

リズム知覚の基礎としての時間知覚に関する精神物理学的研究

中島, 祥好

<https://doi.org/10.11501/3159287>

出版情報 : 九州芸術工科大学, 1999, 博士 (芸術工学), 論文博士
バージョン :
権利関係 :

第 1 章 本研究の背景

1. 1 序論

本論文は、人間のリズム知覚についての理解を深めるために、基礎的な知見として、比較的短い時間長の心理尺度（＝精神物理関数、すなわち物理量と主観量との関数関係）を提案するものである。

様々な形態のコミュニケーションの中で、話し言葉（音声）をはじめとする聴覚を介したコミュニケーションは、日常生活において極めて重要である（Handel, 1989）。知覚心理学の分野において、人間が環境に関する情報の大半を視覚を通して得ているとの議論がしばしばなされるが、これは、人間が集団としてコミュニケーションを図りながら環境に適応している事実を無視した議論であるように思われる。現代においては、人間が、複雑な社会を形成することによって、本来の自然環境に集団として適応しており、個人にとっては、人間の作った社会が二次的環境とも言えるくらいの実体を伴っている。言葉を迅速に正しく伝え、音楽体験（必ずしも高級な芸術体験を意味しない）を共有することは、このような環境に適応するための重要な手段である。

音に関しては、楽器類の研究を別とすれば考古学的な研究が殆ど不可能であり、有史以前に発せられた音について考えるには、極端な思いこみをしないように注意しながら想像をめぐらせる以外にない。人類が集団で大型動物を捕らえることを学んだ頃には、音声によるコミュニケーションが相当高度に発達していたと推測することができる。ヒト以外にも集団で狩りをする動物はいるが、ヒトの場合には、自分よりもはるかに大きくて強く、一対一で戦っては勝てないような動物を、簡単な武器を用いて集団で倒すことができた点に特徴がある。このような場面で、素早いコミュニケーションを、身振り手振りによって行うことは、不可能であり、音声を用いる以外に手段はありえない。この際、現代の集団的なスポーツにおいても見られるような、「すすめ」、「さがれ」、「あぶない」、「たのむ」、「まかせろ」といった程度の内容を持つ音声コミュニケーションは、確立していたであろう。厳しい自然環境に適応する手段として、音声の果たした役割は大きかったと思われる。

チンパンジーのような類人猿が、ある程度の言語能力を持ちうることは、多くの研究によって明らかになっているが、彼らに欠けているのは、多くの種類の音声を次々に発するのための調音器官である。ヒトは、長い咽頭（「のど」のうち声帯より上の部分）を持つことによって、この問題を解決したが、一方では、食物などがのどに詰まる危険を背負うことになった。人類が、進化の歴史においてこのような危険を侵したことは、音声によるコミュニケーションがいかに重要であったかを裏付ける、間接的な証拠である。

話し言葉は、個人の力では知りえないような、環境に関する情報を伝える点でも重要である。我々が暮らしている環境に関する情報は、自分が直接得た情報だけではなく、他の人に教えてもらった情報をも組みこんだものであり、聴覚コミュニケーションは、このような情報を伝えるための最も重要な手段となっている。テレビ番組は、多くの現代人にとって欠かせない情報源となっているが、テレビの音を消して映像だけにすると、字幕入りの外国語映画などを除いて、情報源としての価値が失われてしまう。一方、テレビ番組の映像を見ずに、音だけにした場合には、かなりまとまった情報の得られる場合が多い。例えば、テレビのニュ

ースを見ずに聞くだけで、ある程度世の中の動きを知ることができる。(どのような目的に使われるのかは判らないが、テレビの音を聴くためのチャンネル番号を表示したラジオ・カセットは、市場に普及している。)

このように、われわれの日常生活において、聴覚コミュニケーションの果たす役割は大きい。聴覚コミュニケーションは話し言葉(音声)と音楽とに代表されるが、そのいずれにおいても、リズムが重要な役割を果たしている。ここで言うリズムとは、時間軸上に知覚される形、あるいは形の枠組みのことである

(Fraisse, 1982; 中島, 1994; 末富, 中島, 1998)。聴覚システムにおいては、音の時間方向のつながりをひとまとめに捉えることにより、知覚的体制化がなされており、その結果としてリズムが知覚されると考えられる(van Noorden, 1975)。聴覚コミュニケーションを目的として発せられた音響信号は、多くの場合、生き生きとしたリズムの感じを伴って知覚される。このことは、我々が、音響信号の伝える内容を、素早くかつ正しく把握するうえで役立っていると思われる(Handel, 1989)。例えば、規則的な時間パターンを構成する要素は、不規則的な時間パターンを構成する要素に比べて、順序、性質などが正確に知覚されやすい(Royer & Garner, 1966; Jones, Kidd, & Wetzel, 1981)。また、聴覚的な時間パターンを再生する(模倣する)課題においては、音楽における拍子のような構造をもつパターンのほうが、そうでないパターンよりも再生が容易である(Povel, 1981; Essens & Povel, 1985; Summers, Hawkins, & Mayers, 1986)。

Fraisse (1982) はリズム知覚に関して、その後の研究に対する指針を与えており、隣りあう時間間隔の長短関係が、リズム知覚の大変重要な手掛かりであると論じている。その際に典型的な西洋音楽の楽曲分析を行い、多くの場合、基本となるのは、1:2 あるいは 1:3 の比率をなす2種類の時価(音楽理論上の相対的な時間長)であることを指摘している。西洋音楽においては、時価の比率を定め、適切に産出する(演奏によって作り出す)ことが極めて重要である。したがって、音楽リズムの本質を理解するには、次々に産出される時間長の比率が、どのように知覚されるかを知ることが不可欠である。

時間長を高い精度で聴きとることは、話し言葉のリズムを知覚する際にも重要である。例えば、長い音節と、短い音節とを正しく区別することは、多くの言語において重要である。特に、日本語においては、時間構造の変化によって、促音、撥音、長音という独特の要素が生ずるために、正しいリズムを把握して話し、正しくリズムを聴き取ることができなければ、「おばさん」と「おばあさん」とを取り違えるなどの重大な誤解の生ずる恐れがある。これは、外国語を母語とする者が日本語を習得する際に、難関の一つとなっている(例えば、村木, 中岡, 1990; 粕谷, 佐藤, 1990)。

日常生活におけるリズム体験を理解するためには、次々に隣りあって生ずる時間長の比率がどのように知覚されるかについて、体系的な研究を行うことが重要である。特に、単純な刺激パターンに関する精神物理学的なデータの積み重ねは、音声学や音楽学の分野では行われにくいので、知覚心理学の分野における研究が求められる。本研究は、このような観点からなされたものである。

人間の聴覚的コミュニケーションにおいて、時間構造の知覚は大変重要な役割を果たしている。話し言葉においては、音節の長短や、アクセントの時間的位置などが、言葉の意味を左右するし、音楽においては、拍子、テンポ、音符の長短のパターンなどが全体の雰囲気を大きく左右する。その他の聴覚コミュニケーションを支える音、例えばドアをノックする音や、拍手においても、時間構造の産

