant 1	3.831	0.589
Participant 2	0.768	0.879
Participant 3	2.735	0.666
Participant 4	1.028	0.844
Participant 5	3.083	0.703
Participant 6	1.403	0.826
Participant 7	0.780	0.916
Participant 8	2.553	0.739

付録 D

各実験の実験参加同意書

本研究で使用した実験参加同意書をそれぞれ表す。

第3章での実験

Participant Consent Form

You are invited to participate in a study entitled "*Audio-tactile synchrony and tolerable asynchrony perception*" which is being conducted by Rai Sato, student at Rochester Institute of Technology.

Purpose of the research: These tests are carried out in order to gain information on the human synchronicity perception of reproduced sound and vibration simultaneously; in particular, on the time variation of audio-tactile synchronicity perception when changing an acoustical stimulus.

What is involved in participating: In this study, you will be requested to answer whether you felt auditory modality came first or not when you heard and felt stimuli which reproduced via headphone and vibrator attached on a smartphone. All samples represent typical environmental sound and vibration that are reproduced at normal listening levels and vibration intensity, so there is no chance for hearing damage and neuromotor impairment. During the experiment, you will be asked to listen and feel the stimuli and answer the response using a custom graphical user interface (GUI). Please make your responses using the given screen on the smartphone. When tapping a "Next" button after answering your response by tapping an either "vibration first" or "sound first" button, the next sound sample will be presented. No feedback will be given to the subject at any point during the study, but questions about the study will be answered at the end of the study.

Each session will be completed in approximately 40 minutes or less. During session, you can take a break for as long as you would like. You may discontinue participation in this study at any point during the process.

In oral and written presentations of the results of this study, your responses will be assigned a number and will be grouped with those of the other participants for later analysis. The aim of this test is not to evaluate your individual performance, but to gather findings about human auditory and tactile responses to reproduced stimuli.

If you would like to learn about the results after the end of the study, please contact Rai Sato (rs1641@rit.edu). I would like to thank you for your interest in this investigation.

Participant's Signature: _____ Researcher's signature: _____

Participant's Name: _____ Date: _____

第4章、第5章での実験

第4章、第5章での実験はそれぞれタイトルや日付が異なるだけで、基本的な文言や体裁は同じである。ここでは第4章の実験で使用した実験参加同意書を表す。

主観評価実験参加の皆様

令和2年4月

聴覚と触力覚の時間同期感覚における知覚許容誤差に関する実験参加のお願い

東京芸術大学 音楽学部音楽環境創造科
佐藤 来

本日は「両刺激の周波数間隔に関する聴覚と触力覚の時間同期感覚に対する知覚許容誤差の変動 (Variation of perceptual tolerance for audio-tactile synchrony about the frequency separation of both modal stimuli)」に関する実験にご参加いただきありがとうございます。本研究は、音と振動が同時に提示された際に生じる知覚許容誤差を推測するモデルを構築することを目的としています。今回の実験はこの研究の一部として、両刺激の周波数間隔を変化させた時の、聴覚と触力覚の時間同期感覚に対する知覚許容誤差の計測を実施いたします。

本実験の趣旨を十分理解の上、協力してもよいとお考えの場合には、「同意書」にご署名の上、実験への参加を同意する意思表示をしていただきますようお願い申し上げます。

1. 実験方法

実験参加者は振動デバイスがついたスマートフォンを持ち、用意されたヘッドフォンを頭に装着した状態で、提示される音と振動に対し、刺激が同じタイミングで提示されていたのか、それとも刺激のどちらかが先行して提示されていたのかをスマートフォン上に表示されたボタンを押すことにより回答していただきます。実験に先立ち、実験担当者が実験の詳細、手順などを説明いたします。また、評価実験に先立ち、15分程度の実験説明及び予備施行を実施いたします。実験は実験担当者の指示により行われます。実験時間は合計1時間30分を予定しています。実験開始後、気分がすぐれない場合や、またその他の理由で、実験を中止することができます。その際は、ご遠慮なくご連絡ください。

2. 実験参加への意思確認

実験に参加するかどうかは、自由意志で決めてください。実験に参加する際は、別紙の同意書に署名を頂きます。いったん実験に同意された場合でも、実験担当者に連絡することにより、いつでも取り消すことができます。その場合は、それまでに行われた実験結果及び解析結果は破棄され、いかなる目的のためにも用いられる事はありません。また、実験に協力されない場合でも、不利益な対応を受けることは決してありません。

3. 健康被害の危険性

実験参加者の身体の状態には万全の注意を払います。実験には、音響再生装置及び振動再生装置を用いますが、再生される刺激は健康に支障のない強度で提示されるよう配慮を行っています。

4. 実験の参加により生じる不利益について

本実験に関して、計画上想定される不利益はありません。また、本実験以外においても、実験参加者に不利益が生じないことを前提としています。気になることがありましたら、どのような小さなことでも結構ですので、実験担当者にご連絡ください。

5. 個人情報の取り扱い

本実験では、実験結果が実験担当者以外にもれないように取り扱いを慎重に行います。解析には、実験課題に関する口頭などによる回答、データ情報以外は使用いたしません。また、その他得られた個人情報についても、実験担当者以外には漏れないよう慎重に取り扱います。

6. 研究結果の公表

ご協力によって得られた実験の成果は、学術目的で発表されることがあります。その際、データは集計して個人を特定できない形で、もしくは、個人を特定できないよう性別や年齢などに匿名化した上で発表します。実験データをランダムに割り当てたIDによって管理することにより、氏名などの個人を特定できる情報は実験データとは分離して取り扱います。

7. その他

実験開始後あるいは終了後、万一不利益が生じたとご判断された時は、下記までご連絡ください。

(実験担当者)

東京芸術大学 音楽学部 音楽環境創造科
学部4年 佐藤 来 (s2116232@ms.geidai.ac.jp)

(責任者)

東京芸術大学 音楽学部 音楽環境創造科
准教授 丸井 淳史 (marui@ms.geidai.ac.jp)

なお、文章等で不明な点がございましたら、直接担当者に問い合わせてください。

実験参加に関する同意書

(実験参加者)_____は、

東京芸術大学 音楽学部 音楽環境創造科 佐藤来が行う「両刺激の周波数間隔に関する聴覚と触力覚の時間同期感覚に対する知覚許容誤差の変動」に関する実験について

(実験担当者)_____より、十分な説明を受け、理解しましたので、実験の参加に同意します。

令和_____年_____月_____日

氏名_____

※ボールペンにてご記入ください。