

## 隣接する二つの時間間隔の知覚に関する研究

宮内, 良太

<https://doi.org/10.15017/458891>

---

出版情報 : Kyushu University, 2004, 博士 (芸術工学), 課程博士  
バージョン :  
権利関係 :

# 第1章 序論

## 1.1 はじめに

時間は、日常的に常にその存在を感じる事が出来るものである。しかし、時間が実際に存在していることを証明し、時間とはなんであるのかを説明することは困難である。時間は、人間が外界から受ける様々な刺激の関係のうちに見いだしてきた次元である。色や形、音の高さや音色といった刺激の主観量には、それらの主観量に対応する物理量を知覚する感覚様相が決まっている。対して、時間のみを知覚する感覚様相というものは存在しない。時間知覚は、視覚や聴覚などの感覚様相が事象を知覚した後に行う情報処理の過程で成立するものであり、あらゆる感覚様相と密接な関連を持っているのである。しかし、逆に、ある感覚様相における時間知覚についての研究で得られた結果と他の感覚様相で得られた結果とを比較検討することで、感覚様相間に共通した、あるいは、感覚様相に固有の時間知覚のメカニズムを見つけることが出来る可能性を秘めているとも考えられる。

本研究では、時間知覚に有利な感覚様相である聴覚に注目し、音事象間の時間間隔がどのように知覚されるのかについて基礎的な研究を行った。聴覚は、視覚や振動感覚などと比べて時間知覚における精度が高く (Grondin, 1993; Grondin & Rousseau, 1991)、音刺激を用いることで、刺激間の時間間隔をより細かく制御したパターンに対する主観的な時間を測定することが出来る。まずは聴覚刺激における時間知覚の重要性について概観する。

人間は、他の動物と比べて、他者と聴覚コミュニケーションをとる能力に長けている。もちろん、相手の表情やしぐさといった視覚的な情報も、コミュニケーションをとる上で重要な役割を果たしているが、音声や音楽を用いた聴覚コミュニケーションによって、情報をやり取りする比重が非常に大きいといえる。手軽に情報を交換する手段として、電話やラジオといった聴覚情報のみを伝搬する機器は一般的に広く普及しているが、視覚情報のみでリアルタイムに情報をやり取りする機器があまり存在しないことから、聴覚コミュニケーションが重要であると考えられる。

聴覚コミュニケーションの主な道具である音声や音楽を構成しているのは、時間軸に沿って、現れては消えていく複数の音事象である。しかし、この音声や音楽を構成している一つ一つの音事象そのものに、伝達したい情報のすべてが含まれているわけではない。

重要な情報は、時間上に付置する音事象間の時間構造や、音事象の時間的变化に含まれていることが多い。このような音列パターンの時間的な秩序を表す言葉として、“リズム”という言葉がよく用いられる。

音楽において、リズムは、旋律や和音と並んで重要なものである。音符の長さが、音楽の基本的な枠組みを形成していることを考えれば、リズム構造が、旋律や和音よりも音楽の根底をなす要素といえよう。組織的な音高の変化を持たない打楽器のリズムのみで音楽を奏することはできるが、リズムのない音楽、音高の変化のみで構成された音楽などはありえない。音楽の知覚を解明する上で、リズム知覚の研究が、重要な手がかりの一つとなりうる。

話し言葉にも、リズムが存在する。英語は、一般的に強勢主体のリズム構造を持っている。一方、日本語は、モーラを主体とするリズム構造を持っている。例えば、日本語では、隣り合う単語の時間構造によって、長音・撥音・促音が補完されることがある。「さか」という言葉を発話する際に、「さ」と「か」の間の時間間隔を延ばすだけで、「さっか」や「さんか」といった異なる意味を持つ単語に知覚される場合がある。音声知覚の研究においても、音事象間の時間構造がどのように知覚されているのかを解明することが、重要な課題である。

我々は、次々と呈示される隣接した時間間隔を、知覚的に体制化することでリズムを感じている。よって、リズムがどのように知覚されているのかを解明するためには、リズムを構成する時間間隔がどのように知覚されるのかを調べることも重要である。また、一つ一つの時間間隔の知覚について調べるだけでなく、時間間隔を隣接させたときに、それぞれの時間間隔がどのように知覚されるのかを確かめる必要がある。

このように、精神物理学的な実験によって、音事象群から情報を取り出すための重要な手がかりの一つであるリズムやリズムを構成する時間間隔が、どのように知覚されているのかを体系的に調べ、刺激パターンに含まれる時間情報の処理過程について考察することは、音楽や音声の知覚の研究にもつながる意義のある研究であるといえる。

末富・中島(1998b)は、聴覚におけるリズムの知覚についての古典的な研究から最新の研究までを概観し、知覚心理的な立場から、聴覚におけるリズムの特徴についてよく整理している。彼らは、リズムを時間上のゲシュタルトと考え、聴覚実験によって得られた結果と空間上のゲシュタルト心理学における知見とを比較することで、時間軸上の事象の群化、および階層化が、リズムの形成に重要であることを示している。以後、第1章では、彼らの論文と重複する部分もあるが、時間上のゲシュタルトに固有の特徴である時間的規則性の

原理と分割時間の知覚に特に注目して、これまでに行われたリズム知覚研究や時間知覚研究を紹介し、本論文の目的を明確にする。

## 1.2 リズムについて

まず、リズムという言葉の持つ意味について考えてみる。日常的に用いられる”リズム”という言葉は、様々な状況で使われる言葉であり、前後の文脈や使われた状況によって言葉の意味や性質が異なってくる。音楽的なリズムに限った場合でさえも、次に示すいくつかの文面では、リズムという言葉の持つ性質が変わってくる。

- a. この曲のリズムは、4拍子である。
- b. この曲のリズムは、大きく変化する。
- c. この曲のリズムは、はげしい感じがする。
- d. この曲のリズムは、よい。

各文章中に現れるリズムという言葉は、aの例では、曲の構造や構成のみを表しており、bでは、曲全体の大きな流れや動きといったものを表している。cになると、音列全体から受ける印象や感情を想起させる対象となっており、dでは、優劣を評価できる価値をも含んでいる。Gabrielsson (1982) は、例として示したaからcまでを、それぞれ”リズムの構造”、”リズムの動き”、”リズムの感情面”と定義し、リズムには様々な側面があることを示した。広辞苑第五版(岩波書店, 1998)を見ると、一般的に用いられるリズムという言葉の意味は、次のように定義されている。

1. 律動。進行の調子。
2. 詩の韻律。
3. 音楽におけるあらゆる時間的な諸関係。

広辞苑第四版では、3の部分で、「音楽で、音の強弱が周期的に繰り返される構造。節奏。拍子。」となっており、改訂に際して、さらにリズムという言葉の適用範囲が広げられたことが分かる。また、認知科学事典(共立出版, 2002)では、次のように定義されている。

旋律に対しリズムという場合には、音の長短(音長)関係に基づくグルーピングを意味する。しかし、音高や和声、拍節(ビート)もグルーピングに関与することから、広く音のグループの様相をリズムとしてとらえる見方が一般的になりつつある。

このように、非常に定義の曖昧なリズムを実験的に取り扱う際には、リズムのどの側面に注目して研究を行うかをはっきりとさせなければ、議論が発散し有益な結論を得ることはできない。

Fraisse (1982) は、知覚心理学におけるリズムを、“続いて起こる事象の秩序”と定義した。時間軸に沿って現れては消えていく複数の音事象から情報を得るためには、複数の音事象を一つにまとめるための明確な秩序を時間軸上に形成する必要がある。リズム知覚について研究するという事は、次々に呈示される一つ一つの音事象そのものの知覚について研究するのではなく、聴覚システムが、時間軸上にどのような秩序を形成して複数の音事象を体制化しているのかを調べることである。そこで、今後はリズムを、“時間軸上に知覚される形、あるいは形の枠組み”の事象と捉え、複数の音事象によって作られる音列パターンの時間的な枠組みを、聴覚システムがどのように形成しているのかについて検討する。