出と知覚とが、大変重要である。一方、我々が聴覚を通じて環境を認知する場合にも、時間構造の知覚が重要な役割を果たすことがある。足音、鳥や虫の鳴き声、電車や汽車の走行音などが、その例である。このように考えてゆくと、聴覚心理学を研究するにあたり、時間知覚について相当な関心を払う必要のあることが解る。しかし、聴覚心理学の教科書、参考書などにおいて、時間知覚ないしリズム知覚に関してある程度のスペースが割かれるようになったのは、最近のことである(寺西, 1984; Handel, 1989; Moore, 1997)。本章の以降の部分では、本研究に関連する過去の研究を概観し、本研究の理論的背景について述べる。

1. 2 時間分解能

人間の耳は、音の流れを時間方向にどのくらい細かく分けて聴くことができるであろうか。この問いは、一見単純であるが、一筋縄では扱えないものである。0.1 ms 程度の、極端に短いクリック音を継時的に2つ呈示し、クリック音の、始まりから始まりまでの時間間隔をいろいろに変化させてみる場合、この時間間隔が1~2 ms 以上の長さをもたなければ、音が2つあるようには感ぜられず、1つだけの音があるように知覚される。このことは、古くから知られており(寺西, 1984)、時間知覚の研究者の間では半ば常識になっているが、近代的な批判に耐えるような実験データのかたちでは示されていない。実際に、クリック音の始まりから始まりまでを5 ms 程度にして聴いてみると、確かに音が2つあるように感ぜられることがあるが、その場合、2つの音が同時に鳴っているようにも聴こえ、時間的に分かれて聴こえていると言ってよいのかどうか定かではない。物理的には、音が2つ鳴れば、刺激パターン全体のスペクトルが変化するのであるから、そのことによる音色の変化も、判断の手掛かりになりうる。このような状況において、厳密な知覚実験を行うことには、被験者への教示をも含めて、困難な点が多い。Green (1971) は、このような問題を解決するために、新しいタイプの実験を行った。すなわち、厳密に等しいパワー・スペクトルを持ちながら、周波数成分によって耳への到達時間が異なるような音を被験者に聴かせたときに、どの周波数帯域の成分が遅れているかを音色の違いとして聴き分けるには、どの程度の遅れが必要であるかを測定した。その結果、最低でも2 ms 程度の遅れが必要であることが判った。Green は、その他の実験結果についても考察を進め、聴覚システムが音の時間的な変化を捉えうる最短の限界を1~2 ms 程度であるとしている。このあたりが、時間知覚について考える際の最短の時間となるであろう。

ところが、単純に聴覚システムの時間分解能が1~2 ms であるとは結論づけられない。音の時間構造の違いを聴き分けたとしても、どのような音、成分のあとに、どのような音、成分が続くかということまで、知覚しているのでなければ、真に時間構造を知覚しているとは言えないからである。Hirsh (1959) および Hirsh & Sherrick (1961) は、音の高さや、到来方向の異なる2つの音に、時間的なずれを与え、どのくらいのずれがあれば、どの音が先でどの音が後かと言う関係が正しく知覚されるのかを検討した。その結果、十分に訓練を積んだ被験者においては、20 ms くらいの時間差があれば、前後関係がぎりぎり判断できることが判った。このような前後関係の識別閾には、相当ばらつきがあるものの、聴覚に限らず、視覚、触覚においても同程度の値が得られており、感覚様相の違いに関わらず共通の仕組みが働いていることを示唆している。ただし、聴覚の場合には、より短い時間差で充分であるとするような実験結果もあり、時間の前後

関係が、前後関係そのものとして知覚されたのか、音色の違いとして知覚されたのかを見極める必要がある。そのためには、被験者に与えた訓練などを考慮しつつ、研究者自身による現象観察や、被験者から得られた言語報告などに基づいて考察を進める必要があり、今後の課題を残している。しかし、いま述べたような事柄を考慮した結果、多くの研究者が、前後関係を知覚するために必要な最低限の時間差は 20 ms 程度であると考えている (寺西, 1984; Moore, 1997)。

ここまでは、継時的に 2 つの音が呈示される場合について考えてきたが、継時的に 3 つ以上の音が呈示される場合には、新たな観点が必要である。3 つのクリック音を、音の始まりから次の音の始まりまでに、それぞれ 20 ms の間隔を与えて呈示しても、音が 3 つあるとは感ぜられず、2 つあるように感ぜられることが多い。はっきりと 3 つの音を聴き取るためには、少なくとも 50 ms くらいの間隔が必要である。これは、Cheatham と White (1954) の実験結果に見出されることである。彼らの結果を見れば、さらに多くの音が一定の間隔で次々に呈示される場合、その一つひとつを聴き取ってゆくためには、音の始まりから次の音の始まりまでに、少なくとも 100 ms 程度の間隔が必要であることが判る。音の数を一つひとつ数えて行くには、通常さらに長い時間間隔が必要である。被験者の熟練度、使用言語などによって差の生ずる可能性はあるが、200 ms 程度の時間間隔が必要であることが報告されている (ten Hoopen & Vos, 1981; Warren, 1993)。

二つの音の時間順序を知覚するには、20 ms の間隔で充分であったが、三種類以上の音が、同じ順序で繰り返して (例えば A B C A B C . . . というように) 呈示されるときにその順序を正しく判断しようとする、さらに長い時間間隔を要する場合がある。Warren, Obusek, Farmer と Warren (1969) は、「高い純音 (1000 Hz)」、「ヒス音 (中心周波数 2000 Hz の 1 オクターブ帯域雑音)」、「低い純音 (796 Hz)」、「バズ音 (40 Hz の方形波)」に、それぞれ 200 ms の持続時間を持たせ、時間的空隙を置かず、ここに挙げた順序で繰り返し呈示した。これらの音の順序を判断することを求められたとき、殆どの被験者が、一つひとつの音を聴き取り、4 種類の音を認めることができるにも関わらず、順序については自信を持つことができず、実際に正しく答えることができなかつた。一つひとつの音の持続時間を 700 ms くらいに増すことによって、被験者の半数以上が順序を正しく答えられるようになったが、熟練した被験者でも 300 ms を必要とした。ところが、刺激音に、英語や日本語の母音を用いるばあいには、一母音あたりに 110~150 ms を与えるだけで、順序の識別ができる (Thomas, Hill, Carroll, & Garcia, 1970; 寺西, 1977)。このように、用いる音によって違いが生ずる原因として、音声刺激を用いた場合、当該音声そのまま音の名前とすることができるので、聴取者が音に名前を当てはめる作業が容易になること、音声刺激の方が、より人工的な音刺激よりも、ひとつながりの音脈として聴き取りやすいこと、などが考えられる。これらの原因は、同じ事柄の両面である可能性もある (寺西, 1984)。

このように見えてくると、聴覚系の時間分解能を一つの値で代表させることは難しく、状況に応じて 1~700 ms の値によって表す必要がある。時間分解能について論ずる際には、具体的にどのような場面を想定しているのかをはっきりさせることが必要である。

