

No.

TR-A-0043

乳児の泣き声における音響特徴の分析と  
カテゴリー判断の単位

009

東倉 洋一

片桐 滋

塚本 妙子

Yoh'ichi Tohkura

Shigeru Katagiri

Taeko Tsukamoto

1988. 12. 28

A T R 視聴覚機構研究所

# 乳児の泣き声における音響特徴の分析と カテゴリー判断の単位

塚本妙子・片桐 滋・東倉洋一

## 【まえがき】

乳児の泣き声は、近年、医学、発達、言語の分野において注目されるようになり、その研究もそれぞれの目的のもとに進められてきた。しかし、研究によって用いている用語や定義づけなど、多くの点でまだ統制がされていないために研究結果の比較を行うことが困難な現状にある。また、泣き声にはどのような音響特徴があり、そうした音響特徴の何が示差的特徴として知覚的なカテゴリー差異をもたらすのかについては、まだまだ解明されていない部分が多い。したがって、泣き声の研究を系統的に進めて行くことが重要であり、そのためには、特に、泣き声の知覚判断と物理的性質を対応づけながら研究していくことが必要である。ここでは、まず、泣き声の音響特徴の概観的な分析と、知覚実験ならびにその結果に関連した音響特徴の分析結果を、

第1章 泣き声の音響特徴

第2章 泣き声のカテゴリー判断における単位

として報告する。

## 第 1 章 泣き声の音響特徴

【はじめに】

泣き声の分析に関する初期の研究は、その多くが医学的見地から、中枢神経学的に異常をもつ乳児を早期発見するという目的のもとにおこなわれていた。しかし近年になって、乳児の泣き声が少なくとも部分的に中枢神経系のメカニズムによってコントロールされており、乳児の内的なストレスを反映しているものである（Wolff, 1967 ; Lester, 1978, 1984）ことから、泣き声の音響特徴のなかに乳児の要求や情緒が表現されていることが推察されるようになった。また、最近では、心理学の分野においても、泣き声が非言語的コミュニケーションのひとつとして見直されるようになってきた。さらに泣き声と言語の生成の連続性、つまり韻律情報において泣き声と言語との関連性を確かめるためにも、泣き声の音響特徴が分析されるようになってきた。したがって、現在では、泣き声の分析に関する研究は、主に

（1） 中枢神経学的に問題のある乳児を早期発見する、

（2） 前言語期の母子相互交渉における伝達行動として、泣き声の機能性とメカニズムを解明する、そして

（3） 泣き声と言語との関連性を検証していく、

という3つの観点のもとに行われている。

泣き声の分析方法として、多くの研究がサウンドスペクトログラムを使った音声分析の手法を基礎としている。しかし、泣き声は乳児の呼吸の周期性のもとに発声された複雑な連続音によって構成されており、しかも乳児の調音は未熟であるために、大人の発声における音声の分析がそのまま適用されるものではない。そこで Wasz-Höckertら（1968）は、乳児の一呼吸によって発声された一連の音声を分析の単位として、その単位のなかの音響特徴をサウンドスペクトログラム上に現れた物理的形態に従って分類し、細かく定義づけて標準化を試みている。しかし、サウンドスペクトログラム自体の精度と乳児の調音の未熟さという2つの要因によって、その物理的形態は不安定である。したがって、これらの物理的形

態とそれによって定義づけられた用語との対応関係を明確に行うことはそれほど容易でない。また、サウンドスペクトログラムから基本周波数 ( $F_0$ ) やパワーなどの精密な値を算出することは不可能である。

そこで以上の問題を解決するためには、泣き声を A/D 変換してコンピュータによる多面的かつ精細な音響特徴の分析を行い、より客観的で精度よく表現された物理的形態に従って泣き声の音響特徴の分類を行うことが有効な手段となる。そして、泣き声の音響分析により安定した物理的形態を抽出し、これらの分類を精度よく行うため、さらにこれらの物理的形態や分類を使った統計的分析を効果的に行うためには、泣き声のデータベースを構築することが不可欠となる。

ところで、前言語期における伝達行動として泣き声の機能性とメカニズムを解明するためには、母親が泣き声をどのように認知し、どのような音響特徴を手がかりに泣きの原因を判断するのかという問題を検討することが必要である。そのような立場から泣き声を分析する場合、Wasz-Höckertら (1968) に代表される従来の研究には次のような問題がある。まず、分析に用いられている泣き声の多くは、泣きの原因を明確にするために操作的な実験状況のもとに収録されたものである。そこでは分析されている泣きのカテゴリーが操作的に生じさせられた泣きのものに限られているということや、また、実験者が意図する泣きの原因以外の要因を完全に排除することは困難である。そして、Murry (1980) も指摘しているように、泣きのカテゴリーに、birth cry のような一人の乳児からは 2 度と聞くことがないものや、pleasure cry のような非叫喚が含まれているということから、泣きの分類が適切であるのかという疑問がある。さらに、分析の基準として乳児の一呼吸で発声された一連の音声単位が用いられているが、そうした音声単位のなかでの分析に終始し、泣き声全体の音響特徴が明かでない。つまり、従来の研究においては、

- (1) 分析対象となっている泣き声のカテゴリーが限定されており、
- (2) その多くが操作的な実験状況における泣きのものである、
- (3) 一呼吸という細分化した発声単位を基準とした分析から泣き声全体の音響特徴を表現するのは限界がある、

等の問題点がある。

そこで以上の問題点に対して、

(1) 家庭状況における母子の相互交渉の観察場面において生じた泣きを分析対象とし、

(2) 客観的で精度よく表現された音響特徴をパラメータとした泣き声全体のパターンを分析する

ことが有効だと考える。

泣き声全体のパターンを分析した研究としては、泣き声の律動性に注目した二木(1973)と小林(1986)の報告がある。二木は、サウンドスペクトログラムの分析において cry - interval - cry - interval という発声と非発声の律動性に注目し、生後二日目から3カ月までの乳児60名の pain cry を分析した。その結果、律動性の高い新生児期からその後の移行期(生後3週目から6週目まで)を経て、不規則な律動性を示す時期(7週目以降)へ発達的に変化することを明かにした。そして同時に、律動性の崩れる時期は乳児の行動が spontaneous から social へと発達的に変化する時期に対応していることを述べている。小林は、生後3日目から7日目までの泣き声を pain cry、hungry cry、discomfort cry に分類し、30秒間に示された発声区間の継続時間に基づいた時系列を分析した。その結果、泣き声のリズムがそれぞれの状況によって異なることを確かめた。

二木の報告した泣き声の律動性が崩れる時期は、Illingworth(1955)の「泣き声は、初め単純な生理的変化の表出や環境から受ける緊張の放出にすぎないものが次第に分化して、より心理的な要求や情緒を表出するものになる」時期とも一致しており、泣きの分化と律動性の崩れにおける関連性が示唆される。また、小林の報告からも、新生児期においても泣き声の時系列パターンに乳児の情緒や要求の違いが表現され、泣き声全体のパターンを分析することが重要となる。

本研究では先に挙げた問題の諸点に留意し、

(1) まず、泣き声のサンプルには日常の母子の相互交渉の観察場面におい

て生じたものを用い、

(2) そのサンプルに基づく泣き声のデータベースを構築する。

(3) そして、データベースを用いることによって、泣き声の音響特徴を定量的に分析し、

(4) さらに泣き声の音響特徴による時系列パターンと、

(5) 発声と非発声という2つの明白な違いのある音響特徴に着目して、その2つの音響特徴による時間構造を分析する。

### 【方法】

泣き声の収録 乳児の泣き声は、日常の母子相互交渉の観察場面において生じたものを収集した。観察は、原則的に乳児の生後3週目から満1才に至るまでの約1年間を期間とし、乳児の生後3カ月までは毎週、それ以降は隔週に1度、各家庭を2、3名の観察者が一定の時間内に訪問し、日常の母子相互交渉に近い状況で行った。観察者は、筆者ならびに心理学専攻の学生である。観察の記録は、観察者による逐一筆記記録とVTR録画（SONY BETAMAX SL-F1；SONY CCD-V8 AF2）を用いて行った。なお観察法による記録の補足資料として、初回訪問時には乳児の生育歴に関する質問紙を、そして訪問ごとに「乳児の生活リズム」の記録用紙を母親に依頼した。「乳児の生活リズム」は、観察前日24時間と観察直前までの乳児の睡眠、授乳（離乳食を含む）、おしめの交換に関する時間を確かめるための記録用紙である。

泣きの分類 乳児の不快時における発声は、単発的発声を不快発声（negative vocalization）、連続的発声をぐずり（fuss）、はっきりとした持続的発声を泣き（cry）と定義づけて分類した。観察時に生じた泣きの原因は、数名の観察者（観察場面に参加した観察者と参加しなかった観察者を含む）がVTR録画の再生ならびに観察時における資料をもとにそれぞれ判断し、その一致率をとって評定された。評定された泣きは、その原因となった乳児の状況にしたがって分類された。その際、一致率の低かった泣きは再度検討し評定し直したが、なお80%以上の一致率をみなかった泣きは分類の手続きから除外した。

サンプル 泣きは、新生児期には生理的変化からくる不快感の表出にすぎないが、生後1カ月を過ぎるころになると、より心理的な要求や情緒を表出するものが分化してくる ( Illingworth, 1955 ; 二木, 1979 ; 塚本, 1985 )。また、泣きは単一の状況よりも複合した状況で出現することが多い。しかし、カテゴリカルな差異を明確にするためには、複合した状況の泣きでなく単一の状況のものに限定することが有効である。したがって、サンプルには、分類されたカテゴリーのうち、「空腹」「眠い」「甘え」「怒り」の4つを選択した。サンプルは、それぞれのカテゴリーにおいて、乳児の生後9週から17週までの泣き声を選出した ( 継続時間は、各25~37秒 )。カテゴリーの定義として、「空腹」は乳児が空腹のときのもの、「眠い」は乳児が眠くて不機嫌になったときのもの、「甘え」は母親が乳児のそばに来て相手になることによって泣きやんだときのもの、そして「怒り」は、乳児の泣きの原因に関係なくある状況の泣きから移行したものであり、母親ならびに観察者によって「怒った」と判断され評定されたものとした。サンプルの内訳は Table 1 に示すとおりである。

データベースの構築 サンプルは、VTRからコンピュータ ( MASSCOMP 5600 ) を用いてサンプリング周波数20kHz、16ビットでA/D変換 ( LPF = 10kHz, -96dB/oct ) され、音声ファイルに格納された。これをフレーム長20msec、フレーム周期2.5msecのハミング窓を用いてフーリエ変換し、デジタルスペクトログラム ( サウンドスペクトログラムの狭帯域45Hzにほぼ等しい ) を求めた。また、ケプストラム分析して $F_0$ を抽出した。これらは波形とともにグラフィック表示した。なお $F_0$ の抽出誤りは視察によって修正を加えた。この $F_0$ と波形とからなるグラフィック表示に基づき、発声区間をピッチ構造の周期性の強さと有無によって以下のように分類した。つまり、周期的なピッチ構造をもつ区間を " pp ( periodic pitch ) "、非周期的なピッチ構造をもつ区間を " ps ( pseudo pitch ) "、無周期的なピッチ構造をもつ区間を " ap ( aperiodic pitch ) " とした。そして、非発声区間は、明瞭に知覚される吸気音を " is ( inspiratory sound ) "、知覚されにくい弱い吸気音を " i ( inspiration ) "、呼気音を " e ( expiration ) "、発声の休止区間を " p ( pause ) "、そして乳児の発声以外の音声、例えば物の音・乳児が舌を鳴らした

