

# 音声の物理学とそれを通して考える自閉症

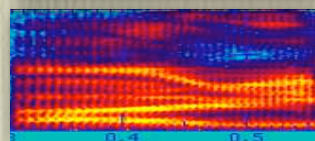
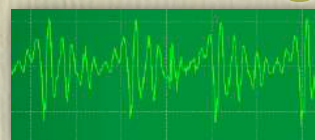
～敏感・鈍感のズレと汎化のズレ～

峯松 信明 & 西村多寿子

## 本日のメニュー

### 🔊 母音とその生成メカニズム

- 口の構え＝舌の位置, 唇の丸め方



### 🔊 音響音声学の「いろは」

- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い

### 🔊 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

### 🔊 鈍感さと敏感さのバランスが崩れるとどうなる？

- 🇲🇻 – 🇵🇹 = ?
- 刺激そのものに敏感に／刺激群の差異に鈍感になるとどうなる？
- 自閉症（＝関係性の病？）に関するいくつかの思考実験

### 🔊 まとめ

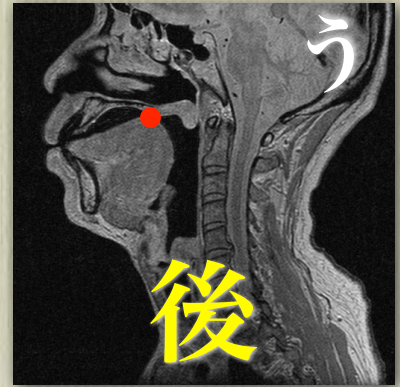
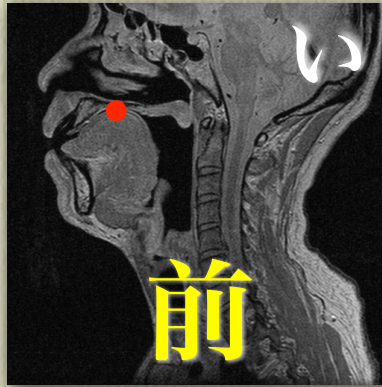
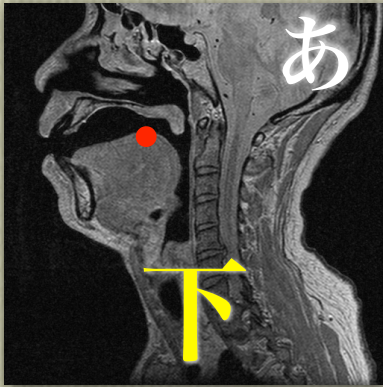


# 母音とその生成メカニズム

## 母音の音声学的定義

- 母音＝肺からの呼気流が妨害を受けずに発せられる言語音
- 子音＝何らかの妨害を受けつつ発せられる言語音

## 母音の生成と分類：どうすれば「あ」は「い」になる？



(提供：ATR人間情報科学研究所)