1. 3 リズムを生ずる時間間隔

2つの音が続けて鳴るだけで、我々は、一種のリズムを感ずることができる。誰かがドアを2回だけノックする場合がよい例である。Nakajima, Shimojo と Sugita (1980) は、継時的に鳴らされる2つのクリック音の時間間隔を様々に変化させ、クリック音の始まりから始まりまでの時間間隔が150~2000 ms くらいの範囲にあるときに、身体の動きを連想させるような、明瞭なリズム感の生ずることを見出した。一方、Bolton (1894) は、物理的に等しい時間間隔において並んだ、物理的に均一な音の列が、知覚のうえでは2つずつや3つずつに群化する現象、すなわち「主観的リズム化」を発見した。この際、2, 3, 4, 6, 8個の音からなる群が安定して知覚されることが多く、大まかに見て、音の始まりから始まりまでの時間間隔が、115~1580 ms の範囲にあるときに、この現象が生じた。このように、リズムの知覚を生じうる時間間隔は、比較的狭い範囲に収まるようである。Warren, Gardner, Brubaker と Bashford (1991) は、規則正しいリズムからなる、馴染み深いメロディー（例えば「キラキラ星」）を、音楽の正式な訓練を受けていない被験者に聴かせ、どの曲のメロディーが鳴らされたのかが判るかどうかを調べる実験を行った。その結果、テンポが、一音符あたりの時間にして160~1280 ms くらいの範囲にあるときに、何のメロディーであるかが認知されることが判った。西洋音楽における実際の楽曲での一音符あたりの時間は、大部分が150~900 ms くらいの範囲に収まると言うFraisie (1982) の分析結果を考えあわせると、音楽のリズムも、BoltonあるいはNakajima が見出した範囲に含まれる時間間隔を基本にしていると考えられる。音声における音節やモーラなどの物理的な時間長も、我々になじみ深い日本語、英語などに関しては、概ねこの範囲に収まっているようである（例えば、Fujisaki & Higuchi, 1980; Han, 1994; Handel, 1989; Warren, 1993）。

ある物理的な時間長が、音楽や音声のリズムの基本単位となりうるためには、音の始まりから始まりまでの時間間隔が全てこの時間長を有するような、規則的な音列を作ったときに、全ての音が聴き取られるくらいに、長くなければならない。したがって、100 ms くらいが短いほうの限界になるであろう。しかし、二つの音の始まりが30 ms 程度離れておれば、その前後関係がはっきりと判るので、30~100 ms くらいの時間長も、リズムを構成する部分となりうる。一方、隣りあう音が知覚のうえで結びつくためには、あまり長い時間間隔があってはならない。Fraisie (1978) の言う<知覚される現在>（ひとまとまりに知覚される時間の広がり）の上限は約5000 ms であるが、精神物理学的な実験で用いられるような単純な条件においては、2000 ms 程度の値が上限となることが多い。例えば、継時的に示される2つの短音が知覚上結びつくには、1500~2000 ms くらいの時間間隔が長いほうの限界になる（Nakajima et al., 1980）。

本論文においては、聴覚コミュニケーションの基礎となるようなリズム知覚、時間知覚に興味の中心があるので、取り上げる時間長に関しては、確実に前後関係の知覚を生ずる50 ms 程度を下限とし、確実に知覚上の結びつきを生ずる1200 ms 程度を上限として、新たな実験、および過去の実験の再分析に基づいた考察を進める。この時間長の範囲は狭いように見えるかもしれないが、リズムの構成要素となる時間間隔の時間長は、殆どがこの範囲に含まれる。

1. 4 時間長の弁別

時間長の弁別は、時間知覚の基本的な問題として古くから取り上げられてきた。特に、時間に関して「ウェーバー比（相対弁別閾；ぎりぎりでは違いが判るために必要な物理量の増減幅の相対値）が一定になる」と言う「ウェーバーの法則」が成り立つかどうか、成り立たない場合にはどのような修正が可能か、と言う精神物理学の古典的な問題には、多くの研究者が興味を示してきた（例えば、Fraisse, 1978）。精神物理学一般における「ウェーバーの法則」はあくまでも大まかな近似に過ぎず、現実のデータがこの法則から系統的に逸脱することは、今日ではむしろ常識である。しかし、「ウェーバーの法則」は、弁別判断を求めると言うごく基本的な手続きを手掛かりにして、多くの異なった物理的次元、あるいは主観的次元を関連づけることを可能にすると言う点で、極めて有効な理論的枠組みである。そのため、法則を表す式に少数のパラメーターを追加して、修正を行う試みがなされている（Baird & Noma, 1978）。

本研究においても、時間長の弁別判断に関して、このような修正が可能かどうかを検討する。寺西（1984）は、聴覚刺激によって示された時間長のウェーバー比を過去の研究から集めて、グラフにまとめている。その結果、400～2000 ms の範囲においてウェーバー比がほぼ一定であるが、より長い時間長、およびより短い時間長に対しては、ウェーバー比が増加するという一般的な傾向が明らかになっている。ただし、2000 ms を超える時間長に関しては、時間長が直接に知覚されるものではなく、今回の考察の対象には含めないことにする。一方、400 ms 以下の範囲においては、話し言葉や音楽に類出するような時間長が関わってくるので、何らかの修正式の当てはめが必要であろう。この問題は、本論文の主要な話題の一つとして、後に取り上げる。

1. 5 時間長の心理尺度

本論文では、物理量と主観量との対応関係のことを、「心理尺度」と呼ぶことにする（「精神物理学的尺度」、「精神物理関数」と同じ意味である。）。物理量と主観量との関係については、時間長に限らず、多くの次元に関して研究がなされている。主として、量推定、量産出の実験結果を記述するために、Stevens（1957）は、「主観量は物理量のべき（冪）関数として表される。」との「べき関数の法則」を提唱した。この法則は、相当な量のデータを集めたときに、粗い近似として成り立つのみであるが、異なる感覚様相にまたがる多くのデータを、簡単な式によって関連づけることができるので、多くの研究者に支持されている。時間長のべき数（主観的な時間長が、物理的な時間長の何乗に比例するかという値）は、典型的には 0.9 くらいである（Eisler, 1976）。もっとも、物理的な時間長と主観的な時間長とが比例する（べき数が 1 になる）との結果や（Bobko, Thompson, & Schiffman, 1977）、べき数が 1 より僅かに大きくなることを示唆する結果も（Gregg, 1951; Stevens & Galanter, 1957）報告されており、統一的な見解は得られていない。全体として、べき数は 1 に近いことが多く、物理的な時間長と主観的な時間長とが正比例するとみなすような、日常的な直観から大きく外れてはいない。（このような直観を受け容れることは、音楽心理学の分野においても珍しくない。）

本研究において取り上げるような数百ミリ秒以下の時間長を中心とした実験は、

あまり広く行われていない。しかしながら、日常生活において、最も精密なリズムの把握が要求される、話し言葉や音楽の知覚においては、数百ミリ秒以下の時間間隔が重要な役割を果たしており、この範囲においても研究を進めることが極めて重要である。Zwicker (1969/70) は、この点に関して、様々な音刺激パターンを用いた先駆的な研究を行い、被験者に対して、比較時間を標準時間の2倍あるいは半分の長さに感ぜられるように調整することを求めた。その結果、充実時間を用いた条件においては、100 ms 以下の時間長に対して、物理量の増加率よりも対応する主観量の増加率の方が少ない（つまり、主観量を縦軸に取った両対数軸座標上で、心理尺度の傾きが1よりも平坦に近い側にずれる）傾向が認められた。空虚時間を用いた条件では、この傾向が更に広い範囲に現れたと言うことであるが、測定値のばらつきが大きいため、データは公表されていない。本研究においては、まさにこのような条件に焦点を絞ることになる。