音・母親の声などを " s ( sound ) " として定義づけた ( Table 2 )。但し、スペクトログラム上の形態ならびに知覚的にも判別が困難であった " i " や " e " は " p " として分類した。この定義に従い、 Fig. 1 に示すように泣き声の音響特徴にラベリングを施すことによってラベルファイルを作成したが、ラベリングはそのラベルに該当する音声区間を D/A 変換することで聴覚的にも確かめて行った。そして、音声ファイルとラベルファイルとからなるデータベースを構築した。手続きは Fig. 2 に示すとおりである。

## 分析

1 基本周波数 (  $F_0$  ) データベースを用いて、周期的なピッチ構造をもち " pp " とラベリングされた区間を検索し、その区間の  $F_0$  の平均値と標準偏差を求めた。

2 時系列パターン Fig. 3 に示すように、横軸に時間、縦軸にラベルの種類を取り、各ラベル区間の継続時間を時系列パターンにして表示した。

3 発声 / 非発声区間による時間構造の分布 発声 ( pp, ps, ap ) 区間と非発声 ( i, is, e, p, s ) 区間の時間構造の連関を見るために、 Fig. 4 に示すように隣接する発声 - 非発声 ( O 印 )、非発声 - 発声 ( X 印 ) の2種類の組合せによる時間構造の分布を表示した。さらに、発声 / 非発声区間それぞれにおいて、継続時間の長さによる頻度分布を求めた。

## [ 結果 ]

1 基本周波数 (  $F_0$  ) 各サンプルの  $F_0$  の平均値と標準偏差をカテゴリーごとにまとめたものが Table 3 である。その結果、総体的に  $F_0$  の平均値は「怒り」が一番高かった。しかし、各カテゴリー内においてサンプルごとの個人差が大きく、 $F_0$  の平均値の範囲にバラつきが示された。そこで、MA と YU を取り上げて、個人内におけるカテゴリー間の差異を確かめ Table 4 に示した。その結果、MA における  $F_0$  の平均値は「怒り」が一番高く、次いで「空腹」「甘え」「眠い」の順であった。YU の場合、「怒り」のサンプルが欠落しているた

めその比較はできなかったが、その他のカテゴリー間に大きな差異が認められなかった。

2 時系列パターン Fig. 3 の時系列パターンから、以下のことが示された。

「空腹」は、全体に発声区間の間隔が狭かった。発声区間のなかでは非周期的なピッチ構造をもつ "ps" "ap" 区間が多くみられた。時に発声/非発声のリズミカルな連続がみられた。「眠い」は、発声/非発声区間ともに継続時間の短いものが多いが、非常に長い継続時間をもつ発声区間がみられた。「甘え」は、発声区間の継続時間の長いものも多く、また pause や吸気も長いものも多くみられた。発声区間のなかでは、周期的なピッチ構造をもつ "pp" 区間が多かった。「怒り」は全体に発声区間の継続時間が短く、pause がほとんどみられなかった。発声区間のなかでは "pp" 区間が少なく、"ps" "ap" 区間が多かった。また、ある一定の継続時間をもった発声/非発声区間のリズミカルな繰り返しが多くみられた。

3 発声/非発声区間による時間構造の分布 隣接する発声-非発声区間と非

発声-発声区間という2種類の組合せによる時間構造の分布を Fig. 5 に示した。ここでは各カテゴリーに含まれる3つの音声データをまとめて表示し、横軸に発声区間の継続時間、縦軸に非発声区間の継続時間を取り、発声-非発声の組合せを○印、非発声-発声の組合せを×印で表示した。この図では発声-非発声、非発声-発声という2つの組合せを用いたために、最初と最後の区間以外は同じ区間を2度測定したことになる。また○×の数は、継続時間の長い区間が多いと減少し、継続時間の短い区間が多いと増加する。Fig. 5 から以下のことが示された。つまり、発声/非発声区間の時間構造の分布において、そのバラツキの最も大きいのが「甘え」、発声/非発声両区間ともに時間構造が短く集約された分布を示したのが「怒り」であった。そして、「眠い」では、非発声区間の時間範囲は短い(1秒以内)が、発声区間の時間範囲は比較的広い分布を示した。しかし、この分布には発声/非発声区間とも時間構造が短く集約された分布とこれを含む大きいバラツいた分布という2つの分布が認められた。さらに、「怒り」と「眠い」の中間的で、比較的バラツキの小さい分布を示したのが「空腹」であった。

そこで Fig. 5 を補足する意味で、発声/非発声区間の継続時間の長さによる頻度分布をカテゴリーごとにヒストグラムにして表示した。時間は 250 ms ごとに区切り、その時間範囲にある区間の数を 3 つの音声データ分すべてに関して集計し、頻度を求めた。実線は発声区間、破線は非発声区間の頻度を示す。結果は、Fig. 6 のとおりである。「空腹」では、発声区間の継続時間は 2 sec までに集中しており、非発声区間の継続時間は 500 ms までに集中がみられた。「眠い」は発声区間の継続時間のピークが 500 ms のところにあるが、3 sec を越えたところまで広く分布がみられた。非発声区間の継続時間では 500 ms までに集中がみられ 1 sec 以上の範囲に分布がみられなかった。「甘え」では、発声区間においても非発声区間においても分布の範囲が広く、明かな集中は見られなかった。「怒り」は、発声区間の継続時間が 500 から 750 msec のところに集中しており、非発声区間の継続時間はそのほとんどが 500 msec までの継続時間で 250 msec にピークがあった。

#### 【考察】

F<sub>0</sub> の分析においては、総体的ではあるが、「怒り」の平均 F<sub>0</sub> 値が一番高いことが示された。しかし、従来の研究結果と同様に、ある継続時間内の平均された F<sub>0</sub> 値からカテゴリー差異は見いだされなかった。また、同じカテゴリー内の平均 F<sub>0</sub> 値において個人差によるバラつきが大きかった。これらの点についてはさらに検討することが必要である。

時系列パターンと発声/非発声による時間構造の分布に示された結果は、次のようにまとめることができる。「空腹」は、発声区間の継続時間にかかわらず、非発声区間の継続時間が短く、発声/非発声がリズムカルに繰り返されることがあった。したがって、発声/非発声区間の時間構造の分布においては比較的バラつきの少ないまとまったパターンを示した。また、発声区間の種類に関して、周期的なピッチ構造のない "ps" "ap" が多かった。「眠い」は、発声区間の継続時間において全体としては短いものが多いが、ときに非常に長い継続時間のものがあり、分布の範囲が広がった。したがって、発声/非発声区間の時間構造において、非常に集約された分布とバラついた分布の 2 種類のパターンを示した。「甘え」は、発声/非発声区間ともに継続時間の長いものが多くみられ、発声/非

発声区間の時間構造の分布では非常に広範にバラついたパターンを示した。また、発声区間の種類に関しては、周期的なピッチ構造をもつ "pp" 区間が多かった。

「怒り」は、全体に発声/非発声区間ともに継続時間が短く、pause がほとんど見られなかった。したがって、発声/非発声区間の時間構造において非常に集約された分布のパターンを示した。以上のことから、乳児の情緒や要求という内的状況の違いによって、泣き声の音響特徴の時間的変化のパターンに差異があることが示された。

特に、発声/非発声区間の時間構造の分布において、「甘え」の泣き声が他の3つの泣き声とはかなり異なるパターンを示したことは興味深い。そのことは、先に述べたように、泣き声がおそらく部分的に中枢神経系のメカニズムによってコントロールされているものである(Lester, 1978; Wolff, 1967)ということや、リズム性の高い新生児期の泣き声が乳児の内的状況によって異なるリズム性をもつ泣き声に分化するという二木(1973)の示唆に従えば、以下のように考察できる。発達的にみると、これらの4つのカテゴリーのなかでは「甘え」が最も遅く分化し出現する泣き声(塚本, 1985)である。また「空腹」「眠い」は乳児の生理的な不快感に伴うものであり「怒り」は原始的な基本感情のひとつであるのに対して、「甘え」は乳児の心理的な不快感に基づく要求から生じた泣き声である。したがって、「甘え」の泣き声は、「空腹」「眠い」「怒り」の泣き声よりコントロールされた高次の泣き声であり、したがって発声/非発声の時間構造に規則性のない独自のパターンを示したと考えることができる。

今回の分析において、泣き声の音響特徴のラベルに従った時系列パターンや発声/非発声による時間構造の分布というように、音響特徴の時間的変化パターンにカテゴリー差異があることが示された。しかし、今後さらに、いろいろな角度からの分析を行うことが必要である。

#### 【要約】

泣き声の安定した分析を行うためには、泣き声のデータベースを構築することが有効な手段となる。また、伝達行動としての泣きの機能性とメカニズムを解明するうえで、泣き声の分析に関する従来の研究の問題を解決するためには、家庭状況における母子の相互交渉の観察場面において生じた泣きを分析対象とし、客