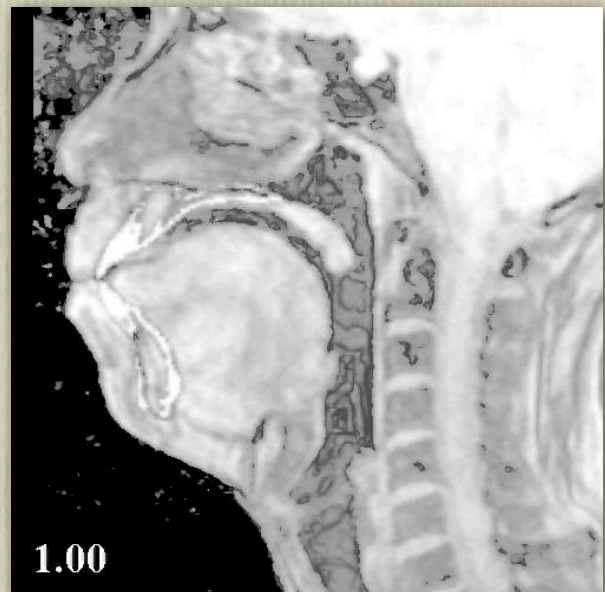
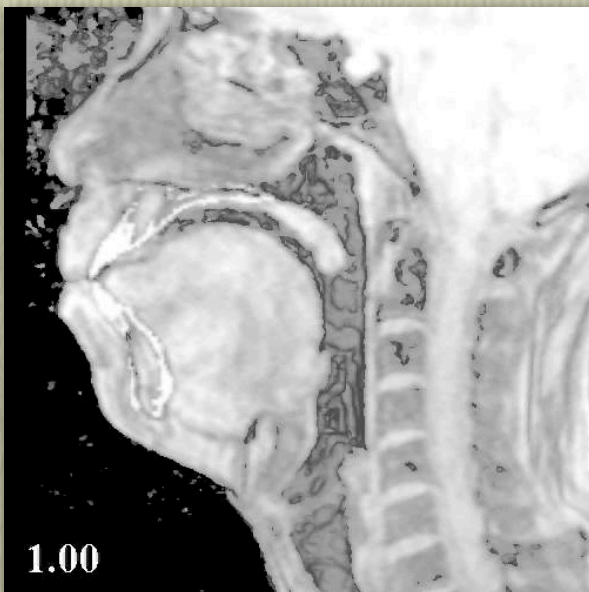
- 口の開き方 (=舌の**高さ**)
- 舌の**前後位置**
- 唇の**丸め**の有無

} 口の構え

# 母音とその生成メカニズム

## 「母音の違い」＝「口の構えの違い」

- ちょっと違った角度からこの等式の意味を考えてみよう。



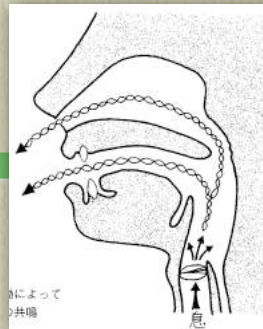
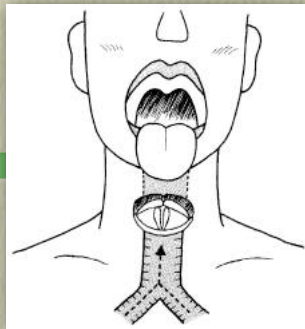
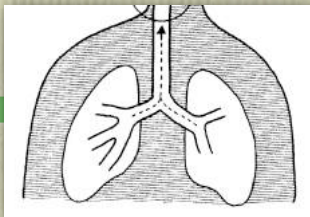
(提供：ATR人間情報科学研究所)

- 「口の構えの違い」→「**口の中のすき間の形状の違い**」



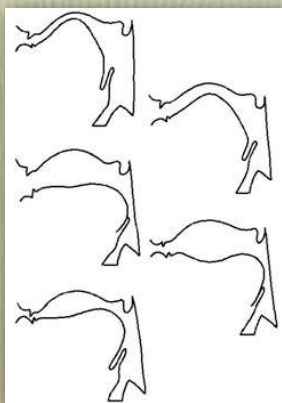
# 声帯で生まれた音源が口を通して母音に

息漏れ「は〜」と母音「あ〜」の違いを考える

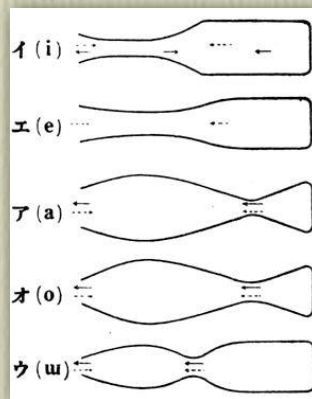


i u  
e o  
a

(出典：萩野&後野「発声のメカニズム」音楽之友社)



曲げ延ばし  
はOK



(出典：I. R. Titze「音声生成の科学」)

## 声帯音源 + 声道 = ブザー + 管



(提供：ATR人間情報科学研究所)





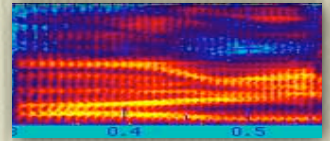
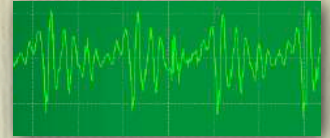
# 本日のメニュー