時間長に関しては、物理量と主観量とが正比例するとの見解が、日常の常識として浸透している。身近な例として、家庭用のコンピューター音楽のシステムは、そのような考えかたに基づいて、楽譜を音に変換している。20～30年前には、専門家の実験室に留まっていたコンピューター音楽が、今や、テレビ放送、ゲーム、インターネットなどを通じて、我々の音楽体験の重要な要素となっているが、芸術表現以前の「正しいリズムの作りかた」に関しては、個々の音楽家、技術者の体験の蓄積に委ねられている。このような現状に対して、理論的な基礎を提供するための手がかりを与えることは、本研究の重要な目的である。本来、物理量と主観量との関わりについて慎重に考察することを求められる、リズム知覚の研究においてさえ、上に述べたような「常識」が通用している場合が多い（例えば、Fraissee, 1982）。このことに関しては大がかりな再検討が必要であろう。

一方、精神物理学の立場から見た場合には、比率判断の結果と弁別判断の結果とを、統一的に解釈することが可能かどうかを検討することが大変重要であり、本研究は、限られた範囲においてはああるが、そのような例を与えるものである。

1. 6 本研究で取り上げる実験状況

数秒よりも短い時間間隔を知覚することは、言語音声、音楽、および身体の動きなどに関連しており、我々の日常生活において極めて重要である。この際に最も大きな役割を果たす感覚様相は、聴覚である。聴覚システムは、時間構造を処理するうえで優れた能力を有すると考えられている。したがって、聴覚刺激によって区切られた短い時間長は、時間知覚に関する理論的基礎を確立するために適切な材料である。時間分解能、時間長弁別などに関する過去のデータには、このことを示唆する例が多い（寺西, 1984）。本研究で取り上げるような数百ミリ秒以下の時間長に関しては、Grondin & Rousseau (1991) および Grondin (1993) が広義の弁別実験を行っており、視覚、振動感覚に比べて、聴覚が時間知覚のために有利な感覚様相であることを示唆するような結果を得ている。

Fraissee (1978) が述べるように、短い時間長の「知覚」、および長い時間長の「評価」の間には、様々な違いがある。最も大きな違いは、2秒を超えるような長い時間長を、記憶を介せずに直接知覚することは難しく、より短い時間長においては、容易に時間間隔全体を知覚しうると言う点である。例えば、2つのクリック音の始まりから始まりまでの時間間隔を、0.5秒くらいから少しずつ伸ばしながら聴いてゆくと、時間間隔が1.5～2秒を超えるあたりで、急にパターン全

体をひとまとまりのものとして知覚することが困難になる。ここでは、時間知覚の基本的な仕組みを探ることが目的であるから、ひとまとまりに知覚されうる短い時間間隔のみを取り上げることにする。

ここで時間長を示す際に、いわゆる空虚時間を用いるか、充実時間を用いるか、あるいは、別種の時間間隔を用いるかが、問題となる。聴覚における充実時間、すなわち、判断の対象となる時間長が音エネルギーで満たされた時間間隔、を実験に用いる場合、音の強さあるいは周波数が、主観的な時間長に複雑な影響を及ぼすことが判っている（例えば、Berglund, Berglund, Ekman, & Frankenhaeuser, 1969; Burghardt, 1973）。そこで、快適聴取レベルのごく短い音（の始まり）によって両端を区切られた空虚時間を用いることが、研究の第一歩として適切ではないかと考えた。ここで、「空虚時間」という用語は、ごく短い刺激の始まりから、次のごく短い刺激の始まりまでに、特に目立つような、刺激、変化がないときに、この時間間隔を示すために用いる。全く刺激の存在しない時間間隔や、持続する刺激に挟まれる時間的空隙を、同じように「空虚時間」と呼ぶことがありうるが、本論文では、用語の意味をこのように限定する。音の始まりから始まりまでの時間長が、リズム知覚に際して本質的な役割を果たすことは、多くの研究者が認めており、そのことを支持する実験データも得られている（Sternberg, Knoll, & Zukofsky, 1982; Handel, 1993）。

数百ミリ秒を超えるような時間長に関する判断を被験者に求める場合、被験者にとっては、全体の長さを2つ以上の部分に分割してタイミングを取るほうが、課題が容易になる場合があり、実験結果を解釈するうえで、不安定な要因になりうる（Fetterman & Killeen, 1990; Grondin, 1992）。一方では、言語音声における音節や、音楽における音符などに相当する時間長の大部分を扱うことができるのでなければ、研究そのものの意味が薄れてしまう。そこで、本論文では主に、聴覚刺激によって区切られた、600 ms以下の空虚時間の心理尺度について検討し、必要に応じて、1秒前後まで尺度の適用範囲を広げることにする。

時間知覚についての様々な実験結果を説明しようとするような、一般的な理論の必要性が以前から指摘されている（Allan, 1979; 寺西, 1984; Killeen & Weiss, 1987）。これまでに提出された定量的なモデルのいくつかは、新しい実験的研究を導くうえで重要な役割を果たしている。しかし、多くの体系的なモデルが提案されているにもかかわらず、広く受け容れられているものは多くない。現在求められるのは、再現性の高いデータに基づいて、解りやすいモデルを構築することではないかと思われる。そのためには、実験条件にある種の制限を加えることも必要である。さらに、神経生理学的な研究とも対応づけられるような理論を確立するために、単純な心理尺度によって、過去の研究をも含めたデータを大局的に記述することが望まれる。すなわち、当てはまりの精度を多少犠牲にしても、多くの異なった状況を扱うことのできるようなモデルを組み立てることが、強く求められている。

本論文においては、9つの実験、および2つの再分析を報告する。多少の例外はあるが、上述の目的に従い、次のように限定された実験状況のみを取り扱う。

- (i) 被験者が、ごく短い短音によって区切られた、2つないし3つの、空虚な（あるいは擬似的に空虚な）時間間隔の関係について判断する。
- (ii) 空虚な（あるいは擬似的に空虚な）時間間隔の大半が、40~600 msの範囲に収まる。

本研究の理論的な出発点は、著者（中島，1979）の提出したくつけくわえ仮説 Supplement Hypothesis>である。この仮説は、「主観的に空虚である時間間隔の主観的な長さは、その物理的な長さに約 80 ms の定数を加えたものに、比例する。」と言うものである。発表当時、この仮説は純然たる作業仮説であり、精神物理学的な研究に不可欠である、種々の数学的関数の当てはまりを比較するという作業を伴っていなかったため、本研究において、本格的にこの仮説の妥当性、発展性を検討することにした。

中島（1979）の研究は、2つの実験からなる。そのうち実験1は、くつけくわえ仮説>を考え出すきっかけとなったものであるから、やや詳しく紹介しておく。この実験は、「分割時間の過大評価」に関するものである。2つのごく短い刺激によって区切られ、それ以外の刺激を含まないような時間間隔を、空虚時間と呼ぶ。時間間隔の両端を区切る2つの刺激の間に、一つまたはそれ以上のごく短い刺激が含まれるならば、その時間間隔は分割時間と呼ばれる。10 s より短い分割時間は、同じ物理的な長さをもつ空虚時間に比べて、より長いと知覚される傾向がある。これが、Hall と Jastrow（1886）によって発見され、その後、多くの研究者によって研究されてきた、「分割時間の過大評価」である（Israeli, 1930; Fraisse, 1961, 1965; 松田, 1965, 1967; Jones & MacLean, 1966; Buffardi, 1971; Thomas & Brown, 1974; Adams, 1977）。