観的で精度よく表現された音響特徴をパラメータとした泣き声全体のパターンの分析を行うことも重要だと考える。

そこで本研究では、泣き声のデータベースを用いることによって、

- (1) 泣き声の音響特徴を定量的に分析し、
- (2) 音響特徴のラベルに従った時系列パターンと、
- (3) 発声 / 非発声区間の時間構造の分布

を確かめることを目的とした。

その結果、時系列パターンと発声 / 非発声区間による時間構造の分布においてそれぞれカテゴリー差異のあることが示された。

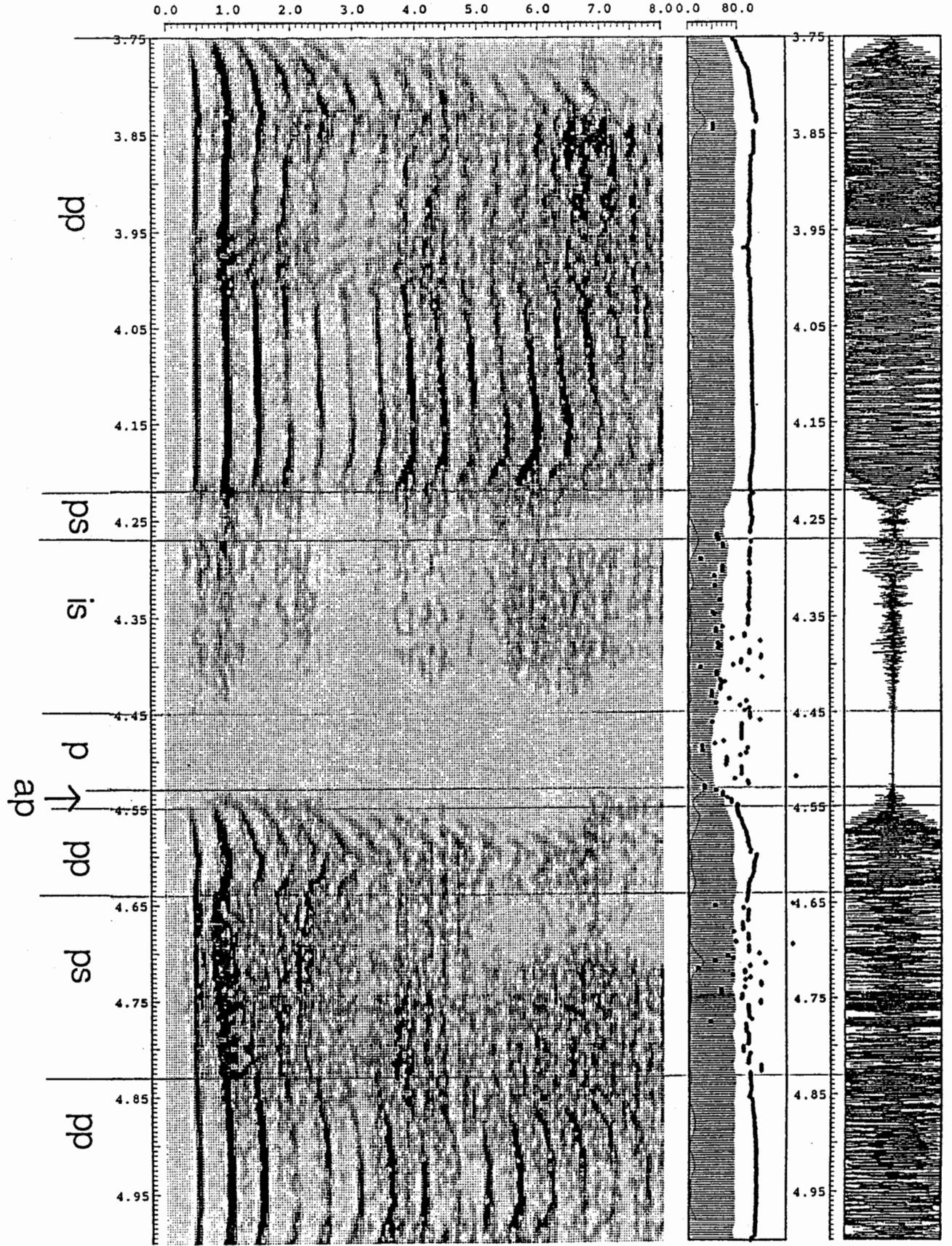
**Table 1 Contents of samples**

Category	Infant	Age (weeks)	Duration (sec)
Hungry	YU	10	27.5
	MA	10	25.2
	MA	12	26.1
	TE	17	28.6
Sleepy	MA	10	26.9
	AS	11	26.3
	YU	17	25.7
Call	TE	11	36.5
	TA	12	26.5
	YU	15	26.0
	MA	15	26.4
Angry	TO	9	26.8
	MA	12	26.1
	MA	13	27.7

**Table 2 Definition of labels**

Labels		Acoustic features
Phonation	pp	Periodic pitch pattern
	ps	Pseudo pitch pattern
	ap	Aperiodic pitch pattern
Non-phonation	is	Inspiratory sound
	i	Inspiration
	e	Expiration
	p	Pause
	s	Oral closure release sound of infant, or some noise

Fig. 1 Example of labeling based upon cry spectrogram and  $F_0$  contour



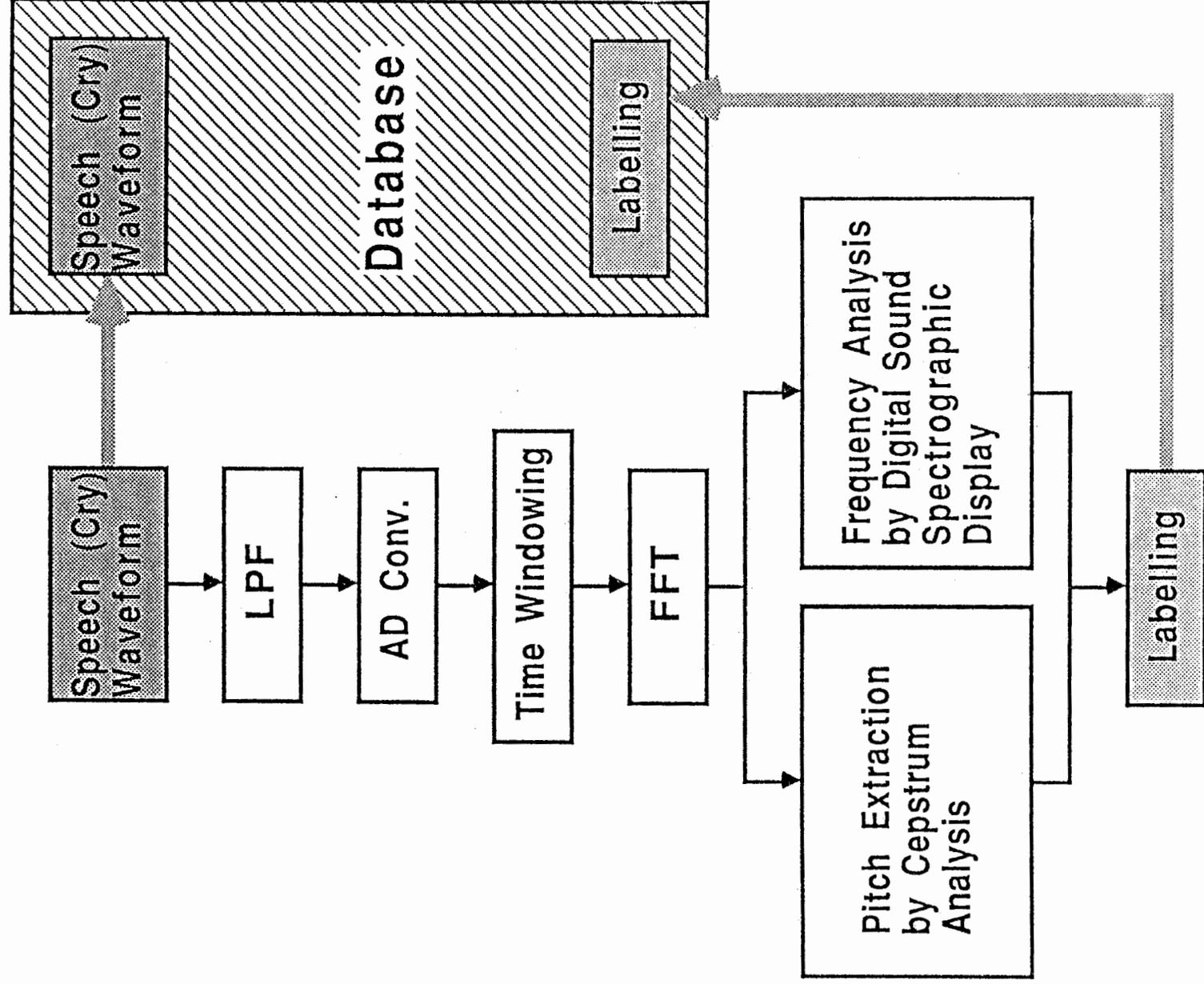


Fig. 2 Blockdiagram of database construction

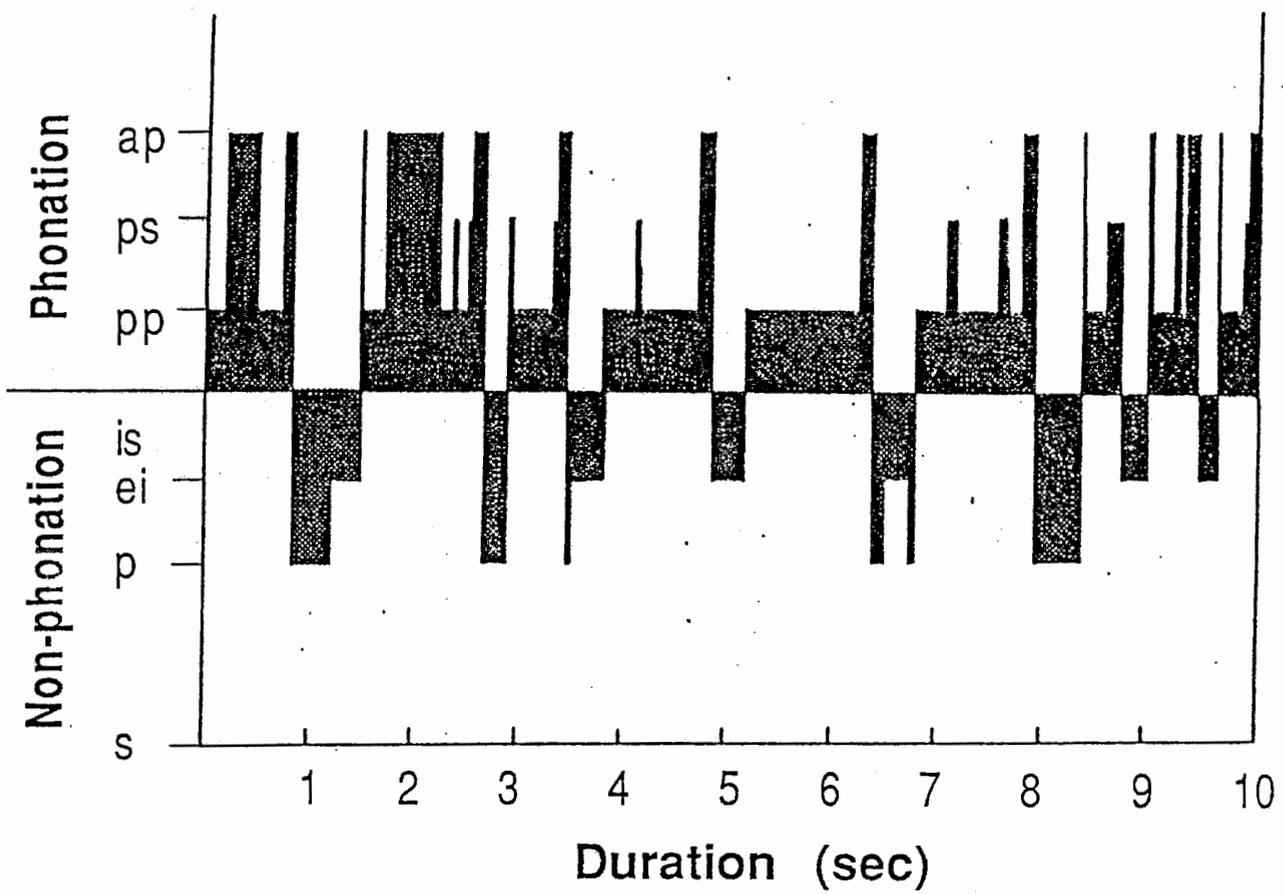


Fig.3 Representation of temporal patterns by labels for an "angry" cry (i.e., MO[13 weeks])