### 1.3 リズムの生じる時間間隔

リズムの知覚について考えるためには、まず、リズムを構成する個々の時間間隔の知覚についてまとめておく必要がある。Fraisse (1978, 1982) は、時間間隔には、リズムの生じうる限界が存在することを、様々な聴覚実験の結果を概観することで示している。音事象間の時間間隔が、20 ms よりも短いならば、二つの音を分離して聞くことは難しく (Hirsh, 1974)、100 ms よりも短い時間間隔で次々に呈示される音を、一つ一つ区別して知覚することは難しい。逆に、音事象間の時間間隔が、1800 ms を超えると、二つの音を一つのパターンとして結び付けることが難しくなる (Nakajima, Shimojo, & Sugita, 1980)。さらに、5秒以上の音列パターンを一つのグループとして群化することは難しい。Fraisse は、この範囲が“心理的現在 (psychological present)”に対応する時間間隔であるとした。心理的現在とは、一連の事象を一つのまとまりとして知覚できる範囲を表しており、音事象間の統合が起こる限界に対応している。これは、聴覚的感覚記憶 (auditory sensory memory) もしくは、音響貯蔵庫 (echoic store) の長さとも対応している。このような貯蔵庫の長さは、約 2 s から 5 s の間であると報告されている (例えば、Treisman, 1964)。これは、時間間隔の知覚が記憶と密接な関係を持っていることを示している。

実験室的な聴覚実験の結果だけでなく、実際の音楽において演奏される音と音との時間間隔や、フレーズの長さを分析した結果から、音楽に頻出する時間間隔の長さには、一定

の傾向があることが分かっている。Fraisse (1982) は、クラシック音楽について、楽譜に含まれる音符の長さを分析した結果から、楽曲のリズムは、主に 150-290 ms 程度の長さを持つ短い音符と、300-900 ms 程度の長さを持つ長い音符とで、構成されていることを報告した。また、Moelants (2002, 2003) は、現代音楽について分析し、多くの音楽のテンポが 128 bpm 付近に集まっていることを報告している。bpm は、音楽のテンポを表すときによく用いられる単位であり、1 分間に基本となる音符がいくつ鳴らされるかを表している。よって、現代音楽では、基本となる音符の長さが、 $60000 \text{ [ms]} / 128 \text{ [bpm]} = 468.75 \text{ [ms]}$  となる。リズムは、多くの場合、この基本となる時間間隔 (一拍) を 2 つ、もしくは、3 つに分割することで作られており (1.6.2 を参照)、現代音楽を構成する音符の多くが 470 ms 以下の時価を持っていると考えられる。

本研究の目的は、聴覚システムが時間間隔を知覚的に体制化する基本的な仕組みを探ることである。よって、二つの音事象間の時間間隔を知覚できないような短い時間間隔を用いることは適切ではない。また、二つの音事象間の時間間隔を一つのパターンとして知覚できないような長い時間間隔を用いることも本研究の目的と合致しない。そこで、二つの音事象を分離して聞くことができる、約 30 ms から、二つの音事象を知覚的に結び付けることができる、約 2000 ms までの範囲に含まれる時間間隔を用いた研究を中心に概観していく。

#### 1.4 単音系列の知覚的な体制化

音楽や音声に関するリズム知覚研究について述べる前に、まずは単純な短い音で構成された音列パターンがどのように知覚的に体制化されているのかについてまとめる。後で詳しく述べるが、知覚されるリズムは、アクセントの位置や旋律の構造によっても影響を受ける。そこで、このような研究では、一つの周波数成分しか持たない純音やサウンドバーストといった、単純な物理的性質を持つ音で構成された音列パターンを用いることで音の高さや音の大きさの要因を出来るだけ除外し、音の始まりから次の音の始まりまでの時間構造のみに注目している。このような単純な音系列が、どのように知覚的に体制化されているかを知ることで、リズムの最も基本的な構造を示すことが出来る。

#### 1.4.1 時間の格子

複数の音列から、時間的な枠組みを知覚的に形成するためには、時間軸上になんらかの枠組みを設ける必要がある。Povel (1984) は、規則的な音列を知覚的に体制化するための一つのモデルとして、“時間の格子 (temporal grid)” という枠組みを用いて音列を体制化しているという考え方を提唱した。時間の格子とは、時間を等間隔に区切った目盛りのようなものである。これは、長さを測るためのものさしや、重さを量るための分銅のような物と考えれば分かりやすい。このモデルでは、次のような過程で、音列に含まれる時間間隔の規則性を検出し、その規則性をもとにして音列パターンを知覚的に体制化していると考えている。

音列に含まれる規則性を検知するために、我々の聴覚システムは、まず音事象によって区切られた複数の時間間隔から等しい長さを持つ時間間隔をいくつか見つけ出す。次に、その中から、音列の構造を最も簡潔に説明する長さを時間の格子として選択する (詳しくは、図 1 を参照)。ほとんどの場合、音楽におけるリズムは、小節や拍子という等間隔の区切りによって成り立っていることを考えれば、時間の格子は小節や拍子の長さに相当するものであると考えられる。ただし、厳密には、時間の格子と小節や拍子の区切りとが全く同じ場所に存在しない場合もある。時間の格子を知覚上に形成した後は、次々と呈示される音事象間の時間間隔と選択された格子の長さとを比較し、格子の長さよりも短い時間間隔が呈示された場合には、時間の格子を分割した時間間隔として知覚される。その際に知覚される分割比は、1:1 や 1:2 といった単純な整数比になりやすい (Povel, 1981)。

Essens and Povel (1985) は、次のような時間構造を持つ音列を繰り返し実験参加者に呈示したあと、タッピングによって音列パターンを何度か繰り返して再生させる実験を行った。

- a. /200 200 **400**/200 400 200/800/800/
- b. /200 200 **500**/200 500 200/900/900/
- c. /400 200 200/**600** 200/**600** 200/800/
- d. /500 200 200/**700** 200/**700** 200/900/

(単位はすべて ms である。/ は、音列から形成される格子 (clock unit) を表す。注目すべき分割パターンを太字で表す。物理的な分割比は、a が 2:1、b が 2.5:1、c が 3:1、d が 3.5:1 である。)

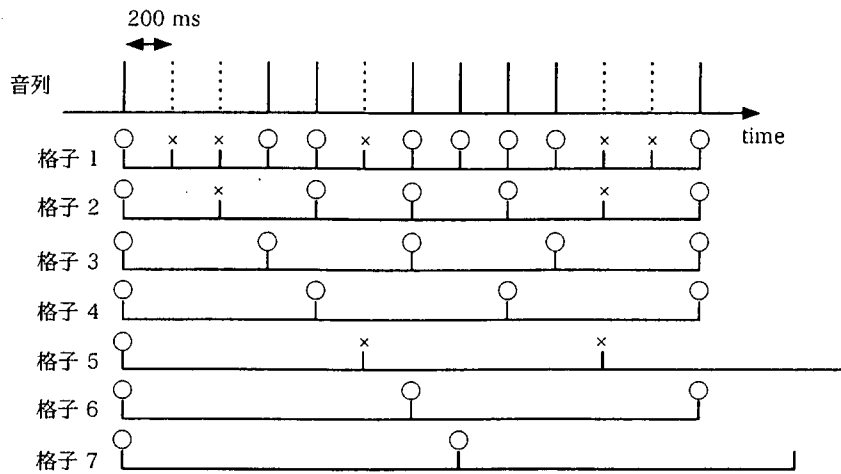


図1 時間の格子の形成過程。600 200 400 200 200 200 600 (ms) の時間間隔で、音列が呈示されたとする。この音列パターンには、いくつかの格子を当てはめることが出来るが、この中から、最も簡潔で効率よく刺激パターンに当てはまる格子が、知覚されやすい。格子1から格子7は、音列に当てはめうる格子である。格子の区切り(格子点)に表示した○は、その格子点に音が呈示されていることを表している。×は、その格子点に音が呈示されていないことを表している。この中から、音列に最も適した格子を選択する場合には、まず、格子点に音がどれくらい存在しているのかを判断する。よって、すべての格子点に音が存在している、格子3, 格子4, 格子6が、形成されやすい。次に、格子間に存在する音列の形が簡潔なものの方が、当てはまりがよいと判断する。格子3では、格子点と格子点との間に音が呈示されていない場合や、等間隔に分割される場合が多い。一方、格子4と格子6では、一つの格子を、二つの異なる長さを持つ時間間隔で分割している場合があり、格子3よりも格子間が複雑なパターンとなっている。よって、この音列からは、格子3が最も形成されやすい(Povel, 1984)。

時間の格子モデルを用いてこれらの音列パターンを分析すると、aとcでは、800 msの時価を、bとdでは、900 msの時価を持つ時間の格子が形成される。そして、これらの音列パターンは、この時間の格子を2つ、もしくは3つに分割した音列パターンとして知覚される。実験の結果を見ると、分割された時間間隔の間の比率が単純な整数比となるパターン(aとc)では、正確に音列が再生された。しかし、時間の格子の分割比が、2.5:1や3.5:1といった整数比では表現できないパターン(bとd)では、再生された音列の分割比が、2:1や3:1に近づいて再生されていた。この実験結果も、時間の格子を分割する際には、単純な整数比として分割されやすいことを示している。