## 母音とその生成メカニズム

- 口の構え＝舌の位置，唇の丸め方

## 音響音声学の「いろは」

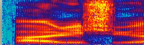
- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



## 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

## 鈍感さと敏感さのバランスが崩れるとどうなる？

-  -  = ?
- 刺激そのものに敏感に／刺激群の差異に鈍感になるとどうなる？
- 自閉症（＝関係性の病？）に関するいくつかの思考実験

## まとめ

# 音の四要素

## 様々な情報＝四要素の組み合わせ

### 音の高さ

- 高い音，低い音

### 音の大きさ

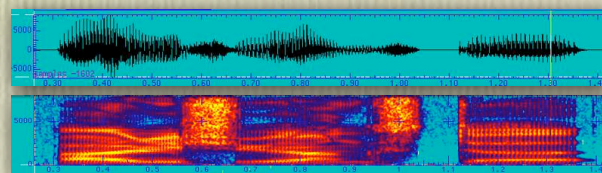
- 大きい音，小さい音

### 音の長さ

- 長い音，短い音

### 音の音色（声色）

- 太い声，細い声，黄色い声，甘い声，渋い声，色っぽい声・・・
- 高さ・大きさ・長さが同じ2音を「違う音」と認識した場合，その2音は音色が異なる。
- 「あ」と「い」      「あ」と「あ」



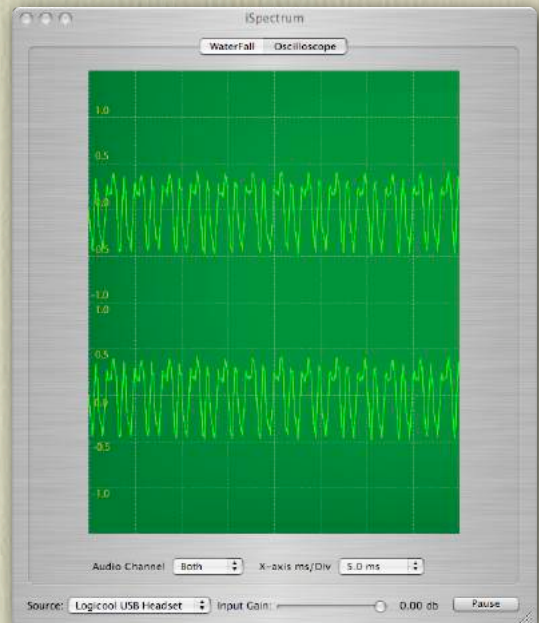
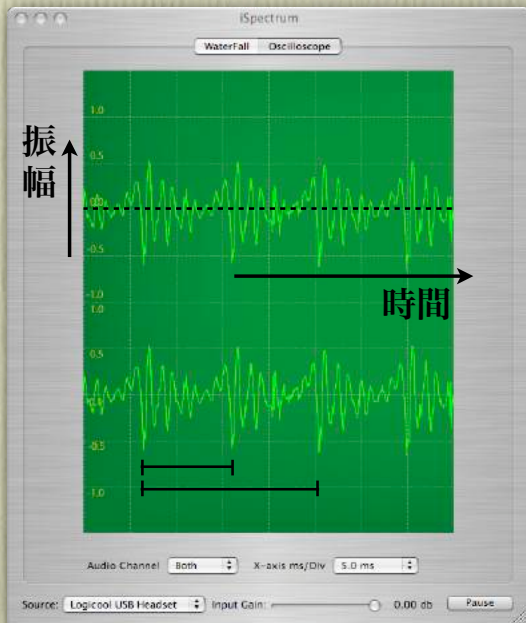
- 1) 高さ， 2) 大きさ  
3) 長さ， 4) 音色