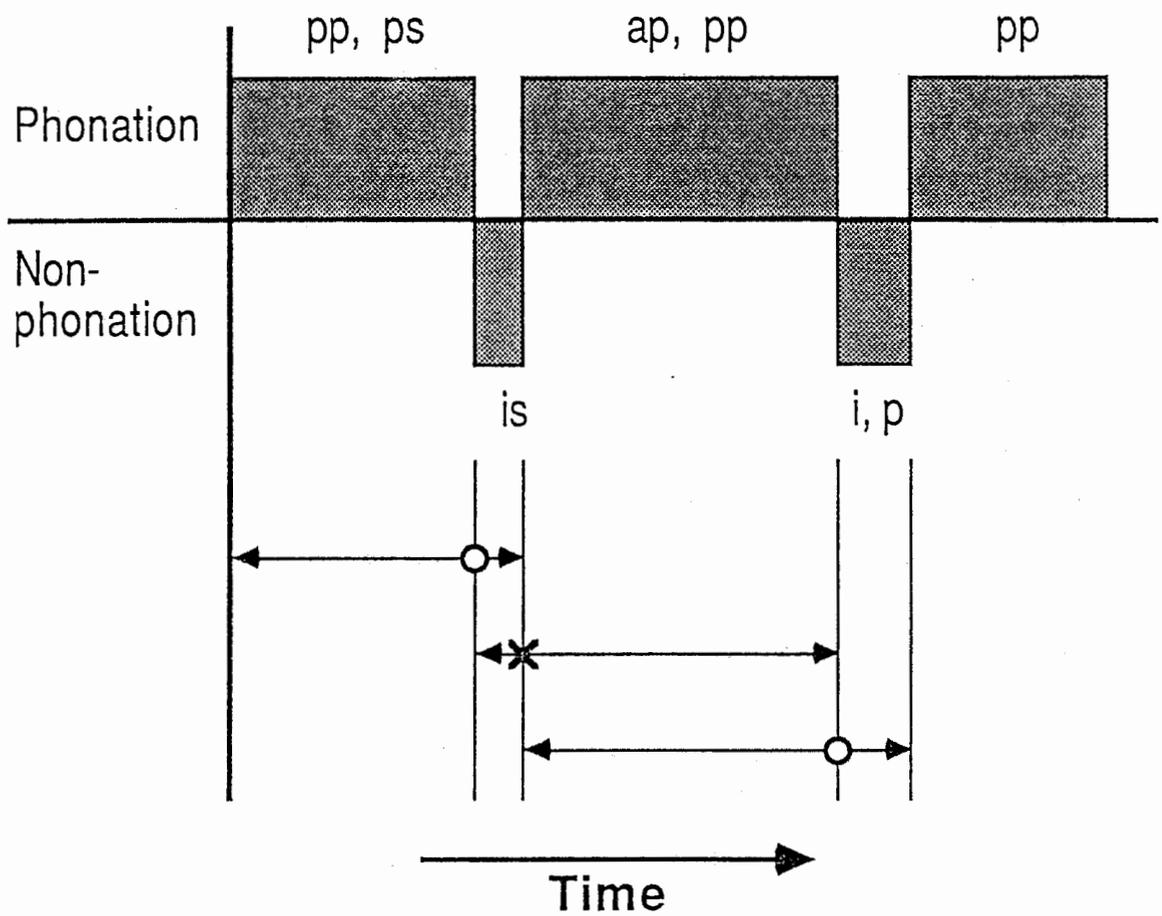


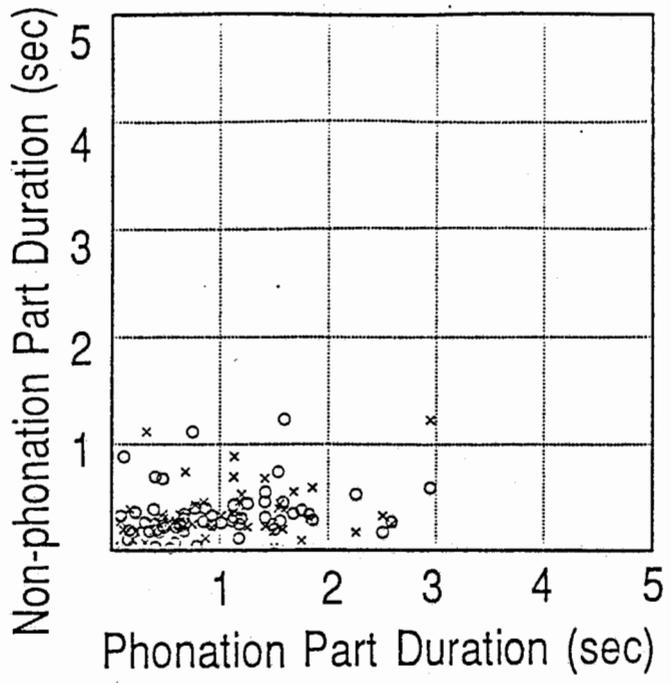
Fig.4 Schematic representation of temporal structure of infant cries using phonation/non-phonation (○) and non-phonation/phonation (×) patterns

**Table 3 Average and SD of fundamental frequency**

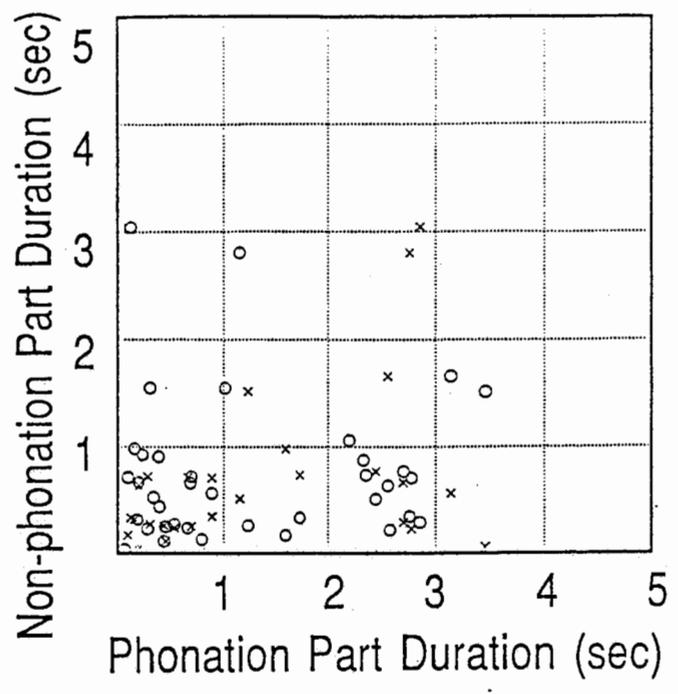
Category	Infant	Age (weeks)	Fundamental frequency	
			Average	SD
Hungry	YU	10	344.0	60.2
	MA	10	384.0	64.4
	MA	12	396.7	63.4
	TE	17	468.8	58.3
Sleepy	MA	10	359.3	50.3
	AS	11	392.3	81.2
	YU	17	359.3	55.9
Call	TE	11	446.2	67.7
	TA	12	383.6	42.2
	YU	15	354.2	41.3
	MA	15	370.6	41.6
Angry	TO	9	455.0	65.3
	MA	12	414.6	50.8
	MA	13	415.5	57.9

**Table 4 Average of fundamental frequency  
for four cry categories by infants,  
MA and YU**

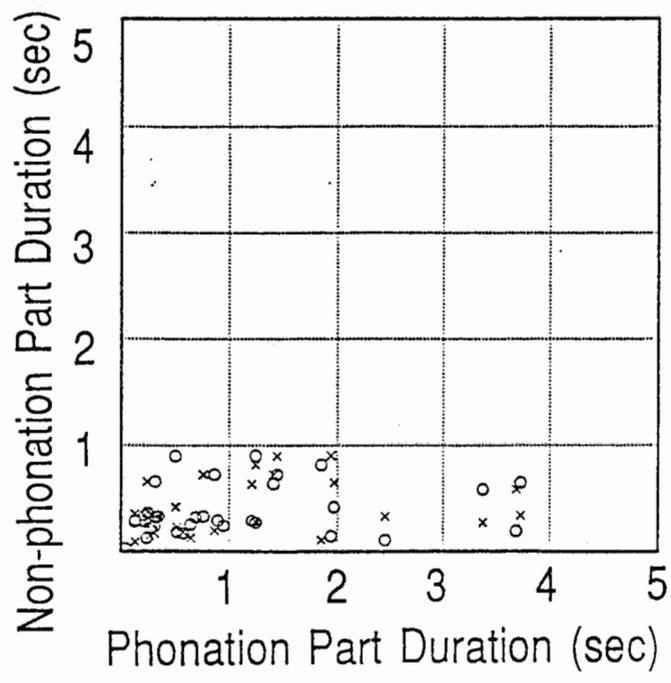
Category	Ave. of fundamental frequency (Hz)	
	MA	YU
Hungry	390.4	344.0
Sleepy	359.3	359.3
Call	370.6	354.2
Angry	415.1	—