しかし、このような時間の格子は、音列のパターンやアクセントの位置などによって可変なものであり (Essens & Povel, 1985; Povel & Okkerman, 1981; Povel & Essens, 1985)、時間の格子がどのようにして知覚上で形成されるのかについて、すべての実験結果を説明しうる体系的な理論は、まだ報告されていない。Povel (1984) は、どのような格子が形成されるかは、“音列の始まり”・“主観的なアクセントの位置”・“テンポ”の3つの要素と密接な関連があり、これらの要素を含んだ体系的な研究を行う必要があると述べている。

#### 1.4.2 階層化

Povel (1984) の提唱した時間の格子の生成過程では、音列パターンを聞いた後に、その音列中に含まれる複数の時間間隔を分析して時間の格子を選択する必要がある。つまり、複数個の時間間隔を含む音列パターンを知覚した後でないと格子を形成できないといえる。しかし、我々が音列パターンを知覚する際には、最初の数音を聞いただけで、なんらかのリズムを感じる事が出来る。さらに、その後の音列に含まれる時間間隔によって、リズムの構造を新たな形に修正したり、形成された格子をより強化したりすることがある。Povel (1984) のモデルは、聴覚システムがリアルタイムで時間の枠組みを形成し、それを随時修正している現象をうまく表現できない。

Longuet-Higgins and Lee (1982) や Lee (1985) は、単音系列のリズムを、音の相対的な長さから階層的にまとめあげていく過程を、音符の長さを用いて説明するモデルを提案した。彼らのモデルでは、次のようにして、隣接する時間間隔を階層的に群化していく (図 2)。まず、系列の最初の2音を聞いた段階で、聴覚システムは、その2音で作られた時間間隔の長さを音列を体制化するために用いる枠組みの長さとして仮定する。さらに、続いて呈示される音の系列を聞いて、その枠組みを修正したり、補強したりする。例えば、図 2 (a) に示したような四分音符の系列を聞いた場合には、まず、最初の音の始まりから二つ目の音の始まりまでの時間間隔の長さを、格子の長さであると仮定する。次に、その後に隣接する時間間隔の長さが、仮定した格子の長さと同じであるならば、格子の長さの2倍の長さを持つ上位の格子を仮定する [図 2 (a)]。つまり、次の階層に知覚的なまとまりを移行させる (彼らは、この行程を “CONFLATE” と呼んでいる)。もし、三つ目の音が、想定した場所よりも前に呈示されたならば、一つ目の音から三つ目の音までの時間間隔を、格子の長さに再仮定し (“STRETCH”)、その格子の長さとして次の時間間隔の長さを比較する [図 2 (b)]。逆に、三つ目の音が、想定した場所になっても呈示されなかったならば、格子の長さの始まる音の一つずらして (“UPDATE”)、その格子の長さとして次の時間間隔の長さ

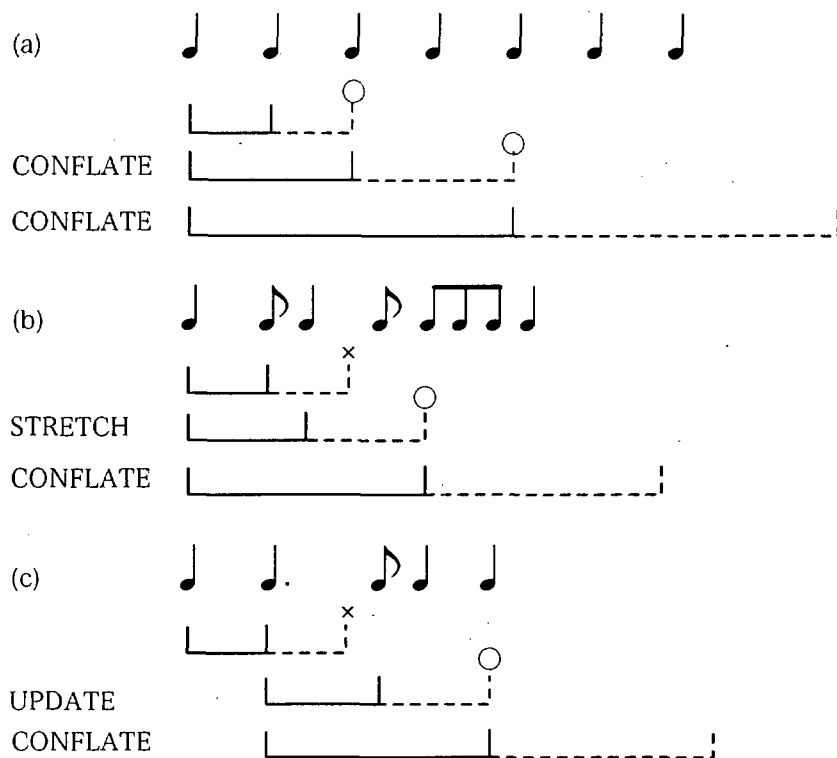


図2 Longuet-Higgins and Lee (1982) が提唱した、単音系列のリズムを階層的に群化していく過程のモデル。

とを比較する [図 2 (c)]。彼らは、次々と呈示される音列に対してこのような処理を繰り返すことで、音列の時間間隔が階層的に群化されていくとした。知覚上に形成された格子の長さをもとにして音列パターンを体制化するという基本的な考え方は、Povel (1984) の提唱した時間の格子を用いたモデルと同じである。しかし、このモデルでは、即時的に音列パターンが体制化されていることをうまく説明することが出来る。また、時間の格子説では、時間の格子で作られる枠組みが、最上位の知覚的な群を形成しており、同じ長さの格子が階層的にまとまって上位の群が形成される過程については、論じられていない。後で述べるが、音楽においては、同じ長さの格子が次々と群化して階層構造をとることが分かっており、このモデルは、音楽におけるリズムの階層性との対応がよい。ただし、彼らのモデルは、楽譜に記譜された音符間の長さのみを機械的に処理することによって組み立てられており、音楽におけるリズムの形成過程をすべて説明しうるものではない。実際に、演奏家によって演奏された楽曲に含まれる、ある音の始まりから次の音の始まりまで

の長さを測定すると、記譜された音符間の長さとは異なる長さで演奏されている場合が多い (Gabrielsson, 1985, 1974; Gabrielsson, Bengtsson, & Gabrielsson, 1983)。しかし、そのような場合でも、我々は、楽譜の表す拍節構造をうまく捉えることが出来る場合が多い。また、演奏するテンポを変化させると、全く同じ楽譜であるにもかかわらず、演奏された音符の長さの比率が変化するという報告もある (Repp, Windsor, & Desain, 2002)。機械的に楽譜の長さから、曲の構造を解析する Longuet-Higgins and Lee (1982) らのモデルでは、このような現象をうまく説明できない。また、実際の音楽では、旋律の変化やアクセントの位置といった要素が、リズムの形成に大きな影響力を持っている。彼らのモデルは、楽譜をもとにモデルを構築しているが、音符の長さのみに注目しており、その他の音楽的な要素を加味していない。これは、単音系列のリズムについてのモデルである。

まずは、単純な短い音を用いた音列パターンが、どのように知覚的に体制化されているのかについての研究について述べてきた。このような研究は、聴覚システムが音列をリズムとして体制化するメカニズムを解明する基礎となりうる。しかし、我々が普段耳にする音楽や音声は、同じ性質を持った音の羅列ではない。それらは、複雑な周波数構造を持っており、個々の音の強さも変化している。また、音楽や音声に特有のリズムというものもあるはずである。そこで、次節からは、音声や音楽のリズムがどのように形成されているのかを扱った研究を紹介する。

## 1.5 音声におけるリズム

音声知覚の研究では、主に、フォルマントの遷移や基本周波数の変化といった、周波数的な分析に重点が置かれる傾向がある。しかし、音声においても、全体の時間構造がどのように知覚されているのかを知ることが重要である。ここでは、英語と日本語のリズムを例として、音声のリズムの特徴について述べる。

### 1.5.1 英語のリズム

英語のリズム構造については、Handel (1987) によって詳しくまとめられている。音声を構成する個々の音には、音の大きさ、長さ、高さがあり、これらが他の音と比べて異なっている場合には、その音が強調される。Handel (1987) は、そのような強調された音を強勢と定義した。そして、強勢が置かれた音 (強拍) と、強勢の置かれていない音 (弱拍) との配列が、英語のリズムの基本となると考えた。

英語のリズム構造を説明するために、まず Handel (1987) は、一つの単語に出現する強拍と弱拍の構造を言語的に分析し、音節が強勢によって階層的に結びついていることを示した。次に、文を発話する際に生じる規則的な強勢のリズムの成立過程が、同じような方法で解析できることを示した。この英語を発話する際に生じる規則的な強勢のリズムを、言語における「拍節の格子 (metrical grid)」と定義した。