ある情報



# 波形の中に見る高さ

## 波形表示ソフトで見る「低いあ」と「高いあ」

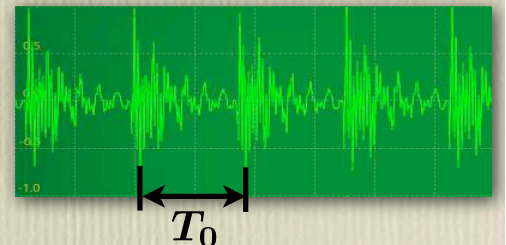
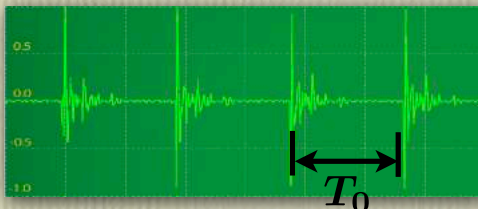
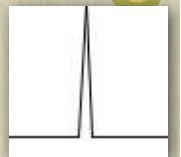


- 同じものが繰り返される時に使われる単位：周波数（振動数）
  - = frequency = 回/秒 = サイクル/秒 = Hz（ヘルツ, Hertz）
  - 1秒当り5回=5Hz, 4秒当り1回=0.25Hz, 1/16000秒当り1回=？

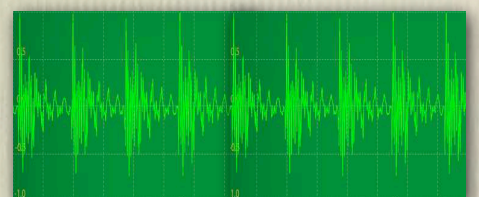
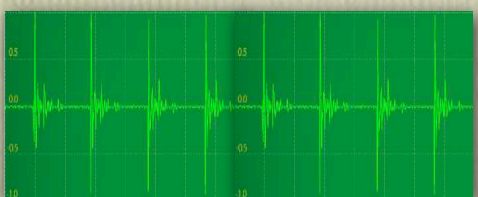
## 高さの違いはブザーの間隔の違い

### ブザー（パルス列的）→ 管 → 「あ～」

- パルス＝瞬時的に振幅が大きくなり、すぐ減衰する波形
- 男性の場合：波形の間隔が長くなる。
  - 1秒当たりのパルスの回数（振動回数）は低くなる。約 100 Hz



- 女性の場合：波形の間隔が短くなる。
  - 1秒当たりのパルスの回数（振動回数）は高くなる。約 200 Hz



- $T_0$  = 基本周期,  $F_0 = (1/T_0)$  = 基本周波数 [Hz]



# 音の四要素

様々な情報＝四要素の組み合わせ

音の高さ

- 高い音, 低い音

音の大きさ

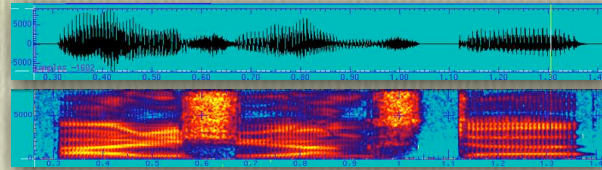
- 大きい音, 小さい音

音の長さ

- 長い音, 短い音

音の音色 (声色)

- 太い声, 細い声, 黄色い声, 甘い声, 渋い声, 色っぽい声...
- 高さ・大きさ・長さが同じ2音を「違う音」と認識した場合, その2音は音色が異なる。
- 「あ」と「い」      「あ」と「あ」