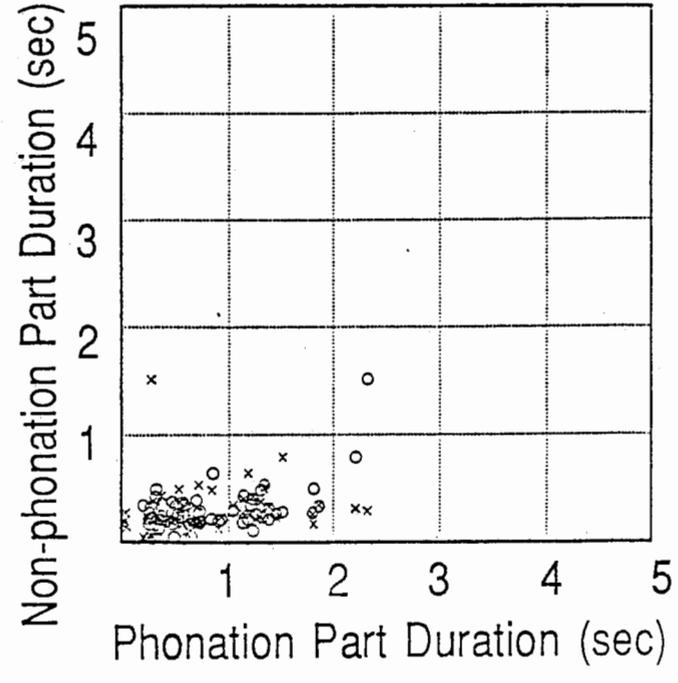
(a) "hungry" cry



(a) "call" cry

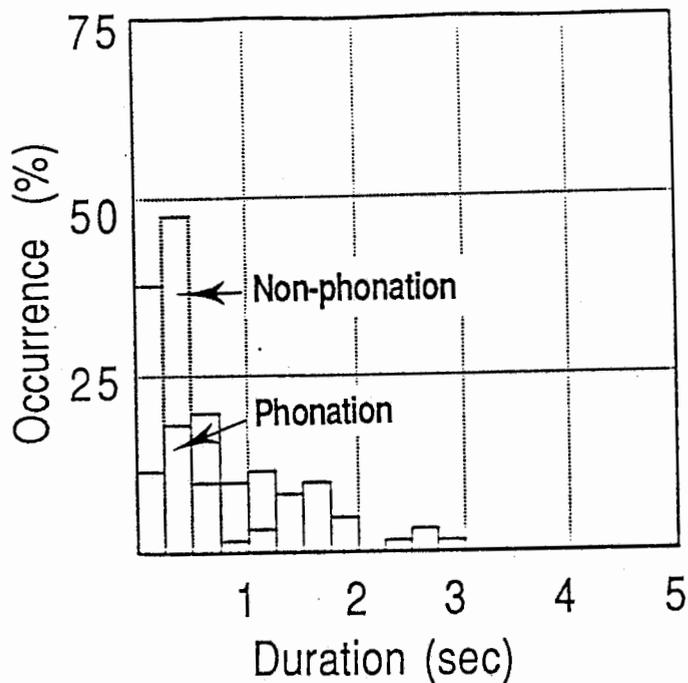


(a) "sleepy" cry

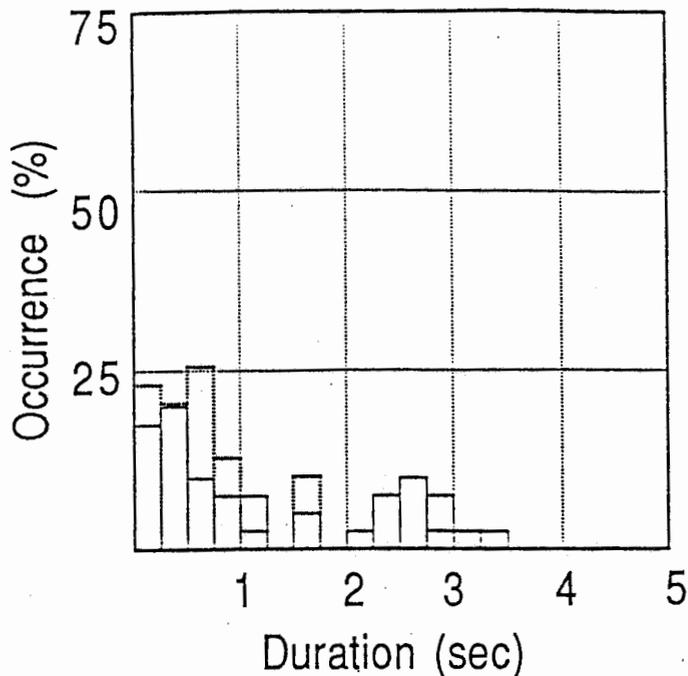


(a) "angry" cry

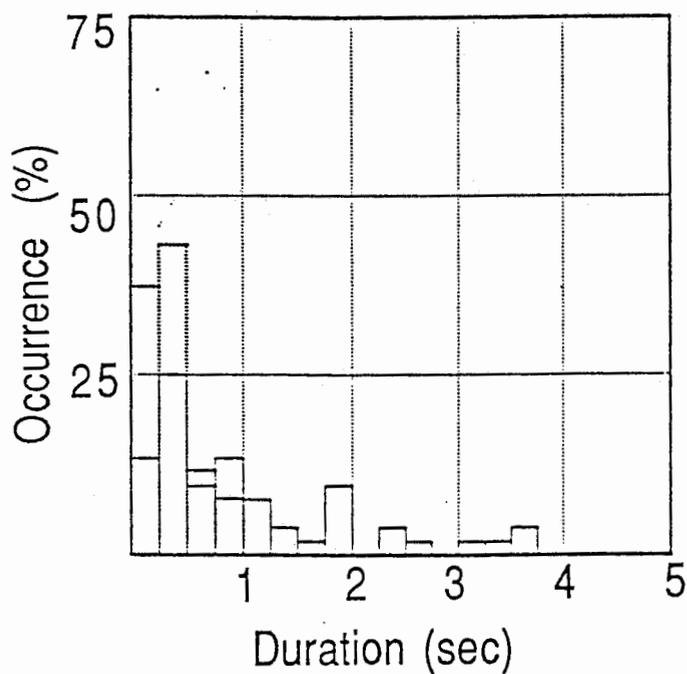
Fig.5 Distributions of phonation/non-phonation part duration for four cry categories



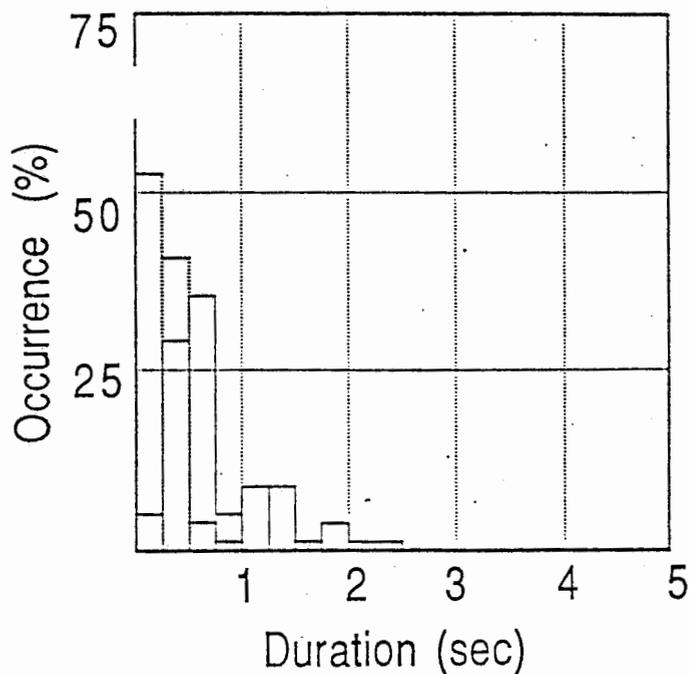
(a) "hungry" cry



(a) "call" cry



(a) "sleepy" cry



(a) "angry" cry

Fig.6 Occurrences of phonation/non-phonation part duration for four cry categories

## 第 2 章 泣き声のカテゴリー判断における単位

【はじめに】

泣き声は、環境に対して養育行動を解発させる信号性をもっているが、それによって泣きの原因を短絡的に判断することは困難である。そこで、従来の泣き声の知覚判断に関する研究は、「泣き声という音響的な情報のみで乳児の泣きの原因をどの程度判断できるか」という点に関心が向けられてきた。しかし、そこには、まだ一致した見解が得られていない。その理由として、Murry (1980) も指摘しているように、実験に用いた泣き声がどのようなものであったかという統制がなされていないことが考えられる。例えば、泣きの原因を明確に説明できる操作的な実験状況を用いた研究が多いが、そのために泣きのカテゴリーが操作的に生じさせられた泣きのものに限られている。また、そのように操作的実験場面において生じた泣きが、必ずしも実験者が意図するものであったかどうかについて疑問の余地がある。そして、泣き声が乳児のどの時期のものであったかによって、同じカテゴリーの泣き声であってもその音響特徴に差異があることが考えられる。さらに、知覚実験においてどのくらいの継続時間の泣き声を被験者に聞かせることが必要であるのか、また、どのような音響特徴をもつ部分を聞かせることがより有効であるのかといった検討がなされていない。つまり、実験サンプルに関して、

- (1) どのような状況（実験状況／自然状況）で収録したものを聞いたか、
- (2) どのような泣き声のカテゴリーを用いたのか、
- (3) 発達的に乳児のどの時期のものであったか、
- (4) どのくらいの継続時間で、どのような音響特徴をもっている部分か、

という4点がそれぞれの研究において異なるからである。そして、特に(4)の問題は泣き声に対する判断を扱う研究において非常に重要な意味をもっていると考える。

ところで、われわれが泣き声をどの程度正確に判断できるかという問題とは別

に、母親が泣き声をカテゴリーカルに判断できるということが報告されている。例えば、Wolff (1969) は、母親が乳児の姿が見えないという実験状況下において、あらかじめ録音された乳児の空腹時の泣き声と痛みを与えた時の泣き声を聞き分けて反応したことを述べている。また塚本 (1983) は、生後3、4カ月の同一男児の「空腹」「不機嫌」「眠い」「甘え」「痛い」という5つの状況の泣き声を用いた弁別と同定により2つの実験結果から、母親がこれらの泣き声をカテゴリーカルに判断したことを報告している。一方、泣き声を分析した研究からは、例えば、乳児の呼吸の周期を単位とした分析から、乳児が一呼吸で発声した音声区間には、泣きの状況に対応したメロディータイプがあることが確かめられている (Wasz-Höckert et al., 1968 ; Rosenhouse, 1977 ; D'Odorico, 1985)。そして、乳児の一呼吸という小さな単位でなく、より包括的に分析した研究からは、泣き声の発声/非発声区間による時系列パターンにおいてカテゴリー差異があることが見いだされている (小林, 1986 ; 塚本, 片桐, 1988, 第1章参照)。以上から、乳児の泣き声には個体差をこえて泣きの状況の特徴づける音響特徴が含まれており、母親はそれらを手がかりにしてカテゴリー判断をしたことが推察される。

では、泣き声のカテゴリー判断において手がかりとなる示差的音響特徴はいかなるものだろうか。その問題については、まだ十分な研究が行われていない。それを解明していくうえで、先に述べた(4)の問題は極めて重要であり、知覚判断とサンプルの物理的性質の関係として確かめ検討することが必要である。したがって本研究の目的は、

(1) 泣き声のカテゴリー判断に必要とされる泣き声の単位(ユニット)を確かめるとともに、

(2) その単位の音響特徴を分析することによって、カテゴリー判断における手がかりが何であることを検討する。

#### 【方法】

泣き声の収録と分類 日常の母子相互交渉の観察場面において生じた4人の乳児(YU, TE, TO, MA)の泣き声は、VTR (SONY SL-F1 ; SONY CCD-V8AF2)によ

って収録した。それらは、泣きの原因となった乳児の状況によって分類された。分類は、VTR録画の再生と観察時における逐一筆記記録などの観察資料をもとに、数名の観察者が乳児の泣きの原因を各々に判断し、評定した結果に基づいておこなわれた。（詳細は、第1章参照のこと。）

サンプル サンプルには、「甘え」「空腹」「怒り」の категорияに分類された生後9週目から17週目の乳児の泣き声を用いた。それらは、計算機（MASSCO MP 5600）を用いてサンプリング周波数 20 kHz で A/D 変換（LPF = 10kHz, -96dB/oct）し、泣き声の知覚判断の実験および音響分析に用いる刺激音とした。

被験者 健常聴力をもつ大学生、大学院生および社会人。うち、男性23名、女性29名。年齢は18才から40才である。

実験 被験者に泣き声の何らかの示差的音響特徴を学習させるという意図のもとに、テスト試行の前に学習試行を施行した。

#### <学習試行>

1 刺激音 学習試行で用いる刺激音の内容を Table 1 に示す。「甘え」「空腹」「怒り」の categoriaごとに1つずつ合計3つのサンプルを選択し、刺激音として用いた。

categoria判断において手がかりとなる音響特徴には、segmental なものと supra-segmental なものの2種類があると考えられる。そこで、刺激音の継続時間は、これらの2種類の音響特徴を十分に含んでいると思われる長さを用いた（継続時間は、26～30秒程度である）。また、これら3つの刺激音にはすべて異なる乳児（YU, TE, TO）のサンプルを用いた。

2 手続き 実験手順は以下に示すとおりである。

(1) 刺激音を「甘え」→「空腹」→「怒り」の順で2度ずつ繰り返し呈示した。このとき、被験者にはこれらの刺激categoriaを前もって知らせ、categoria学習を促した。

(2) (1)と同じ刺激音をランダム順に2度ずつ呈示し、その泣き声などのカテゴリーのものであったかを判断させた。判断には、「甘え」・「空腹」・「怒り」の3つの選択肢を与え、強制選択させた。なお、この判断結果は(1)における学習達成の指標として用いた。