図3に示した「a hundred thirteen men」という文を例として、英語のリズムの生成過程を説明する。まず、文章を音節ごとに分け、それぞれの音節の下に一つ丸印(●)をつける[図3(a)]。これが、リズムの最も小さな階層である。次に、言語的な階層性の法則に則って、強拍と弱拍の結びつきを表す階層木を描く。基本的には、隣り合う強拍と弱拍とが結びついて一つ上の階層を構成する。このときに、強拍をS、弱拍をWと表し、強拍から上に伸びる枝を太い線で表す。また、音節に繋がる太い枝の数だけ、音節の下に丸印を増やしていく[図3(b)]。この丸印の数が多い音節は、他の音節と比べて、強勢が相対的に強くなることを表している。図3(b)は、言語的な処理によってこの文章を分析した結果の階層木を表している。これを見ると、強勢を相対的に強くすべき、つまり丸の数が多い「teen」と「men」とが隣接しており、「強勢の衝突」が生じている。文章を発話する際には、強勢の衝突を嫌う傾向があるため、このように強勢が衝突した場合には、その衝突を緩和する方向に強勢の入れ替わりが生じる。この例では、「men」に繋がる枝は一本である。つまり、途中で分岐していない。よって、「men」の相対的な強さを小さくすることは出来ない。衝突を緩和するためには、「teen」の相対的な大きさを小さくする必要がある。「teen」に繋がる枝を上位の階層から見えていくと、まず、「a」と「hun dred thir teen」とに分岐している。ここで強勢を入れ替えることが出来れば、「teen」の強勢が小さくなる。しかし、「a」に繋がる枝はすべて弱拍に繋がっており、途中で急に強拍に転じることはできない。次に、「hun dred」と「thir teen」との分岐を見ると、「hun dred」の方には、「hun」に繋がる太い枝が存在する。よって、ここで強勢が入れ替わることになる[図3(c)]。一回の入れ替えで強勢の衝突がなくなる場合は、さらに強勢の入れ替えが生じる[図3(d)]。最終的に、強勢は、「hun」と「thir」と「men」の位置に置かれ、規則的に強勢が並ぶリズムが出来上がる。

このように、英語では、強勢と強勢との間が等間隔になるような拍節の格子が形成されやすい。よって、英語の音声を発話する際には、時間の格子に当てはまるようなリズムの形成が、拍節の格子によって促進されるといえる。

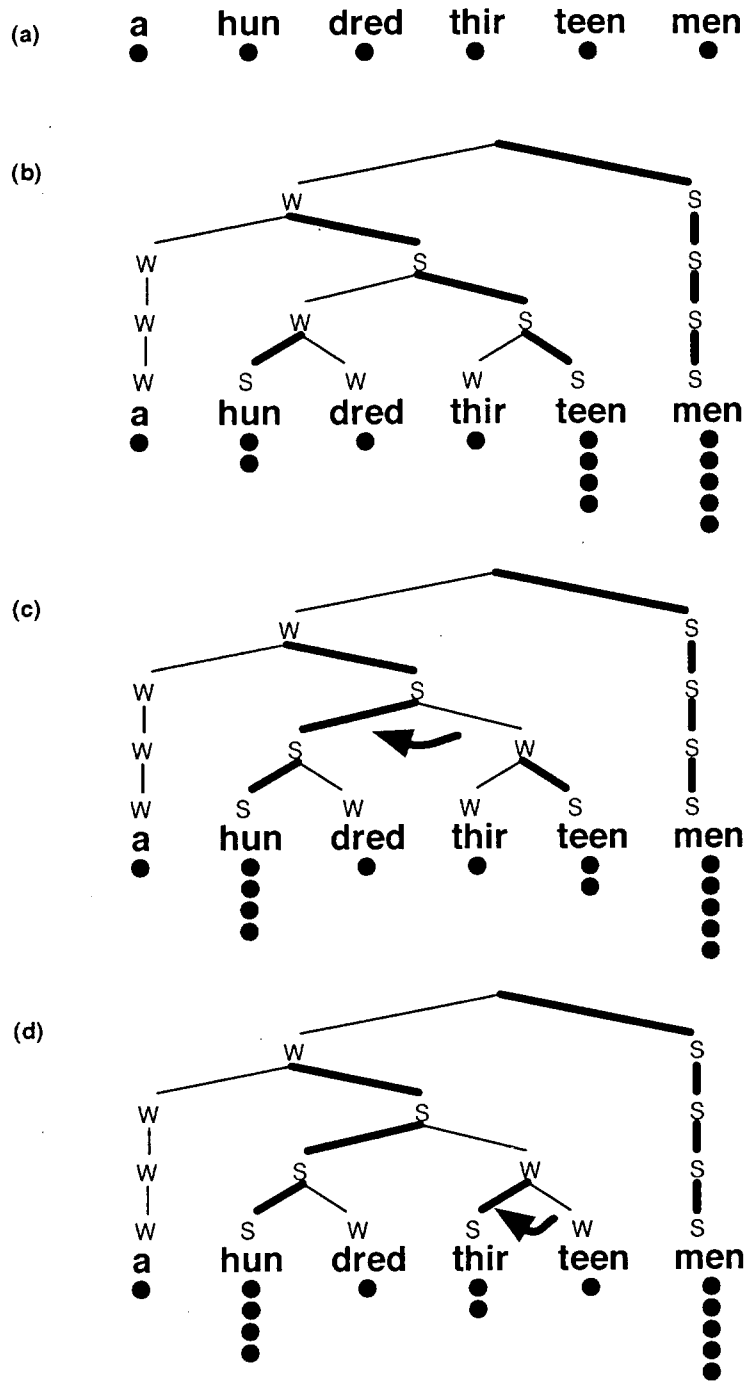


図3 音声のリズムの生成過程 (Handel, 1987)。“a hundred thirteen men” という文章から、音節の階層性と、強勢の衝突による強勢の入れ替えをもとにして、強勢が等間隔に並んだ拍節の格子が形成される過程を図示した。

## 1.5.2 日本語のリズム

一方、日本語は、“モーラ”を一つの単位としたモーラ基準のリズムを持っている。1モーラは、短母音を含む1音節の長さに相当し、ほとんどの場合、仮名一文字が1モーラとなる。1モーラの長さは、話者の発話速度や話す状況などによって90から250msくらいの間で変化するが、160msくらいが典型的な値である (Port, Dalby, & O'Dell, 1987)。このモーラの長さで、連続的な音声を次々と区切り、それぞれのモーラに離散的な仮名文字を当てはめることで音声から言語情報を抽出している。よって、音声のモーラ構造が、言語情報の伝達に重要な役割を果たしているといえる。

日本語では、母音や子音の持続時間を延長する、つまり音声の時間構造のみを変化させるだけで、新たな音素が加えられたときに生ずると同様の意味の違いが生ずる場合がある (Port et al., 1987; 藤崎・杉藤, 1977)。Fujisaki, Nakamura, and Imoto (1975) は、「居た /ita/」に含まれる無声破裂音 /t/ の直前の無声部分の長さを変化させると、/i/ の終わりから /a/ の始まりまでの時間間隔が169msの時に、「行った /itta/」と無声部分に促音が知覚されることを示した (図4)。この他にも、先行する単語の持続時間を変化させることで、「尼 /ama/」が「按摩 /amma/」に、「甥 /oi/」が「多い /o:i/」に、「伊勢 /ise/」が「一畝 /isse/」に変化する、つまり、音の持続時間を延長するだけで、物理的には存在しない撥音・促音・長音といった特殊拍が知覚されることを示した。さらに、前後の母音の長さを長くすると、ほぼそれに比例して、特殊拍が知覚されるために必要な時間間隔も長くなることが分かった。これは、全体のモーラ構造が変化することで、特殊拍が知覚される境界も変化することを表しており、モーラの構造が言語情報の伝達に重要である確かな証拠といえる。

## 1.6 音楽におけるリズム

次に音楽におけるリズムを扱った研究について概観する。リズムは、音楽を構成する要素の中で重要なものである。当然のことながら、心理学者が音楽のリズムを体系的に研究する以前から、音楽を創造する音楽家や演奏家は、音楽のリズムに注目してきた。彼らは、様々な西洋音楽を分析することで、いくつかの法則によってリズムが構成されていることを示している (Cooper & Meyer, 1960; Yeston, 1976; Kramer, 1988)。これらは、音楽理論としてリズムの特徴をまとめたものであるが、非常に体系的に西洋音楽のリズムを分析しており、心理学的に見ても有用な法則が多く含まれている。