この錯覚の量が、調整法によって測定された。実験に用いられた分割時間は、全て、両端の区切音および1つの分割音からなっている。すなわち、継時的に現れる3つの短音（A、B、C）が、隣接する2つの空虚部分（空虚時間；AB および BC）を示している。この3つの音は、いずれも 1 ms の短音であった。この分割時間の全体の物理的な長さ（AC）は、120、240 あるいは 480 msであった。物理的な分割比、すなわち、最初の空虚部分の長さの全体の長さに対する比率（AB / AC）は、1/12、2/12、4/12、6/12、8/12、10/12、あるいは 11/12 であった。このパターン全体が調整法の標準時間（標準刺激）となる。一方、両端を 1 ms の短音（D、E）によって区切られた空虚時間が、調整法の比較時間（比較刺激）として、標準時間の前あるいは後の、時間的に充分離れた位置に呈示された。ここで、標準時間のほうが先に呈示されるような時間順位を「第1時間順位」と呼び、比較時間のほうが先に呈示されるような時間順位を「第2時間順位」と呼ぶ。被験者は、比較時間の長さ（DE）を、標準時間の全体の長さ（AC）と等しく感ぜられるように調整し、その結果生じた比較時間の物理的な時間長を主観的等価値とした。それぞれの被験者について、標準時間のパターンごとに、また、時間順位ごとに、上昇系列と下降系列との測定が、1試行ずつなされた。

いま述べた「分割条件」の試行以外に、上記の分割時間（ABC）から分割音（B）を取り除いた、空虚時間（AC）を標準時間とする、「対照条件」の試行が設けられた。各条件の時間チャートを図1に示す。ここで、分割条件における主観的等価値が、対照条件における主観的等価値よりもどのくらい大きいかを示す値が、分割音によって生じた過大評価の量を示すと考えることができる。対照条件においては、標準時間の時間長ごとに、また、時間順位ごとに、上昇系列と下降系列との測定が、2試行ずつなされた。被験者数は7名であった。

この実験の結果、(i) 物理的な分割比が 4/12~8/12 の範囲にあり、かつ、(ii) 第1時間順位が用いられているような場合に、平均値において、数十ミリ秒に達するような過大評価が見られた（図2）。なお、空虚時間を標準時間とした

対照条件では、物理的等価値と主観的等価値との差が殆ど生じていない [いずれの時間順位においても、平均値のうえで 10 ms 以上の差が生じた条件はなかった。]。そこで、次の分析 (図 3) においては、分割条件における主観的等価値が、物理的等価値に比べてどのくらい大きいかを計算し、この値を、分割音によって生じた過大評価の量であるとみなす。

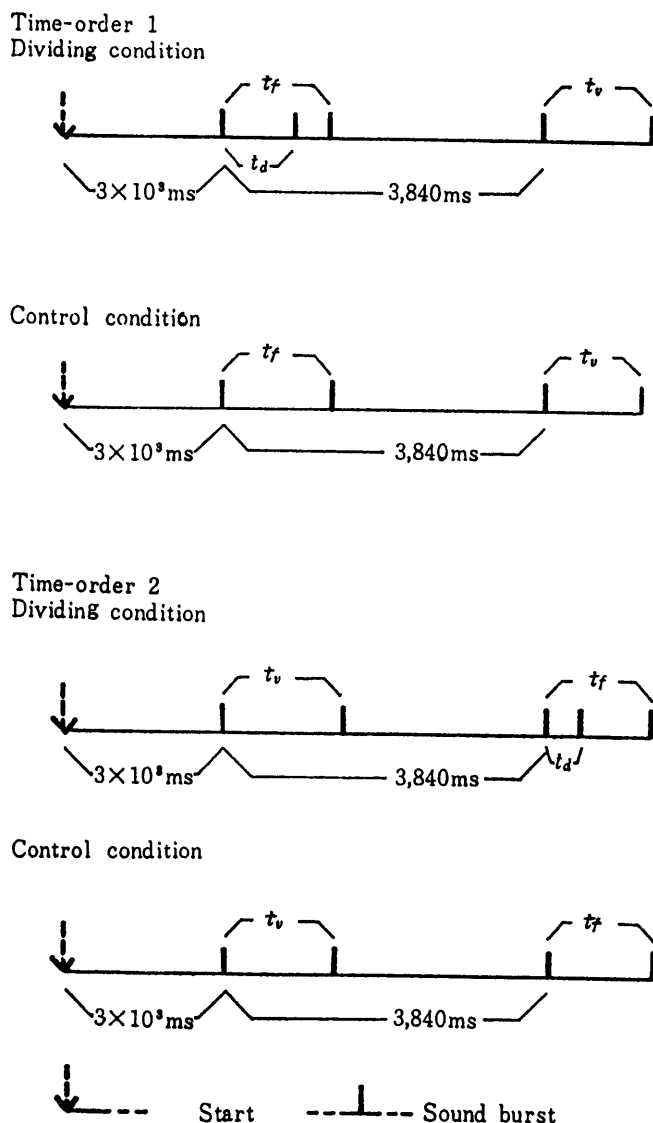


図 1 中島 (1979) の実験 1 の時間チャート。被験者がボタンを押すと、呈示が開始される。 t_f : 標準時間の物理的な時間長 (A C)、 t_d : 第 1 区切音の始まりから、分割音の始まりまでの、物理的な時間長 (A B)、すなわち、物理的な分割位置、 t_v : 比較時間の物理的な時間長 (D E)。被験者は t_v を調整することにより、標準時間と比較時間とが主観的に等しい時間長を持つようにする。