(3) 確認過程として被験者に刺激カテゴリーを知らせたうえで、もう1度、刺激音を「甘え」→「空腹」→「怒り」の順で呈示した。

### <テスト試行>

1 刺激音 テスト試行で用いる刺激音の内容を Table 2 に示す。刺激音には、学習試行で用いた3名の乳児とは異なるひとりの乳児(MA)の「甘え」「空腹」「怒り」の状況における3つのサンプルを用いた。これら3サンプルは、それ自身も刺激音(これを全分節ユニットと呼ぶ)として用いた。また、これら全分節ユニットの刺激音における継続時間は、学習試行で用いた刺激音の継続時間にほぼ対応する長さとした。次に、全分節ユニットの刺激音は、乳児の一呼吸による発声を基準とする分節に分割された。分割は音声波形編集用ソフトウェアを搭載した計算機上で行い、泣き声の波形とサウンドスペクトログラムに基づいて視察により行った。またその際、波形による音声区間をD/A変換し、聴覚的にも確かめて行った。乳児の一呼吸での発声は、乳児が呼吸した直後の発声から再び呼吸した直後の発声の前までとし、それを1分節とした(参照例を Fig.1 に示す)。この結果、上記サンプルに関し、「甘え」は9分節(26.4秒)、「空腹」は18分節(22.7秒)、「怒り」は16分節(25.5秒)に分割された。これらの各分節は、それ自身1分節ユニットの刺激音として用いられた。さらに、順序は保持した条件で分節を組合せ、2分節ユニット、3分節ユニット、5分節ユニット、7分節ユニットからなる刺激音を作成した。その際、Fig.2に示すように、2分節ユニット、3分節ユニットのものは1分節ずつ、5分節ユニットのものは2分節ずつ、そして7分節ユニットのものは3分節ずつ時間的にシフトさせた。したがって、1、2、3、5、7分節ユニットのそれぞれに対し、43、43、43、22、15個、計166ユニットが作成された。この166ユニットと3つの全分節ユニットの刺激音を併せ、計169ユニットの刺激音を実験に用いた。

2 手続き テスト試行では、169の刺激音をランダム順に呈示したが、刺激系列は先頭に10個のダミー刺激を付加し、合計179の刺激音で構成した。なお、呈示の順序効果を考慮して異なったランダム順による3種類の刺激系列を作成した。被験者はカウンターバランスし、それぞれ3種類のランダム順刺激系列のうちの一つが呈示された。学習試行の(2)と同様に、被験者は各刺激音を聞いた後、与えられた「甘え」・「空腹」・「怒り」の3つの選択肢のいずれかを強制選択した。

以上の実験において、刺激系列のISIは4秒であった。また、実験はSound Boothを用いて行い、刺激系列を計算機よりDAT(DIGITAL AUDIO TAPE DECK、SONY DTC-1000ES)を介して、ヘッドフォン(STAX SRM-1 /MK-2)により被験者の両耳に呈示した。その際の呈示強度は、被験者が聞きやすい、泣き声として自然な強度とした。実験所要時間は、学習試行が約10分、テスト試行が約40分であった。

#### 【結果・考察】

正判断率 学習試行の結果によって、まず、被験者を2群に分類した。すなわち、すべての刺激音を正しく同定できたものをL群、そして完全な同定ができなかったものをN群として分類した。L群は44名で全体の85%であり、N群は8名であった。ここで、学習試行は被験者に泣き声の示差的特徴を学習させる意図のもとに行われたことを考慮すると、L群は呈示された3つの刺激音における何らかの示差的音響特徴を学習ができたものであり、N群はそのような学習ができなかったものと考えられることができる。したがって、N群は分析の対象から外した。

L群は、テスト試行における3つの全分節ユニットの判断結果によってさらに2群に分類した。すなわち、学習試行で用いた刺激音に対応すると思われる全分節ユニットの刺激音にすべて正判断できたものをA群、そしてそのうちの一つでも誤判断したものをB群とした。A群は20名、B群は24名であった。

Fig. 3に、A群とB群それぞれにおいて、各分節に対する正判断の比率を3つのカテゴリー別に、また、ユニットの分節数ごとに示した。この結果、A、B

両群ともに「甘え」においては、ユニットの分節数とは無関係に、1ユニットであっても非常に高い正判断率が得られた。しかし、「空腹」と「怒り」においては、A群では分節数の増加とともに正判断率の上昇がみられたが、B群では安定した上昇がみられなかった。そこで、A群の判断の性質を知るために、各カテゴリーごとに誤判断の傾向を調べ、Fig. 4に示した。その結果、ユニットの分節数の増加に伴って誤判断される比率は減少し、特に「空腹」を「甘え」、「怒り」を「空腹」とする誤判断の場合にこの傾向が顕著であった。また、全体としては「空腹」と「怒り」の間の誤判断が多かった。B群においては、「空腹」と「怒り」の間の誤判断がさらに多く、また、誤判断がユニットの分節数の増加に伴って減少する傾向はA群ほど顕著に認められなかった。

以上のことから、A群とB群は、学習試行において泣き声の3つのカテゴリーにおけるそれぞれの示差的音響特徴を学習したという点においては同様であったが、学習した示差的音響特徴の内容が異なっていたことが推察される。そして、A群が学習した内容は、同一乳児による3つのカテゴリーの泣きを判断するというテスト試行において利用することができたものであるのに対して、B群が学習した内容は、テスト試行に利用できないものであったということが考察される。

テスト試行における被験者の反応は、泣き声の何らかの示差的音響特徴を学習しカテゴリーカルな判断をしたと思われるA群においてのみさらに詳細に分析された。まず、[カテゴリー] × [ユニットの分節数]を2要因とした枝分かれ配置の分散分析をおこなったところ、どちらの主効果においても有意差がみられ ( $P < .001$ )、A群が有意に泣き声をカテゴリー判断したことが示された。また、カテゴリーにおいては「甘え」、「怒り」、「空腹」の順にカテゴリー判断されやすく、その正判断率はユニットの分節数の増加にしたがって上昇することが示された。上昇は、「甘え」に関しては2分節ユニット、「空腹」と「怒り」に関しては3～5分節ユニットで一時的に飽和し、その後緩やかに上昇することが示された。ここで1分節の継続時間が、カテゴリーによって異なることに着目すると、3つのカテゴリーとも平均的に6～8秒程度となり (Table 2 参照)、これがカテゴリー判断に必要とされる最小単位であることが示された。

次に、判断に混同がみられた「空腹」と「怒り」のカテゴリーをとりあげて、各ユニットごとの正判断の比率を調べた。「空腹」に関する例を Fig. 5に示す。

その結果、1分節ユニットにおいても非常に高い正判断率を得た分節があり、ユニットの分節数の増加にしたがってこのような分節を多く含む多分節ユニットの正判断率も高くなることが示された。これとは逆に、ユニットの分節数が増加しても、1分節ユニットにおいて非常に低い正判断率しか得られなかった分節を多く含む多分節ユニットの正判断率は上昇がみられないことが示された。このような現象は、「怒り」より「空腹」において顕著であったが、これは「空腹」では低い正判断率しか得られなかった1分節ユニットが時間的に連続している傾向がみられたのに対して、「怒り」では1分節ユニットにおいて低い正判断率しか得られなかったものが少ないこと、また、これに関連してこれらの1分節ユニットが時間的に連続せず、前後に高い正判断率をもつ分節が存在したためと思われる。このため、多分節ユニットにおいて高く正判断されたユニットが「空腹」より「怒り」に総体的に多かったものと思われる。

**音響特徴** 泣き声の知覚におけるカテゴリー判断は、どのような音響特徴に基づいてなされるのであろうか。ここでは、カテゴリーの正判断率が分節に依存して変化する点に着目して、パラメータには基本周波数 ( $F_0$ ) とパワーを取り上げて分析した。

まず、全分節ユニットに対して求めた  $F_0$  とパワーの平均値を Table 3 に示す。 $F_0$  (Hz) は、ケプストラム分析を使ったピッチ抽出法によって 2.5 ms ごとに求め、また、パワーは、20 ms ハミング窓を用いて、2.5 ms ごとに dB 値で求め、それぞれカテゴリーごとに平均値と標準偏差を算出した。 $F_0$  の平均値は「甘え」「空腹」「怒り」の順に高くなり、それぞれ 370.7 Hz、392.8 Hz、437.3 Hz であった。パワーの平均値は「甘え」「怒り」「空腹」の順に大きかったが、差異は少なかった。

次に、「空腹」と「怒り」のカテゴリーにおいて、正判断率が 80% 以上のユニットと 50% 以下のユニットの  $F_0$  とパワーを分節数ごとにまとめ、その差異を調べた。「怒り」の例を Table 4 に示す。ここで示した  $F_0$  とパワーの平均値の算出では、まず、ユニットの平均値を求め、さらにユニットの分節数ごとにユニット平均値の平均を求め、これをそれぞれの分節数に対応したカテゴリー平均値とした。その結果、「怒り」においては、高い正判断率を得たユニットの  $F_0$

の平均値は、436.3 Hz から 443.4 Hz の高い数値を示していた。それに対して低い正判断率しか得られなかったユニットの  $F_0$  の平均値は、414.7 Hz から 423.7 Hz の低めの数値を示しており、「空腹」の全分節ユニットの平均  $F_0$  値に近い数値を示していた。「怒り」に対する誤判断の大部分が「空腹」への誤りであったことから、 $F_0$  の値が判断の手がかりになった可能性が示唆される。ところが、 $F_0$  の平均値が高くてパワーの値が低い場合やその逆の場合においても高い誤判断率が示されたことから、パワーのユニット平均値が補助的要因として判断の手がかりになっている可能性も推察される。

しかし、泣き声のカテゴリー判断においてどのような音響特徴が手がかりにされたかという問題は、今回の分析に用いられた  $F_0$  とパワーという2つのパラメータからは説明しきれないものもあり、さらに多角的な分析をしていくことが必要である。

#### 【要約】

本実験において得られた結果を以下にまとめる。

(1) 学習試行において、泣き声の示差的音響特徴を学習したと思われたものは、そこで学習した内容をテスト試行に活用できたものとできなかったものに分類された。

(2) そこで、学習した内容をテスト試行に活用できたA群の判断結果を分析したところ、正判断率はユニットの分節数が増加するにつれてそれぞれ上昇することが示された。この上昇傾向は、「甘え」に関しては2分節ユニット、「空腹」と「怒り」に関しては3～5分節ユニットで、一時的に飽和し、その後は緩やかに変化する。すなわち、これらがカテゴリー判断に必要な最小単位の一つと考えると、その継続時間は平均6～8秒であった。

(3)  $F_0$  とパワーという2つの物理的性質が泣き声のカテゴリー判断において1つの手がかりになっている可能性が示唆された。

(4) しかし、 $F_0$  やパワーのユニット平均値だけでは、ユニットに対するカテゴリー判断の結果をすべて説明しきれず、動的な音韻特徴（例えば、Wasz-Höckertのいうメロディータイプ）など、さらに多くの角度から示差的特徴を検討