### 1.6.1 音楽の階層構造

音楽は、個々の音事象が結びついて動機 (motive) となり、動機がフレーズ (phrase) に、フレーズが楽節 (period) に、という具合に、小さなグループが段々と大きなグループにまとまって一つのリズムを形作っている (Cooper & Meyer, 1960)。すなわち、音楽のリズムは、階層的な構造を持っているといえる。Cooper and Meyer (1960) は、このような音楽の階層は、大きく分けて3つのレベルから成り立っていると述べている (図5)。まず、彼らは、リズムの基礎となる階層を第一次リズムレベル (primary rhythmic level) と定義した (図5 1)。この階層のリズムは、一個の強拍と一個以上の弱拍とがひとまとまりになっており、完全なリズムグループとして機能している。この第一次リズムレベルは、いくつかの小さな音価の組み合わせから成り立っており、小さな音価によって、第一次リズムレベルよりもさらに短い部分的な時間パターンが形成される場合がある。ただし、こうした部分パターンは、そのみでは一つの完全なリズムグループを実現することができない。このような部分パターンを下位リズムレベル (inferior rhythmic level) と定義した (図5 1)。さらに、第一次リズムレベルを構成するいくつかのグループが一つにまとまることで、上位リズムレベル (superior rhythmic level) が構築されるとした (図5 2)。上位リズムレベルを構成するグループがさらにいくつかまとまることで、より高次の上位リズムレベルとな

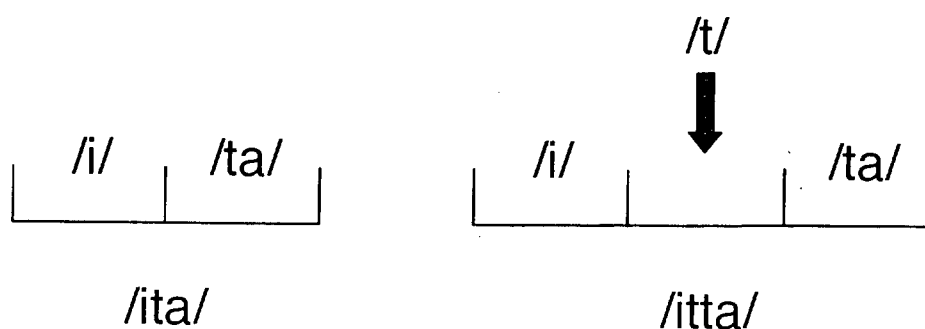


図4 日本語のモーラ構造。/いた/は、2モーラの単語である。/い/と/た/の間に1モーラ分の空虚な時間間隔を挿入すると、3モーラの単語として知覚される。/い/と/た/に挟まれたモーラには、物理的な音韻情報は全く存在していない。しかし、一つのモーラには一つの音素があるとした方が日本語の構造として自然である。そのため2モーラ目に促音 /つ/ が補完されて、/いった/ という単語が知覚される。

a. Presto

EXAMPLE 53

— : 強拍      ∪ : 弱拍

図5 音楽の階層性。図中の1.で示した枠組みが第一次リズムレベルを、2.で示した枠組みが上位リズムレベルを、i.で示した枠組みが下位リズムレベルをそれぞれ表している (Cooper & Meyer, 1960)。

る。音楽のリズムは、このようにして、等しい長さを持つ小さなリズムグループが次々と高次のリズムグループにまとまることで階層的な構造を成している。

### 1.6.2 リズムのカテゴリ化

次に部分的なリズムパターン、つまり、Cooper and Meyer (1960) の定義した下位リズムレベルを構成する音列パターンが、どのように知覚されているのかについて考えてみる。このレベルのリズムは、隣接する二つ、ないしは三つの時間間隔で構成されるような、最も小さなリズムグループであるが、音楽の階層性を形作る基礎となっており、音楽のリズムがどのようにして知覚されているのかを知る上で重要な意味を持っている。

Fraisse (1982) は西洋音楽の楽曲を分析し、一つの曲の中で使われる音符の80%以上が二つの時価のいずれかを持っており、その二つの音符の時価が、1:2、もしくは1:3となっていることを示した。西洋音楽の楽曲では、隣り合う時間間隔の比が、1:3, 1:2, 1:1, 2:1, 3:1といった単純な整数比となる場合が多いといえる。これは、楽譜を分析した結果であるが、演奏者によって生成された音列を分析する実験においても、隣り合う音の長さが単純な整数比となるパターンが重要であることを示唆する結果が示されている。Fraisse (1982) は、実験参加者に、5つか6つのタッピングによるリズムパターンを自由に生成させる実験を行い、実験参加者の作り出したリズムパターンに含まれる隣り合う時間間隔の比率が、1:1と2:1の周辺に集まっていることを示した。また、Povel (1981) は、物理的な長さを系統的に変化させた二つの時間間隔からなる刺激パターンを16回繰り返して呈示し、その後、実験参加者にその刺激パターンをタッピングによって再生してもらう実験を行った。その結果、呈示した隣接する時間間隔の物理的な比率が、1:2よりもずれてい



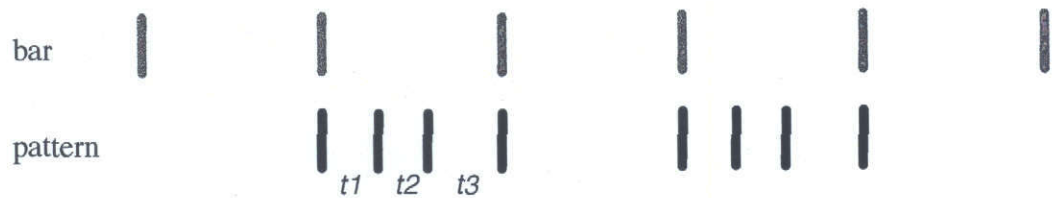
たとしても、実験参加者の再生した刺激パターンに含まれる隣接する時間間隔の比率は、1:2に近づくことを報告している。これらの研究は、隣接する時間間隔の関係が単純な整数比で表されるリズムカテゴリが存在する可能性を示している。

このようなリズムカテゴリが存在することを確かめるために、Desain and Honing (2003) は、1000 ms の時間間隔を体系的に3つに分割した刺激パターンを実験参加者に呈示し、音符を使ってその音列パターンのリズムを回答してもらう実験を行った。その結果、各刺激パターンに含まれる3つの時間間隔の物理的な長さを一定の細かい変化幅で段階的に変化させたにもかかわらず、実験参加者の回答が単純な整数比で表されるカテゴリに集まる結果が得られた(図6)。また、音列パターンの拍構造を二拍子や三拍子に変化させることで、各カテゴリ間の境界線の位置が変化すると報告している。この結果から、音楽的なリズムを知覚する場合には、音と音との時間間隔をそのまま知覚するのではなく、知覚した刺激パターンの時間構造を、音符の長さで表現できるような単純な整数比で表されるリズムカテゴリに当てはめることでリズムを知覚していると考えられる。

しかし、実験参加者の回答を音符によって得る方法が、適切であるかどうかは、疑問である。音符の長さは、16分音符、8分音符、4分音符といったように、それ自体が単純な整数比で構成されている。確かに、付点をつけたりすることで、複雑な時間パターンを表現することは出来るが、連続的な指標であるとはいえない。また、実験刺激と同時に、拍子を誘導するような強勢によるアクセントをつけた等時音列を呈示していることも[図6(a)]、カテゴリの形成に寄与している可能性が大きい。このような等時音列を同時に呈示することで、実験参加者は、3つの時間間隔が一拍を分割したものであると予測しやすくなる。そのため、その拍子において音楽的に違和感のない音符の組み合わせが優先的に選択される可能性が高い。つまり、Desain and Honing (2003) の結果は、音列パターンを音楽のリズムとして捉えたときに、音符によって表現できるリズムカテゴリが知覚上に形成されることを示しており、このリズムカテゴリの形成には、音楽の規則性や音楽経験が影響している可能性が大きい。

Drake (1993) は、リズムカテゴリが、音楽経験や年齢の影響を受けるかどうかを確かめるために、子供(5歳児と7歳児の2群)、音楽の訓練を受けていない成人、音楽的な訓練を受けている成人という、4種類の実験参加者群に対して、様々なリズムパターンの再生実験を行った。その結果、実験参加者によって再生されたリズムパターンの時間構造と教示した音符の時間構造との関係は、すべての実験参加者群において、次のような傾向を示していた。