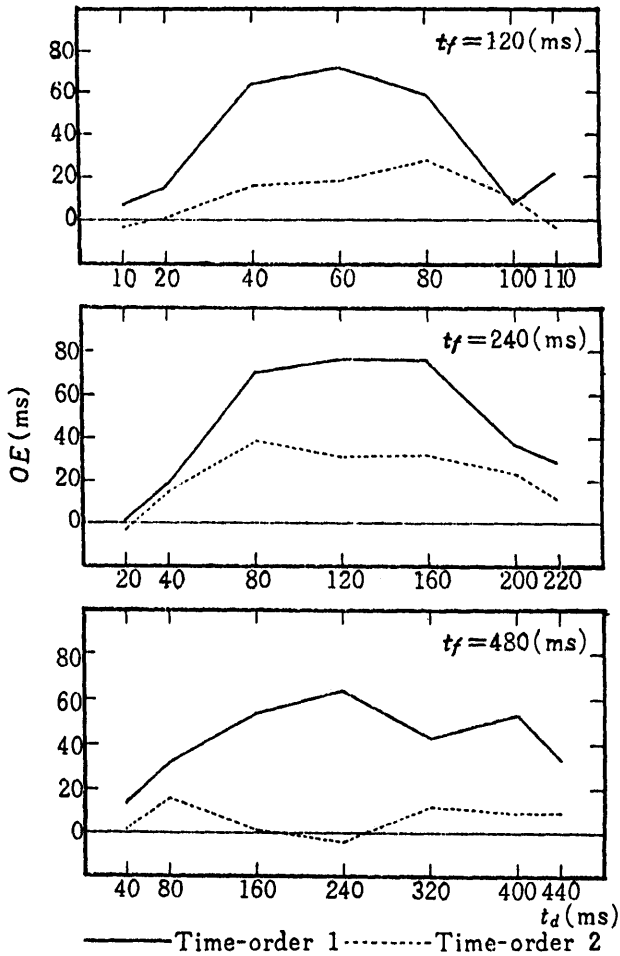


図2 中島 (1979) の実験1において分割音がもたらした過大評価。分割条件における主観的等価値から、対照条件における対応する (t_i が等しい場合における) 主観的等価値を減じた値を、過大評価 OE として示す。各点の値は、分割条件における該当する刺激パターンに関して、7名の被験者から得た14個の主観的等価値を算術平均することによって求めた。著しい時間順位効果が見出されたので、それぞれの時間順位におけるデータを別々に処理した。

第2時間順位において生じた過大評価の量が、第1時間順位において生じた過大評価の量に比べて、半分にも満たない場合が多く見られたので、どのような事情でこれほど大きな時間順位効果が現れたのかを調べるために、分割条件において、標準時間ごと、時間順位ごとに、主観的等価値の測定値を集め、その頻度分布を調べた(図3)。ただし、物理的な分割比が $4/12 \sim 8/12$ の範囲にある条件のみを、この分析の対象としている。第1時間順位における、主観的等価値の頻度分布には、明瞭に異なる2つの知覚のモードが示されている。過大評価量は、第1のモードにおいては約 20 ms であり、第2のモードにおいては約 100 ms で

ある。これらの2つのモードの中間の場合は認められなかった。第1のモードにおける約 20 ms の過大評価は、様々な実験的ノイズおよび、副次的な要因によって起こるとみなされた。それに対して、第2のモードにおいては、約 80 ms だけ過大評価が増しており、この分が標準時間の「主観的な分割」によって生ずる本来の錯覚現象に相当するとみなされた（この値は、あくまでも大まかな推測から得られたものである）。すなわち、分割時間の知覚は、視覚における反転図形の知覚のように、どちらかの状態が起こるが、中間的な状態が起こらない性質のものであると考えられる。主観的分割は、いずれの物理条件においても完全な確率で生ずるわけではない（毎回生ずるわけではない）ので、これまで多くの研究者が行ってきたように過大評価量の平均値に基づいて議論を進めることは、危険であることが判る。主観的分割が一旦生じたならば、少なくとも 50 ms よりも大きい過大評価を引き起こすように見受けられる。

この主観的等価値の頻度分布に関して、両時間順位を比較すると、第2時間順位の場合に、（約 100 ms の過大評価に相当する）第2のモードを示す分布の山が殆ど認められず、大部分の値が（約 20 ms の過大評価に相当する）第1のモードを示す分布の山に属していることが判る。平均値（図2）に見られる時間順位効果は、主として、2つの知覚のモードのうちいずれが生ずるかの、確率の相違によると考えられる。このような時間順位効果については、二つの原因が考えられる。変化するものが先立つために、調整作業の内容が変わることと、先に空虚時間が知覚されることによって、分割時間の知覚が変化することである。著者自身の内観によれば、空虚時間と分割時間とをこの順で聴いて比較する場合、分割時間をも空虚時間に似たものとして捉えやすく、分割音（B）を、時間間隔の両端を示す区切音（A、C）と知覚のうえで結びつけることが難しいようである。結局、第2時間順位においては、物理的に分割された標準時間が、空虚時間であるかのように聴こえることが多かったのではないかと推測される。

Adams (1977) の実験によると、分割音と、両端の区切音との知覚的な結びつきを弱めるような刺激操作によって、分割時間の平均的な過大評価量が減ずる。このことは今の考察を支持している。Adams の実験においては、1 s 前後の時間間隔が、500 Hz、20 ms の短音によって両端を区切られ、1個から5個の 1000 Hz、20 ms の短音によって分割された（短音の波形は不明である）。被験者は、このような分割時間の全体の長さを、カテゴリー判断によって評定し、分割音の数が評定値に与える影響を見ることにより、分割時間の過大評価の有無、度合が推定された。この手続きは、似たような条件において、分割音の数が多きほど過大評価の大きいことが、既に確認されていることに基づいている（Buffardi, 1971）

[Adams は、上述の分割時間から分割音を取り除いて対照条件を設けているが、対照条件における測定を、別の被験者群に対して行っており、このことが結果の解釈をやや不明瞭にしている。]。呈示される時間長と時間長とのあいだに、分割音と同じ短音の系列を挿入したり、分割音と、時間間隔の両端の区切音とをヘッドホンの異なる側から呈示したりすることにより、分割音と、両端の区切音との知覚的な結びつきを弱めることを試みたところ、いずれの場合にも、分割音の効果が減ずることが示された。

中島 (1979)、Adams (1977) の結果は、分割時間の過大評価が生ずるかどうかが、分割時間の置かれた文脈に影響されうることを示しており、この点については、今後、同一の被験者から大量のデータを得るような実験を行い、より広い視野から検討することが望まれる。いずれの実験結果についても、分割時間が知覚