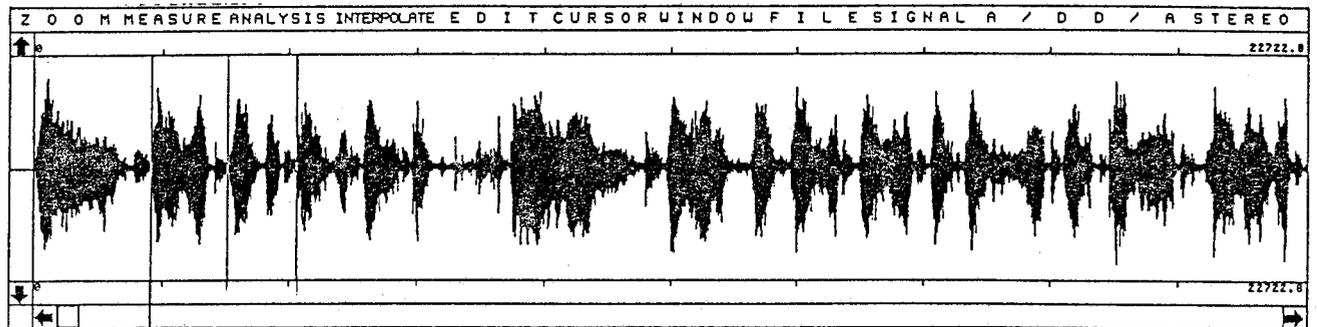
する必要がある。

**Table 1 Cry stimuli used in learning sessions**

Category	Infant [Age in weeks]	Duration (sec)
Call	YU [17]	26.0
Hungry	TE [17]	28.6
Angry	TO [ 9]	26.8

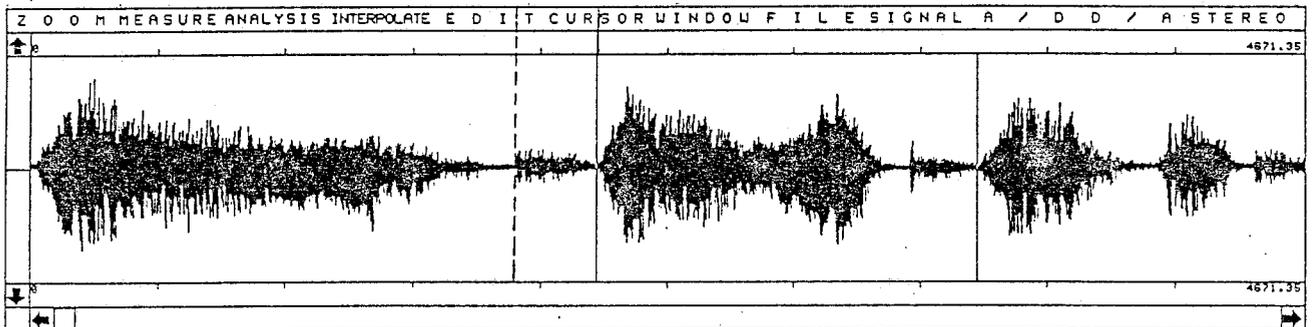
**Table 2. Cry stimuli used in testing sessions**

Category	Infant [Age:weeks]	Duration (sec)	Number of Segments	Segmental Duration (sec)		
				Ave.	Max.	Min.
Call	MA [15]	26.4	9	2.93	8.82	0.53
Hungry	MA [12]	22.7	18	1.26	2.79	0.31
Angry	MA [13]	25.5	16	1.60	3.02	0.52



(a) Amplitude envelope of a full unit of "hungry" cry (infant; MA)

1-segment  
inspiratory sound



(b) 3-segment unit excised from the full unit shown above

Fig.1 Segmentation of cry waveforms. A full unit cry (a) is segmented into an n-segment unit based upon inspiratory sound portions

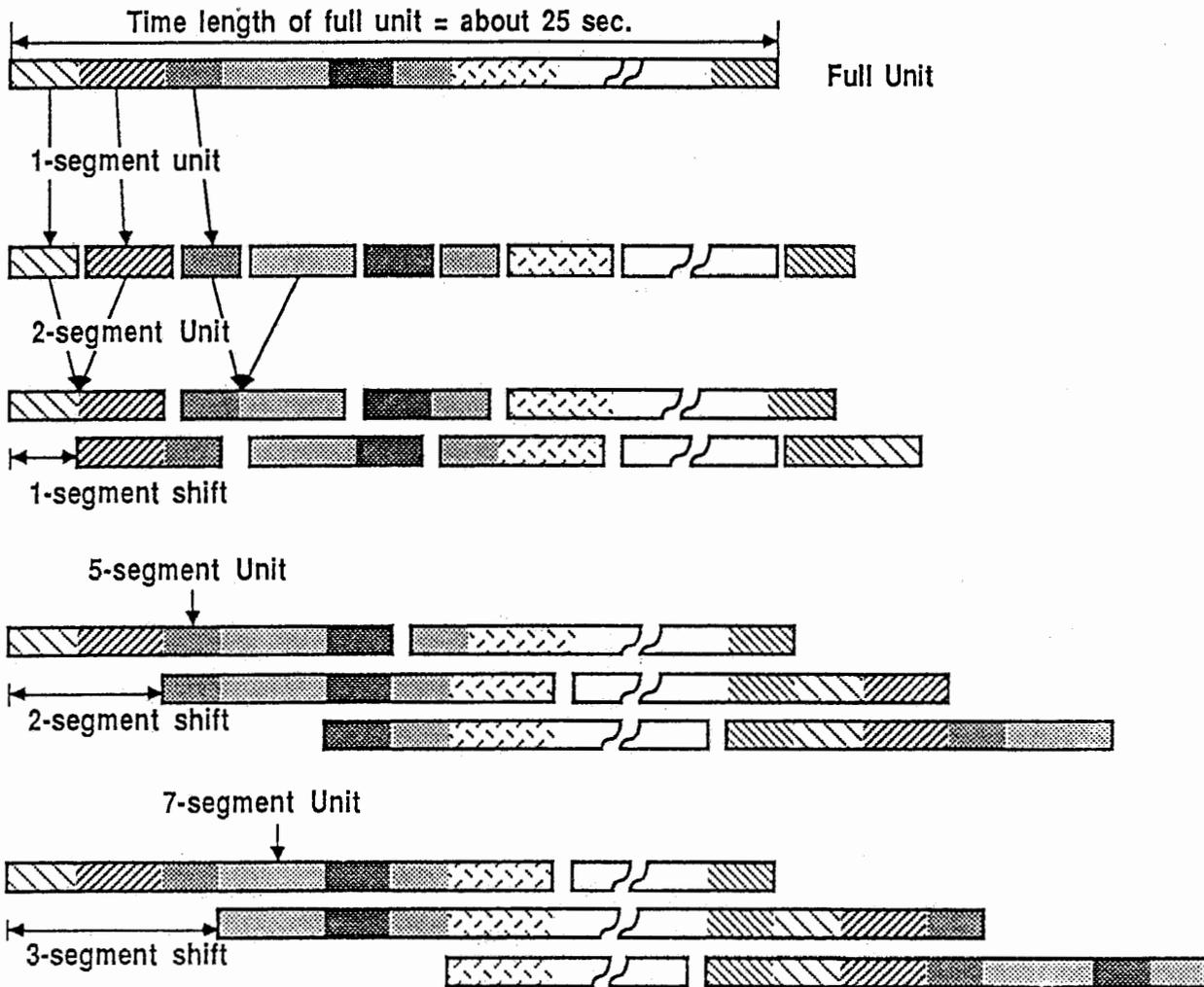


Fig.2 Cry segments and units used in category identification tests. A full cry unit is divided into  $n$  cry segments. A 1-segment unit has only one segment. An  $i$ -segment unit has successive  $i$  segments. In the identification tests, 1-, 2-, 3-, 5-, 7-segment and full units are used as stimuli.

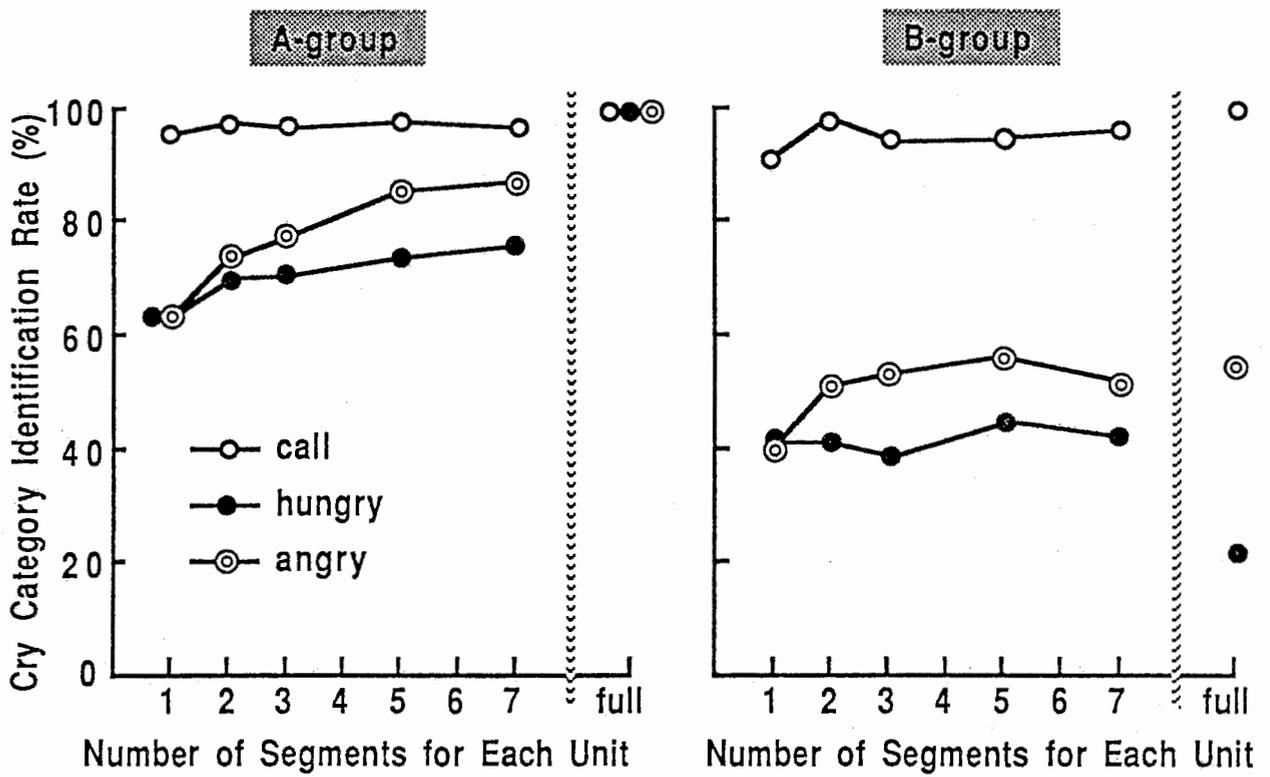
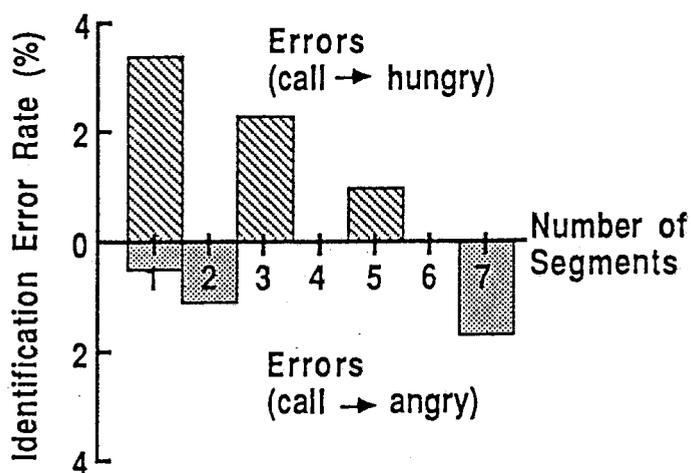
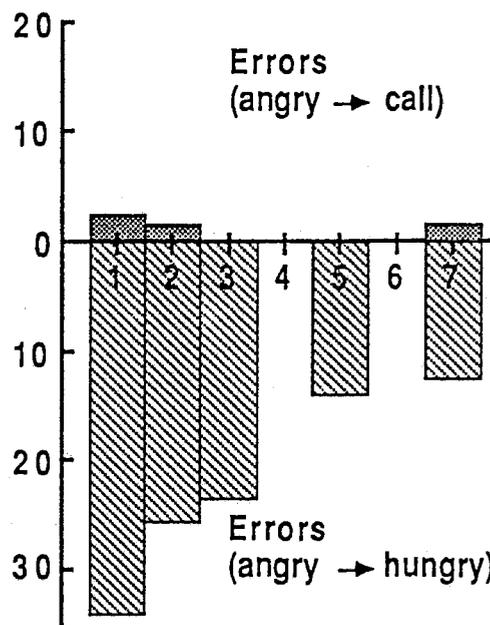