(a) Stimulus Pattern



(b) Result

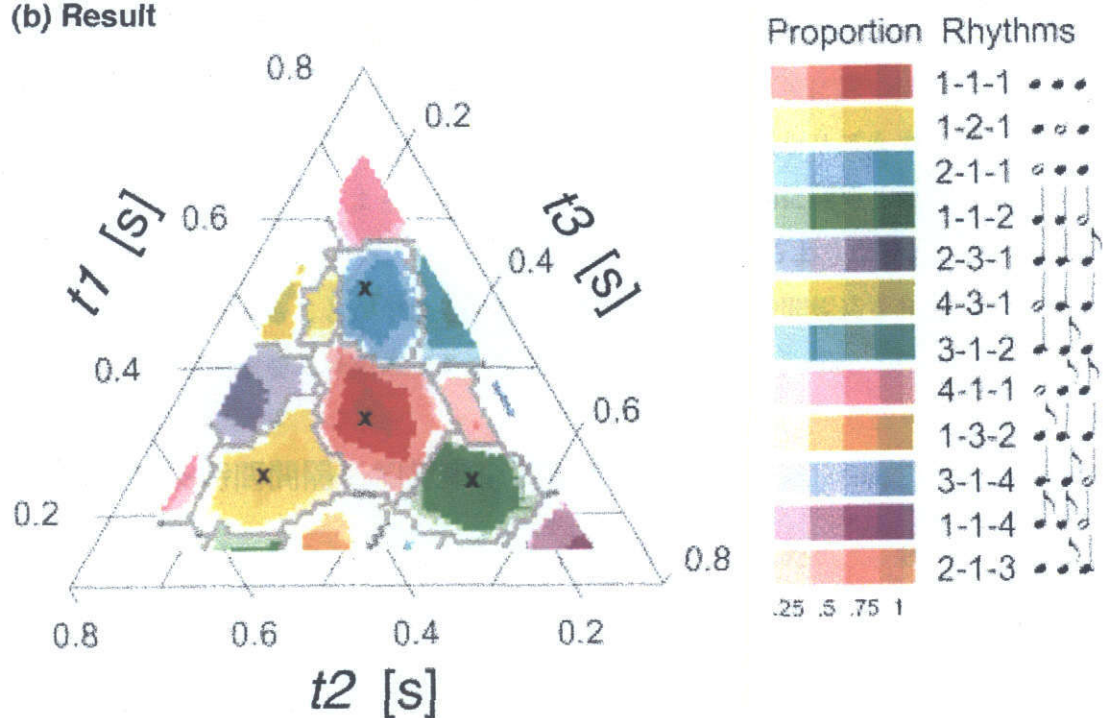


図6 Desain and Honing (2003) の示したリズムカテゴリーの判断境界。(a)は、実験で使用した刺激パターンの模式図である。彼らは、拍子を誘導する等間隔の音列 (bar) と実験参加者の評価する実験刺激 (pattern) とを異なる音を用いて同時に呈示した。(b)は、実験刺激を構成する三つの時間間隔の物理的な長さをもとに、実験結果を表した図である。3つの軸は、それぞれ(a)で示した $t_1$ 、 $t_2$ 、 $t_3$ の物理的な長さを表している。また、右に示した楽譜で表される回答が得られた確率を色の濃淡で表現し、灰色の線でそれらのカテゴリー間の判断境界を表している。

- (1) リズムパターンに含まれる時間間隔の比率が、2の倍数となっているパターンの方が、3の倍数となっているパターンよりも正しく再生される。
- (2) 2種類の時間間隔で構成されるパターンの方が、3種類の時間間隔で構成されるパターンよりも正しく再生される。

例えば、リズムパターンに含まれる時間間隔の長さの関係が、1:2の比率で表されるパターンでは、5歳児においても、80%以上の割合で正確に再生ができていた。このリズムパターンは、上述の二つの条件を両方とも満たしている。しかし、音楽的な訓練を受けている成人の結果を除いて、この二つの条件の一つでも満たしていないと、正確に再生できる割合が、50%以下にまで下がってしまう結果が得られた。また、音楽的な訓練を受けていない成人の結果と、7歳児の結果を比べると、ほとんど結果に違いはなかった。この結果を見ると、Desain and Honing (2003)の示したような複雑なリズムカテゴリーの枠組みは、音楽経験や前後の文脈によって可変なものであり、後天的に獲得した能力によって影響を受けやすいといえる。一方、一拍を1:2や1:3に分割するようなリズムパターンは、幼児でもすでに形成することができる枠組みであり、音列の時間構造を知覚的に体制化するための基礎となる重要な枠組みであると考えられる。

## 1.7 隣接する時間間隔の知覚の重要性

次々に呈示される音事象をリズムとして知覚的に体制化するために、我々は、等間隔の格子をもとに音列パターンを階層的に群化していることを述べてきた。このようなリズムの階層性は、音楽や音声に共通してみられる構造だといえる。さらに、リズムを細かく分析すると、隣接する時間間隔の比率が、単純な整数比となるパターンが、リズムを知覚する上で重要となることを述べた。

このような単純な整数比で表されるようなリズムカテゴリーの形成には、高次の段階における処理が大きく影響を与えられられる。これは、リズムカテゴリーの境界が、音列パターンの枠組みである拍子によって変化することや (Desain & Honing, 2003)、音楽経験のような個人の特性によって変化すること (Drake, 1993) などからも明らかだといえよう。よって、このリズムカテゴリーは、音列パターンを構成する時間間隔の長さを聴取者がどのように知覚しているのかを直接反映した結果とはいえない。

では、聴取者は、次々と呈示される時間間隔の長さをどのように知覚して、音列パターンを体制化しているのでしょうか。ここで、Longuet-Higgins and Lee (1982)の提案した

単音系列の知覚的体制化に関するモデルをもとに、音列の処理過程について、もう一度、考えてみたい。まず、二つの音が呈示された段階で、聴取者は、時間間隔の長さを知覚する。そして、この主観的な長さを格子の長さとして仮定し、次々と呈示される時間間隔と格子の長さを比較することで音列を体制化している。もし、直後に隣接する時間間隔の長さが、格子の長さ、つまり直前に呈示された時間間隔の主観的な長さと同じであれば、その長さをもとに、さらにその後の時間間隔の長さを処理していく。一方、直後に隣接する時間間隔の長さが、直前に呈示された時間間隔の主観的な長さとは異なる場合には、基準となる長さに変更を加える。このような処理を行うためには、隣接する時間間隔の主観的な長さを比較し、二つの長さが同じであるか、もしくは異なっているかを瞬時に判断する必要がある。よって、音列パターン処理過程について論じるためには、隣接する時間間隔の等時性というものが、どのようにして検出されているのかを明らかにすることが必要である。

## 1.8 分割時間の知覚

中島 (1979) は、隣接する二つの時間間隔のみで構成されるような単純な音列パターンを用いて、二つの時間間隔の主観的な長さの関係について定量的な分析を行い、隣接する二つの空虚な時間間隔の主観的な長さは、その物理的な長さに約 80 ms の定数を付け加えたものに比例するという「つけくわえ仮説」を提唱した。つまり、物理的な比率が  $t_1:t_2$  である二つの時間間隔の主観的な比率は、 $(t_1 + 80):(t_2 + 80)$  となる。Nakajima, Nishimura, and Teranishi (1988) は、50 ms から 1200 ms という広範囲の時間間隔を用いた比率判断の実験を行い、つけくわえ仮説の妥当性を確かめている。この仮説の注目すべき点は、隣接する二つの時間間隔の物理的な比率から、一意的に主観的な比率を予測することができるという点である。これは、隣接する時間間隔の比率を、単純な整数比で表されるカテゴリに分類して知覚しているというカテゴリ知覚の考え方と異なっている。Nakajima (1987) は、リズムカテゴリは、前後の文脈によって生じるものであり、そのような文脈が形成されない音列パターンでは、物理的な長さの比率の増加に応じて、主観的な比率も増加するとした。

ただし、隣接する二つの時間間隔の比率が、ほぼ 1:1 に感じられる時間条件、つまり、隣接する二つの時間間隔の物理的な長さの差が小さい条件では、実験結果が、つけくわえ仮説によって予測される値から逸脱している。Sasaki, Nakajima, and ten Hoopen (1998)

は、隣接する二つの時間間隔それぞれの主観的等価値を測定する実験を行い、二つの時間間隔の差が小さい場合には、それぞれの主観的等価値が近づいており、ほぼ 1:1 に感じられるカテゴリが形成されることを報告した。この 1:1 カテゴリに含まれる時間条件の範囲は、後の時間間隔が前の時間間隔よりも長い条件の方に広がっており、時間的に非対称的となっている。例えば、80 ms と 160 ms の時間間隔が隣接している場合は、主観的な比率が、ほぼ 1:1 となっているが、時間間隔の長さを反転させた、160 ms と 80 ms の時間間隔が隣接している場合は、主観的な比率は、1:1 に感じられない。一方、1:1 カテゴリに含まれない時間条件では、物理的な比率の増加に従って主観的な比率も変化しており、1:1 カテゴリ以外のリズムカテゴリは生じなかった。前後の文脈による影響がほとんどない単純な刺激パターンを用いたにもかかわらず、1:1 カテゴリのみがはっきりと生じていることから、単純な整数比で表されるカテゴリと 1:1 カテゴリとは、その生成過程が異なる可能性がある。