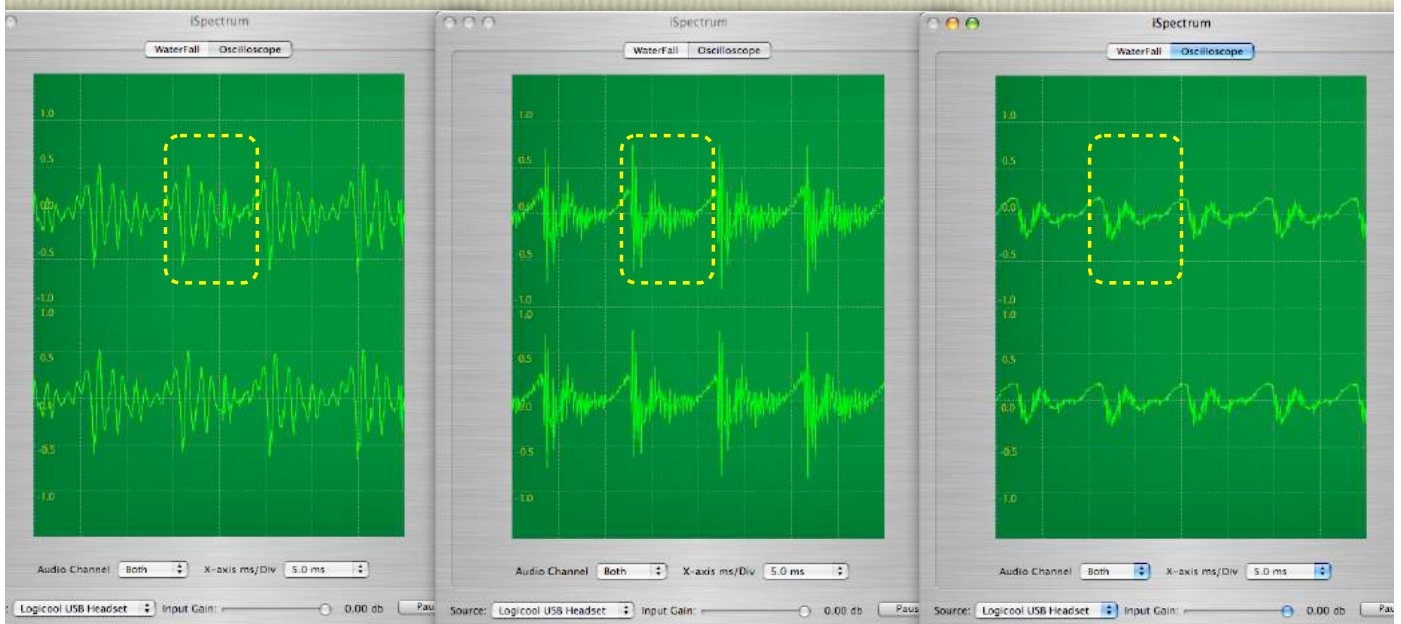
1) 高さ, 2) 大きさ  
3) 長さ, 4) 音色

ある情報

## 波形の中に見る音色

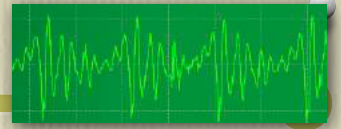
波形表示ソフトで見る「あ」と「い」と「う」

- 同じ高さ = 同じ周期 = 一秒間当りの振動回数が同じ
- 一つ一つの波形形状の違いが音色の違い？



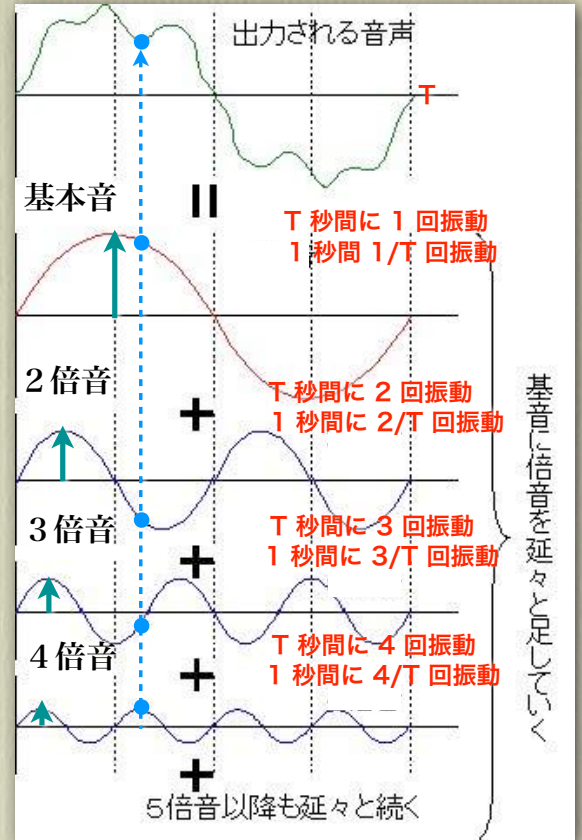
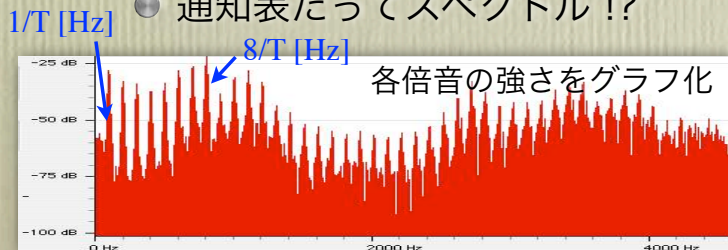