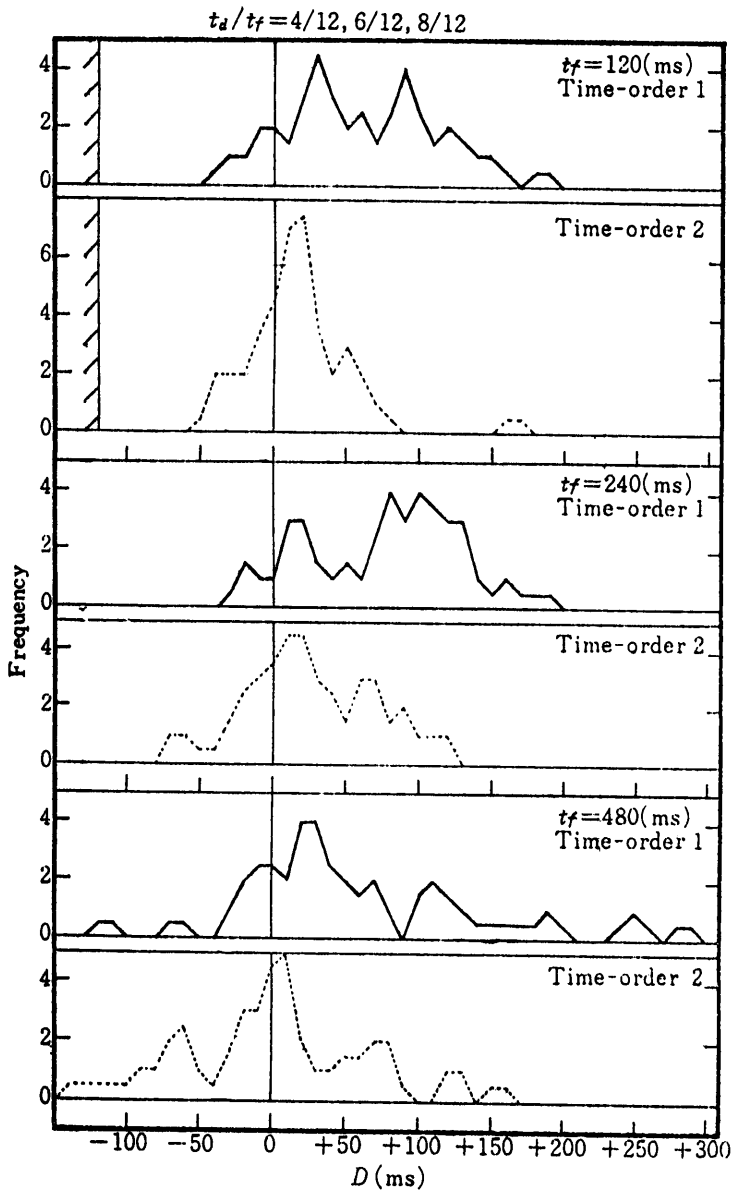


図3 中島(1979)の実験1の分割条件において得られた主観的等価値の、物理的等価値に対する超過量の頻度分布。それぞれの物理的等価値 t_i について、それぞれの時間順位ごとに、 $t_d/t_i = 4/12, 6/12, 8/12$ の場合における42個の主観的等価値を集め、そこから物理的等価値 t_i を減じた値 D について頻度分布図を描いた。ここでは、 D を、過大評価量とみなすことができる。範囲 $[x-10, x)$ における頻度と、範囲 $[x, x+10)$ における頻度とを平均し、点 x における頻度として示した。第1時間順位に関しては、分布の山が少なくとも2つあり、第2時間順位においては、そのうちの一方のみが明瞭に現れている。

される際に、分割音がその知覚内容から取り除かれたかのように振る舞う場合があると考えることによって、結果を解釈することができる。ten Hoopen (1996) は、このような実験結果について、注意の仕組みが、時間軸上の一部を抑制するフィルターのように働く例であると論じている。

ここでは、時間間隔として一つのまとまりをなす知覚内容に、分割音が含まれるか否かによって、時間間隔全体の主観的な長さが明確に変化することが、強く示唆されているが、このような考えかたの妥当性について検討するには、過大評価の生ずる具体的な仕組みまで考察する必要がある。中島 (1979) は、分割時間の全体の物理的な時間長を変化させても、分割時間の「主観的分割」から生ずる過大評価の絶対量（「パーセント」ではなく「ミリ秒」で示される過大評価）があまり変化しないことに着目した。そして、素朴な類推として、主観的に空虚な部分は、1枚の「包み紙」に包まれた食品のようなものであり、「包み紙」の分まで含めた「重さ」が、主観的な時間長に相当するとの着想を得た。すなわち、この1つの「包み」を、2つの「包み」に分割すると、同じ「包み紙」がもう1枚要るので、その分だけ全体が「重く」なり、過大評価が生ずると考えた。この「包み紙」が約 80 ms に相当すると考えるならば、主観的に空虚な部分を1つだけ含む時間間隔と、主観的に空虚な部分を2つ含む時間間隔とを比較すると、常に後者のほうが、約 80 ms に相当する分だけ過大評価されることになる。

このような類推の妥当性、発展性を検討するには、分割によってできた2つの「包み」の「重さ」の比率を見る必要がある。すなわち、分割時間をなす2つの空虚な部分の時間長が、どのような比率を持つように知覚されるかを確認する必要がある。このような目的で、物理量と主観量との関係を本格的に検討するには、精神物理学の歴史を踏まえた、体系的な実験、考察が必要であり、本論文の第2章、第3章の内容がこれに当たる。

中島 (1979) が上記の着想を発表した時点では、その実験1で用いた刺激パターンに的を絞って、時間長の比率の知覚に関する実験的検討を行っているので、簡単に紹介しておく。中島 (1979) の実験2は、分割時間の2つの空虚な部分 (A B および B C) が知覚されるときに、その時間長の比率がどのように知覚されるかを、調べるものであった。4名の被験者が、実験1において用いられた全ての分割時間に関して、その分割比を評定した。すなわち、分割時間の分割比に等しい比率が感ぜられるように、評定用紙の上で、60 mm あるいは 90 mm の視覚的な線分を分割した。(視覚的な線分の、物理的な分割比と主観的な分割比とは、ここでの目的に関する限り、等しいとみなして差し支えない。) その結果、分割時間の評定分割比 (第1の部分の占める割合で示したもの) が、対応する物理的な分割比よりも $1/2$ (1:1 の比率に相当する) に近づく傾向が見られた。分割時間全体が物理的により短いとき、このような物理的な分割比と評定上の分割比とのくい違いが、より著しくなった。この場合、評定上の分割比は、主観的な分割比の推定値であるから、この実験によって、主観的な時間長が、単純に物理的な時間長に正比例するものではないことが示されたことになる。

このような実験の結果を説明するために、上述の「包み紙」の類推を整理して、仮説の形にまとめた。これが、本研究の出発点となった〈つけくわえ仮説〉である。その内容を以下に記す：

- (i) 主観的に分割された (分割音などで分割されているように知覚される) 時間間隔の、全体の主観的な長さは、その (主観的に) 空虚な部分の主観的な

長さの総和に等しい。(つまり、時間長の全体が部分に分かれて知覚される
とき、全体は部分の総和として知覚される。)

(ii) 主観的に空虚である時間間隔(あるいは分割時間の部分)の主観的な長さは、その物理的な長さに約 80 ms の定数を加えたものに比例する。

(iii) 被験者が通常の状態を取る限り、物理的に空虚な時間間隔は、主観的にも空虚である。

この仮説の論理的枠組みは、Thomas と Brown (1974) の提出した<チャンクモデル>の枠組みとほぼ同じである。ただし、彼らは、定量的なモデルを提出するほど明瞭な根拠を示していない。第一に、<チャンク・モデル>そのものについて考えれば、分割時間の知覚において、主観的な分割が生ずるか否かの違いにより、少なくとも2つのモードが生じうると推測されるにもかかわらず、彼らのデータ処理法は、主観的な分割が殆どいつも同じように生ずるとの仮定から逃れていないように見受けられる。この仮定は、Adams (1977) によって疑問視され、中島 (1979) によって否定されているので、Thomas と Brown に従って何らかの考察を進めるには、その実験を追試することが不可欠であろう。第二の問題として、再生法(呈示された標準時間と等しい長さに感ぜられるような時間間隔を、キーを押すなどの動作によって作り出すことを、被験者に求めるような測定法。作り出された時間間隔を再生時間と呼ぶ。)の手續きにおいて、標準時間および再生時間の両方が空虚である場合にさえ、両者の物理的な時間長の間に比例関係を見出すことができなかつた点が挙げられる。このような、一見奇妙な結果に対して納得のゆく説明がつかない限り、その実験データに基づいて錯覚現象の仕組みについて考察することには無理がある。第三の問題として、<チャンク・モデル>の基本をなす「チャンク」の大きさ(「主観的に空虚な部分」の長さに相当する)がどのように知覚されるのかについて、分割比評定に相当するような実験の裏付けが必要ではないかと思われる。このように、<チャンク・モデル>は、無視しえない弱点を残している。とは言うものの、本研究との関連において、大変示唆に富むモデルであり、先駆的な着想として尊重したい。