## 1.9 時間縮小錯覚

Sasaki et al. (1998) の研究でみられた 1:1 カテゴリは、先行する時間間隔によって、後の時間間隔の主観的な長さに著しい過小評価が生じることによって形成されていた。Nakajima, ten Hoopen, and Wilk (1991) と ten Hoopen et al. (1995) は、このような後の時間間隔に過小評価が生じる現象を ”時間縮小錯覚 (time-shrinking)” と名付け、定量的な実験を数多く行うことで、この錯覚現象の生じるメカニズムをモデル化している (Nakajima et al., 2004)。

Nakajima et al. (1991) は、ある時間間隔 ( $S$ ) の前に  $S$  よりも物理的な長さがわずかに短い先行時間 ( $P$ ) を隣接させた刺激パターンを実験参加者に呈示し、 $S$  の主観的等価値を測定した。その結果、 $P$  を隣接させない条件と比べて、 $P$  を隣接させた条件では、 $S$  が短く感じられる現象、すなわち時間縮小錯覚を発見した (図 7)。時間縮小錯覚については、 $S$  および  $P$  の物理的な時間間隔の長さを様々に変化させた実験が行われており、 $S$  の物理的な長さ  $S$  と  $P$  の物理的な長さ  $P$  との関係が、 $0 < S - P < 80$  [ms]、かつ、 $P < 200$  [ms] の条件を満たすときに、 $S$  に過小評価が生じることが分かっている (Nakajima et al., 1991; ten Hoopen et al., 1995)。また、時間間隔を構成する音の性質 (音の長さ・周波数・音声) を変えることによって、過小評価量に若干の違いが生じるものの、どのような音を用いても同様の範囲で過小評価が生じており (Yamashita & Nakajima, 1999; 山下・白川・石浦・中島,

1998; 小森・中島, 1998)、強固な錯覚現象であるといえる (Suetomi & Nakajima, 1998a; Remijn et al., 1999)。さらに、音の代わりに、光を用いた実験も行われ、過小評価の生じる範囲が、聴覚での結果とは異なるものの、同様の過小評価が生じることも報告されている (Arao, Suetomi, & Nakajima, 2000)。このような時間縮小錯覚が生じることで、後の時間間隔の主観的な長さが先行時間の長さに近づいており、この現象は、時間間隔の知覚に生じる同化現象の一種であると解釈することができる。

Nakajima et al. (2004) は、定量的データを分析し、この時間縮小錯覚の生じるメカニズムをモデル化している。彼らは、モデルの基本に、ある時間間隔の長さを判断するために必要な "付加的な処理時間 (additional mental processing time)" というものを想定している。これは、先に述べたつげくわえ仮説の考え方をもとにしている (Nakajima, 1987; Nakajima et al., 1988)。このメカニズムを簡単に説明すると、まず、時間間隔の主観的な長さは、時間間隔の物理的な長さに、 $\alpha$  という長さの付加的な処理時間をつげくわえた長さになる。そのため、二つの時間間隔 ( $t_1, t_2$ ) を隣接させた場合の主観的な比率は、 $(t_1 + \alpha):(t_2 + \alpha)$  となる。しかし、 $t_1$  と  $t_2$  との関係が、 $0 \leq t_2 - t_1 \leq \alpha$  の条件を満たすならば、 $t_2$  の長さを符号化する途中で、 $t_1$  とのたまかな等時性が判定され、残りの  $t_2$  の処理

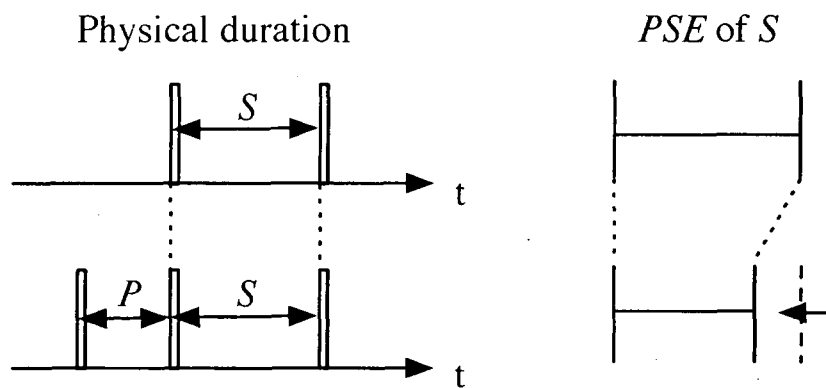


図7 時間縮小錯覚。時間間隔 ( $S$ ) を単独で呈示したときと、 $S$  の直前に先行時間 ( $P$ ) を付け加えたときの主観的等価値 ( $PSE$ ) を測定すると、 $S$  の物理的な長さは等しいにもかかわらず、 $PSE$  は、 $P$  を付け加えたときの方が、短くなる。Nakajima et al. (1991) は、この錯覚現象を、時間縮小錯覚と名付けた。時間縮小錯覚による過小評価は、 $S - P = 80$  [ms] の時に最も過小評価量が大きくなり、 $S$  と  $P$  の差が、100 ms を超えると、急に過小評価が生じなくなる。

時間が短縮される。そのために、短縮された処理時間の長さの分だけ  $t_2$  の長さに過小評価が生じる。このような知覚メカニズムを考えることで、 $t_2$  に生じる過小評価の傾向をおおまかに説明しようとした。このモデルを多くの定量的なデータに当てはめることで、 $\alpha$  は約 80 ms の長さを持つことが示されている。詳細には、この基本的な考え方のみでは、説明できない現象もあり、Nakajima et al. (2004) は、二つの時間間隔の差を判断するために必要な“絶対弁別閾 (absolute difference limen)”の考え方を導入し、モデルに修正を加えている。この点については、本研究の考察で重要となってくるので、第 4 章において詳しく取り上げる。

この時間縮小錯覚の特徴的な点は、まず、錯覚の生じる時間条件が、時間的に非対称なことである。例えば、隣接する二つの時間間隔が、100 ms と 150 ms の場合には、 $t_2$  の主観的等価値が 111 ms となり、39 ms の過小評価が生じている。しかし、刺激の時間条件を反転させた、150 ms と 100 ms の条件では、 $t_2$  の主観的等価値は、100 ms となり時間縮小錯覚が生じていない (Nakajima et al., 2004)。

また、 $t_2$  の時間間隔のみに錯覚現象が生じる、つまり、同化が片方向にのみ生じていることも特徴的である。一般的に、ある事象間の性質の差異が少なくなるように感じられる同化は、その効果が双方の事象に同等に生じることが多い。これに対して、時間縮小錯覚では、後の時間間隔に著しい過小評価は生じるが、先行する時間間隔に生じる同等の過小評価や過大評価は見つかっていない。

## 1.10 本論文の目的

時間軸上に点在している単一の音事象が、複数個集まることによって、音と音との時間間隔が知覚される。さらに、複数の時間間隔が知覚的にまとまることで、リズムが知覚される。このように、時間軸上に点在する複数の音事象間の時間関係をリズムとして体制化し、一つの音脈として知覚する一連のメカニズムがなくては、音楽や音声を使った人間特有の聴覚コミュニケーションは成立しない。時間間隔やリズムの知覚を研究することは、聴覚の基本的な知覚メカニズムを解明するために、有益である。

我々は、次々に呈示される音事象間の時間間隔を、等間隔の格子をもとに階層的に群化することで、リズムを知覚している。このようなリズムの階層性は、音楽や音声に共通してみられる構造だといえる。さらに、リズムを細かく分析すると、隣接する時間間隔の比率が、単純な整数比となるパターンが、リズムの最も基本的な階層を形成している。

一方、より単純な隣接する二つの時間間隔のみで構成される刺激パターンを用いた時間間隔の知覚に関する研究では、時間縮小錯覚によって、後の時間間隔が過小評価される時間条件が存在することが分かっている。その結果として、二つの時間間隔の主観的な長さが近づいており、隣接する二つの時間間隔の長さが、ほぼ 1:1 に感じられるカテゴリが形成されている。一方、そのカテゴリから外れた時間条件では、隣接する二つの時間間隔の主観的な比率は、物理的な比率に応じて変化しており、音列パターンのリズムに関する研究で発見されたリズムカテゴリが形成されるような証拠は見つかっていない。よって、音列の時間間隔を体制化するときには、1:1 カテゴリ、つまり、隣り合う時間間隔が同じかどうか重要な枠組みであり、その他のリズムカテゴリは、音列パターンの知覚的な体制化がある程度すすんだ高次の段階で形成されるものであると考えられる。