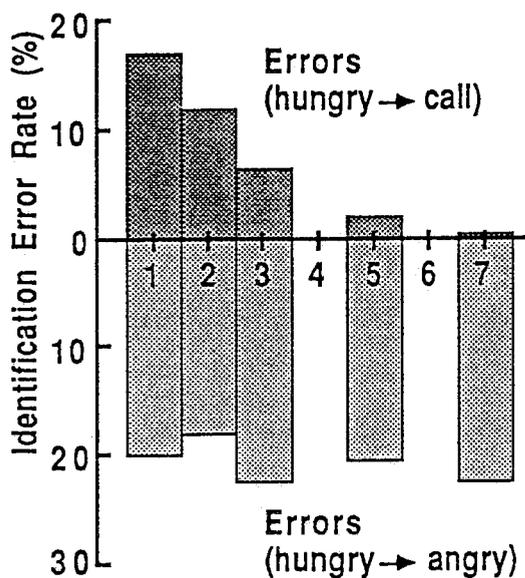
Fig.3 Cry category identification rates for A- and B-group. A-group consists of 20 subjects who correctly identified each full unit of three cry categories by 100%. B-group consists of 24 subjects who failed to identify full unit cries correctly.



(a) Errors for "call" cry



(c) Errors for "angry" cry



(b) Errors for "hungry" cry

Fig.4 Error properties in identifying cry units of three categories for A-group. Error rates are shown as a function of the number of segments for each cry unit. Two kinds of errors -- misjudgement of "hungry" as "call" and misjudgement of "angry" as "hungry" -- decrease significantly as the number of cry unit segment increases.

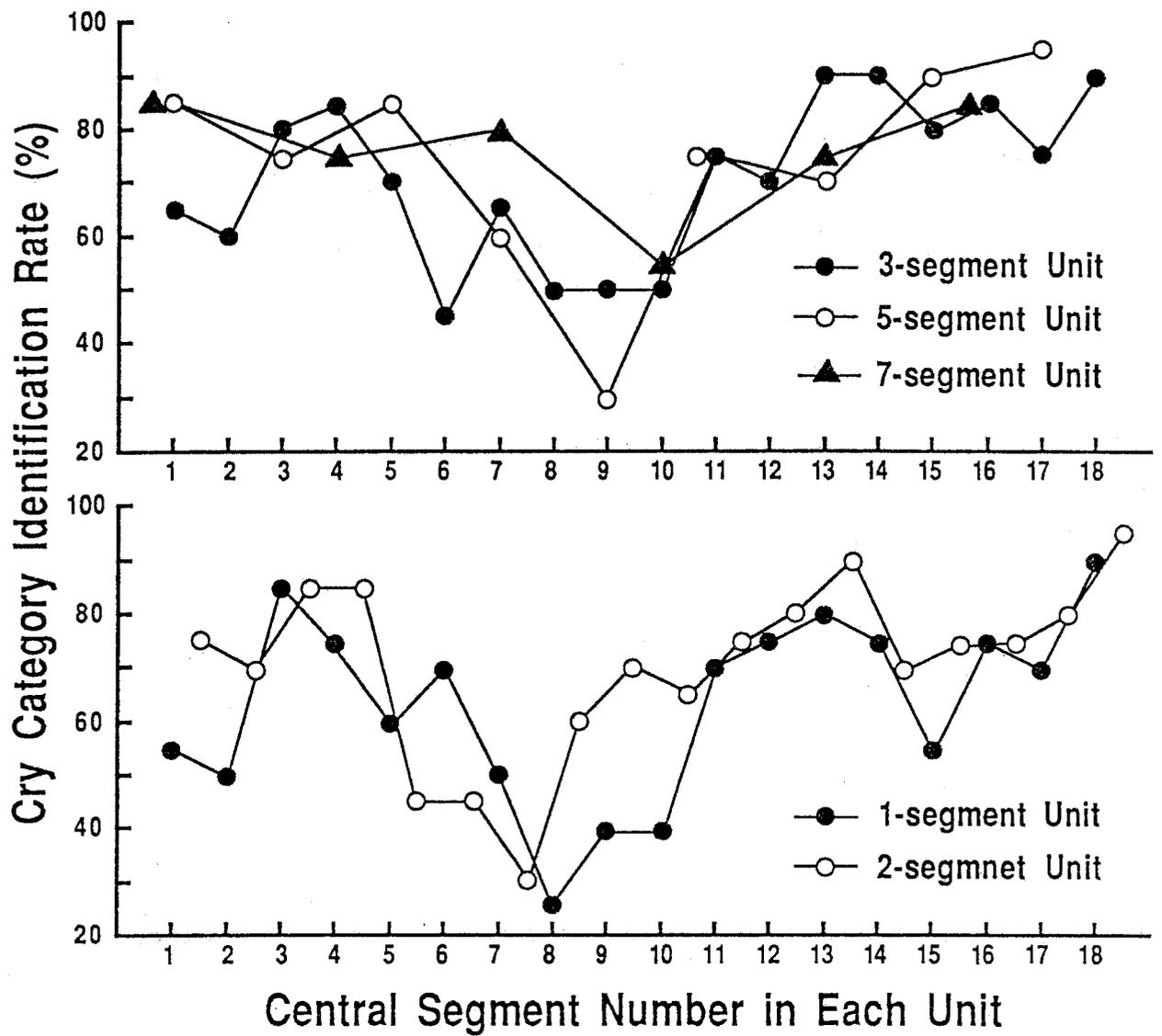


Fig.5 Category identification rates for "hungry" cries as a function of segment numbers in each unit for each of 1-, 2-, 3-, 5- and 7-segment units.

**Table 3 Fundamental frequency ( $F_0$ ) and power of full unit cries**

Category	$F_0$ (Hz)	Power (dB)
Call	370.7	65.8
Hungry	392.9	67.2
Angry	437.3	66.6

**Table 4 Fundamental frequency ( $F_0$ ) and power comparing "angry" cry units with high identification rates and those with low identification rates**

(a) "angry" cry units with high identification rates ( $\geq 80\%$ )

Number of Segments	$F_0$ (Hz)			Power (dB)		
	Mean	Max	Min	Mean	Max	Min
1	444.3	460.1	408.6	66.7	69.3	64.5
2	438.2	453.5	411.3	66.3	68.0	65.3
3	439.4	453.2	424.4	66.1	67.2	65.3
5	436.3	445.0	424.9	66.6	68.0	65.5
7	436.7	441.5	428.4	66.8	67.4	65.9

(b) "angry" cry units with low identification rates ( $\leq 50\%$ )

Number of Segments	$F_0$ (Hz)			Power (dB)		
	Mean	Max	Min	Mean	Max	Min
1	423.7	460.3	397.6	66.4	67.9	65.0
2	414.7	---	---	68.1	---	---

【文献】

- D'Odorico, L. 1984 Non-segmental features in prelinguistic communications : An analysis of some types of infant cry and non-cry vocalization. *Journal of Language*, 11, 1, 17-27.
- 二木恒夫 1979 新生児期における泣き声とその発達の意義 児童精神医学とその領域, 20, 167-177.
- Illingworth, R.S. 1955 Crying in infants and children. *British Medical Journal*, 1, 75.
- 小林健二 1986 新生児の泣き声の時系列パターンの検討 医学電子と生体工学, 24.
- Lester, B.M. 1978 The organization of crying in the neonate. *Journal of Pediatric Psychology*, 3, 3, 122-130.
- Lester, B.M. 1984 Infant crying and the development of communication. In N.A.Fox & R.J.Davidson (Eds.) *The psychobiology of affective development*. Hillsdale, New Jersey, London.
- Michelson, K. 1980 Cry characteristics in sound spectrographic cry analysis. In T. Murry & J. Murry (Eds.), *Infant communication : Cry and early speech*. College-Hill Press.
- Murry, T. & Murry, J. 1980 *Infant communication : Cry and early speech*. Colledge-Hill Press.
- 塚本妙子 1983 乳児の泣きに対する母親の認知 甲南女子大学大学院心理学年報, 2, 31-44.
- 塚本妙子 1985 乳児の泣きの発達とメカニズム 甲南女子大学大学院心理学年報, 4, 1-20.
- 塚本妙子, 片桐 滋 1988 泣き声の音響特徴の定量的分析 日本心理学会第52回大会発表論文集.
- Rosenhouse, J. 1977 A preliminary report : An anlysis of some types of a baby's cries.
- Wasz-Höckert, et al., 1968 The infant cry. A spectrographic and auditory analysis. *Clinics in Developmental Medicine*, 29. Spastics Inter-

national Medical Publications, Heinemann, London.

Wolff, P.H. 1967 The role of biological rhythms in early Psychological Development. Bulletin of the Menninger Clinic, 31, 4, 197-218.

Wolff, P.H. 1969 The natural history of crying and other infant vocalization. In B. M. Foss (Ed) The Determinants of Infant Behavior 4, London, Methuen.

## 【あとがき】

泣き声にはどのような音響特徴があり、そうした音響特徴の何が示差的特徴として知覚的なカテゴリー差異をもたらすかを明らかにするための研究の一環として、泣き声の分析と知覚実験を行った。その結果を、

第1章 泣き声の音響特徴

第2章 泣き声のカテゴリー判断における単位

として報告した。

第1章では、音響特徴のラベルに従った時系列パターンと発声/非発声区間による時間構造の分布においてカテゴリー差異が認められた。そして、第2章では、ユニットの分節数が増加するにつれて正判断率が上昇するが、その上昇傾向からカテゴリー判断に必要な最小単位は各カテゴリーにおいてそれぞれ平均6～8秒の継続時間であることが示された。また、 $F_0$ とパワーという物理的性質がカテゴリー判断における手がかりのひとつになっている可能性が示された。

しかし、泣き声のカテゴリー判断において手がかりにされる示差的音響特徴を解明していくためには、今後さらに泣き声を知覚判断と物理的性質の関係において多くの角度から検討することが必要である。