1. 7 <つけくわえ仮説>の詳細

分割時間の過大評価が発見されて以来(Hall & Jastrow, 1886) ほぼ一世紀のあいだ、その生ずる仕組みについて、定量的な検討に耐えるような仮説、モデルが見出されなかつた。<つけくわえ仮説>は、大雑把ではあるが、この錯覚現象に関する定量的な仮説としては初めてのものである。この仮説を簡便に記述するために、以下のような数学的表現を用いる:

(1) 空虚時間の物理的な長さが t [ms] であるとき、その主観的な長さを、 $\tau(t)$ と表す。

(2) ある時間間隔が、いくつかの空虚な部分に主観的に分割され、これらの空虚な部分に相当する物理的な長さが t_1, \dots, t_n [ms] であるとき、その全体の主観的な長さを、 $\tau(t_1, \dots, t_n)$ と表す。

<つけくわえ仮説>は、次のように、記述される:

$$\tau(t) = k(t + \alpha), \quad \alpha \doteq 80 \text{ [ms]}; \quad k > 0 \quad (1)$$

$$\tau(t_1, \dots, t_n) = \tau(t_1) + \dots + \tau(t_n) \quad (2)$$

ここで、式 (1) の定数 k は、物理的次元を主観的次元に変換するためのものである。式 (1) のような心理尺度は、知覚実験の結果を説明するために相対的な値を示すだけであるから、任意の正の値を k に与えることができる。 α は、仮説の本質を表すくっつけくわえ定数であり、本論文においては、その値に関して種類の検討を加える。

主観的に2つの空虚な部分に分割された時間間隔の、一つひとつの部分の物理的な長さが t_1 [ms] および t_2 [ms] であるとする、これらの部分の主観的な長さは、式 (1) から $k(t_1 + \alpha)$ および $k(t_2 + \alpha)$ となる。式 (2) から、この時間間隔の全体の主観的な長さは、 $\tau(t_1, t_2) = k(t_1 + t_2 + 2\alpha)$ となる。これは、主観的に空虚であり、物理的な長さが $t_1 + t_2 + \alpha$ [ms] であるような時間間隔の主観的な長さ ($= k\{(t_1 + t_2 + \alpha) + \alpha\}$) に等しい。

$$\tau(t_1, t_2) = \tau(t_1 + t_2 + \alpha) \quad (3)$$

これは、分割時間の全体の長さに、約 80 ms の過大評価が生ずることを意味する。すなわち、分割時間の過大評価が説明された。

上記の分割時間において、2つの空虚部分の主観的な長さの比率は、 $\tau(t_1) : \tau(t_2) = (t_1 + \alpha) : (t_2 + \alpha)$ となり、対応する物理的な比率である $t_1 : t_2$ とは一般に一致しない。物理的な分割比 $R_o = t_1 / (t_1 + t_2)$ と、主観的な分割比 $R_s = \tau(t_1) / \tau(t_1, t_2)$ との関係は、

$$R_s - (1/2) = [(t_1 + t_2) / (t_1 + t_2 + 2\alpha)] \cdot [R_o - (1/2)] \quad (4)$$

となる。すなわち次のことが導かれる：

- (i) 主観的に空虚な2つの部分からなる分割時間の、主観的な分割比 R_s は、その物理的な分割比 R_o よりも $1/2$ に近くなる[ただし、 $R_o = 1/2$ のときには、式 (4) の右辺が0となるから、 $R_o = R_s = 1/2$ となる。]
- (ii) 分割時間の全体の物理的な長さ $t_1 + t_2$ [ms] が短いほど、(i)の傾向は著しい。

中島 (1979) の実験2は、確かにこのような傾向を示している。この予測に関する、本格的な検証は、本研究の基本をなす話題である。

1. 8 本研究の目的

時間知覚についての様々な実験結果を説明しうるような、一般的な理論を導く第一歩として、再現性の高いデータに基づいて、可能なかぎり単純な心理尺度を構成し、神経生理学的な考察とも結びつけうるようなモデルを構成することを試みる。また、このような心理尺度を適用することが可能な条件、範囲についても考察する。

ここでは、議論が発散することを避けるために、次のように限定された実験状

況を、主な考察の対象とする。

- (i) 被験者が、ごく短い短音によって区切られた、2つないし3つの、空虚な（あるいは擬似的に空虚な）時間間隔の関係について判断する。
- (ii) 空虚な（あるいは擬似的に空虚な）時間間隔の大半が、40～600 ms の範囲に収まる。

本研究の理論的な出発点は、中島（1979）の提出した〈つけくわえ仮説〉である。この仮説は、「主観的に空虚である時間間隔の主観的な長さは、その物理的な長さに約 80 ms の定数を加えたものに、比例する。」と言うものである。発表当時、これは純然たる作業仮説であり、精神物理学的研究に不可欠な、種々の数学的関数の当てはまりを比較するような作業も行われていなかったので、本研究において、本格的にこの仮説の妥当性、発展性を検討することにした。

本論文では、様々な状況において〈つけくわえ仮説〉の妥当性を吟味し、この仮説が、何らかの時間知覚の仕組みに結びつきうるものかどうかを検討する。具体的には、比率判断、リズム産出、弁別判断を取り上げる。特に比率判断に関しては、大量のデータに基づいた定量的な分析を行い、心理尺度の候補となりうるいくつかの式に関して、その得失を比較する。単純な仮説を、ある程度広い範囲に適用しうるならば、時間知覚の仕組みを理解するための、理論的な枠組みとなりうるであろう。

1. 9 本論文の構成

第1章においては、聴覚におけるリズムの役割に関して、文献調査に基づき、問題点を整理した。以降、第2章において、線分尺度を用いた空虚時間の比率判断の実験について報告し、続いて、第3章において、数を用いた比率判断の実験について報告する。数を用いた比率判断に関しては、充分多くのデータを得ることができたので、心理尺度を簡単な数式によって近似することが妥当であるかどうかを、様々な角度から検討する。以上の実験により、1秒程度までの空虚時間に関して、その主観的な長さが、物理的な長さに約 80 ms を加えたものに比例するとの仮説、すなわち〈つけくわえ仮説〉を検証する。第4章では、音楽のリズムに関連した考察を行う。時間長の比率判断は、西洋音楽における音符の相対的な長さに関連づけることが容易であるので、楽譜に記された簡単なリズムの演奏、ないし調整による産出を求めるような実験について報告し、仮説として提出した心理尺度の当てはまりを検討する。また、数を用いて指定された比率を、調整によって産出することを求めるような条件において得られたデータについても、比較のために報告する。精神物理学の理論について検討する際には、弁別実験の結果についても考慮することが望ましいので、第5章では、空虚時間の弁別判断に関する過去のデータを集めて再分析し、本研究の文脈に即して検討する。上記の仮説が、ここで取り上げるような様々な実験状況において、包括的な説明を与えるものであることが判ったので、第6章において、本論文において検討した実験結果を全体として説明しうるような時間知覚の仕組みがありうるかどうかを検討する。第7章では、結論を述べる。