本研究では、このような隣り合う時間間隔の長さが、ほぼ 1:1 に感じられるような時間条件に注目した。音列を体制化するためには、隣接する時間間隔の等時性を瞬時に検出することが重要となってくる。このようリズムの基礎を形作ると考えられる時間パターンについて定量的なデータを大規模に収集し、それを細かく分析することは、時間間隔の知覚的な体制化の仕組みを解明するための基礎となりうる。

隣接する二つの時間間隔を構成する時間間隔のうち、後の時間間隔の知覚メカニズムについては、時間縮小錯覚のモデルを用いることでおおまかに説明しうる (Nakajima et al., 2004)。しかし、データを細かく分析すると、いくつか、時間縮小錯覚のモデルでは、説明できない現象が生じている。特に、先行する時間間隔の知覚に注目した研究はあまり行われていない。Sasaki et al. (1998) の行った実験では、先行する時間間隔にも有意な過大評価や過小評価の生じた時間条件が見つかっており、時間縮小錯覚の生じる時間条件のみでなく、隣接する二つの時間間隔の知覚について包括的に詳しく研究する必要がある。

そこで、本研究では、先行する時間間隔と後の時間間隔との両方に注目した精神物理学的な実験を行い、得られたデータから、時間間隔を隣接させたときの主観的な長さや時間間隔を単独で呈示したときの主観的な長さとのずれが、時間縮小錯覚とは異なる知覚現象によっても生じていることを確かめた。さらに、そのデータを用いて、隣接する時間間隔の知覚的な処理過程を説明しうるモデルについて考察した。



## 1.11 本論文で用いた刺激条件について

時間間隔の知覚について実験心理学的手法を用いて調べる場合には、時間的な変化が少ない音や、非常に短い音を使ったパターンを用いることが多い。このような、可能な限り単純な実験室的状況から出発して知覚の基本的法則を見だし、それをより複雑な現象に順次拡大していくというやり方が堅実である。リズムの研究をする上で、興味の対象となるのは、音楽や音声の知覚であり、実際の楽器演奏音や音声を用いて実験を行う方が、興味深い結果が得られる可能性が大きい。しかし、実験に用いた刺激を構成する音が複雑になればなるほど、刺激パターンを体系的にコントロールすることが難しくなる。また、得られた結果へ影響を与えている要素について考察する際にも、議論が発散してしまい、結論が曖昧になってしまう可能性が大きい。そこで、本研究で題材としている状況や用いる言葉の意味をここで明確に定義しておく。

まずは、研究対象とした時間間隔について詳しく説明する。一般的に、時間知覚研究では、音の始まりから次の音の始まりまでの時間間隔 (inter-onset interval, IOI) に注目していることが多い。これは、音事象自体の性質よりも、音事象のオンセット間の時間間隔が、知覚される時間間隔の長さに大きく影響を与えるからである (Handel, 1993)。そこで、本研究においても、音の始まりから次の音の始まりまでの長さを時間間隔と定義した。

ただし、知覚的な音の始まりを明確に定義することは難しい。音の立ち上がり部分や知覚的な音の始まりについての研究は、多く行われている (Gordon, 1987; Vos & Rasch, 1981)。特に、音声の知覚的な始まりは、perceptual center (P-center) と呼ばれ、それについては、様々な研究が行われている (Morton, Marcus, & Frankish, 1976; Jong, 1994)。しかし、いまだに明らかになっていない部分が多い。

このような音の始まりの問題を回避するために、本研究では、純音やパルス列といった単純で物理的な性質の変化が小さい短い音を用いて時間間隔を構成した。このような音を使うことで、音の始まりを明確に定義できなくても、時間間隔を構成する音の知覚的な始まりが、相対的には保たれることになる。

本研究で行った実験で用いる刺激パターンの実験条件は、すべて、三つの音で区切られた隣接する二つの空虚な時間間隔によって構成した。なお、空虚な時間間隔とは、二つの音の間に音事象がなく、背景のノイズとの SN 比も十分にとれており、擬似的に無音と見なせる時間間隔のことである。以後、最初に呈示される時間間隔を  $t_1$ 、次に呈示される

時間間隔を  $t_2$  と呼ぶ。刺激パターンの時間条件は、 $t_1$  の物理的な長さ  $t_2$  の物理的な長さを用いて、 $/t_1/t_2/$  (単位は、ms) と表記する。例えば、 $t_1$  が 80 ms、 $t_2$  が 120 ms の条件を、 $/80/120/$  と表す。時間間隔を構成する音は、区切り音と呼ぶ。区切り音には、20 ms 以下の長さを持つ短い音を用いた。ほとんどの場合、それぞれの実験で用いた区切り音は、すべて同じ物理的な性質を持つ。ただし、実験 1 と実験 3 では、区切り音の音圧差による影響も調べたため、音圧のみ異なる 2 種類の区切り音を用いた。このような区切り音によって区切られた時間間隔を、1 ミリ秒の精度で制御して、体系的に変化させた。このように、時間間隔のみを精密に変化させた刺激パターンを使うことで、時間間隔の物理的な長さが知覚に与える影響のみを測定できるように配慮した。

## 1.12 本論文の構成

第 1 章では、本論文の背景となるリズム知覚研究と時間知覚研究の動向を概観し、本論文の目的を述べた。次章から、実際に行った実験の結果を報告する。第 2 章では、隣接する二つの時間間隔の主観的な比率を、数字や線分間の距離を用いて測定した。その結果、隣接する二つの時間間隔の物理的な比率が 1:1 ではないにもかかわらず、主観的な比率がほぼ 1:1 に感じられる時間条件が存在することが分かった。この 1:1 カテゴリの境界を明確にするために、時間条件を体系的に変化させた刺激パターン間の非類似度を測定し、その結果をもとにクラスタ分析を行った。クラスタ分析の結果から、刺激パターンが三つのクラスタに分類されることが分かった。そのうちの一つは、主観的な比率がほぼ 1:1 に感じられる時間条件で構成されており、1:1 カテゴリのカテゴリ境界を客観的な統計手法で示すことが出来た。第 3 章では、第 2 章で発見された 1:1 カテゴリに含まれる時間条件を構成する二つの時間間隔の長さが、それぞれどのように知覚されているのかを確かめた。このような隣接する二つの時間間隔を知覚する際には、 $t_2$  に著しい過小評価が生じる時間条件があることが分かっている。この  $t_2$  に生じる時間縮小錯覚については、多くの定量的なデータから過小評価量を表すモデルが提案されている。しかし、先行研究の実験結果を詳細に検討すると、時間縮小錯覚のモデルのみでは説明できない現象も生じている。そこで、隣接する二つの時間間隔の長さを体系的に変化させた刺激パターンを用いて、 $t_2$  だけでなく  $t_1$  の主観的等価値も測定した。さらに、先行研究で行われた実験よりも多くの実験参加者から定量的なデータを集め、隣接する二つの時間間隔に時間縮小錯覚とは異なる知覚現象が確かに生じることを統計的にも明らかとした。その結果から、1:1 カテゴリ

内の時間条件では、 $t_2$ だけでなく $t_1$ にも過大評価や過小評価が有意に生じていることが分かった。 $t_1$ と $t_2$ に生じる過大評価や過小評価によって、二つの時間間隔の主観的等価値は近づいており、二つの時間間隔に双方向の同化が生じているといえる。また、1:1カテゴリーを外れた時間条件では、対比が生じている可能性を示すことも出来た。これらの結果をもとに、第4章では、隣接する二つの時間間隔における知覚現象を説明しうる仮説を提案する。仮説では、時間縮小錯覚と本研究で見つかった双方向の同化とが、異なる知覚過程において生じていると考えた。時間縮小錯覚は、早期の知覚段階で時間間隔の長さの処理時間が短縮されるために生じている。つまり、即時的に生じる錯覚現象であると考えられる。その後、 $t_1$ の長さと時間縮小錯覚によって過小評価された $t_2$ の長さとの関係を単純化して記憶に格納するために、双方向の同化や対比によってこの二つの長さに等質化が生じるのではないかと考えた。実験の結果にこの仮説を当てはめることで、 $t_1$ に生じる過大評価や過小評価の傾向をうまく説明することができた。最後に、第5章で、本研究で得られた知見をまとめるとともに、双方向の同化の生じるメカニズムをさらに詳しく調べていくための今後の研究指針を提案して本論文の結びとする。