# 波形の分解と合成

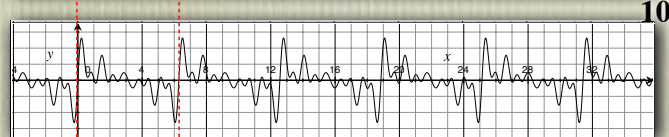
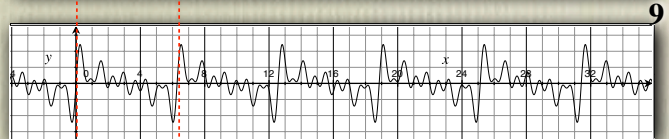
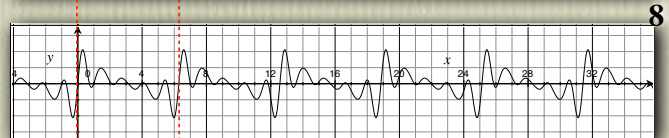
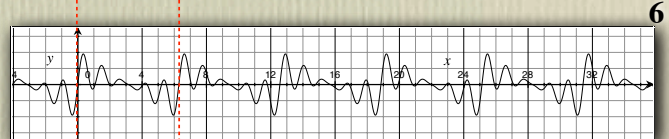
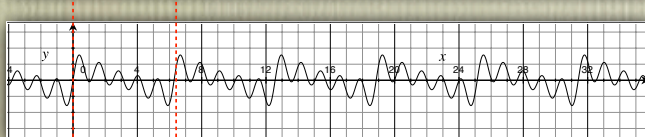
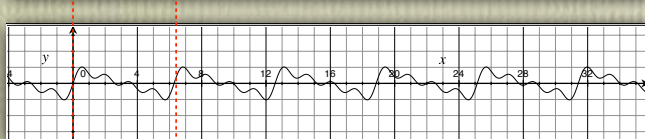
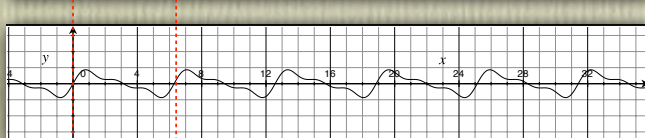
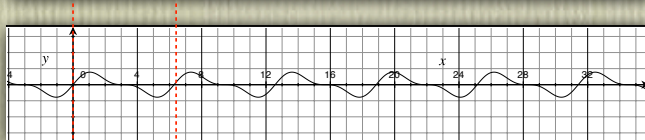
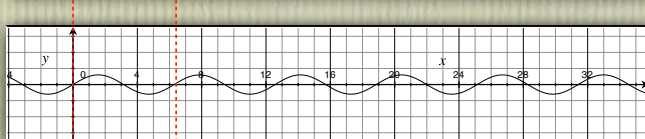
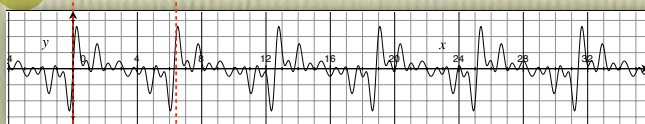


## 基本音とその倍音の足合わせ

- 波 = 基本音 + 2倍音 + 3倍音 + ...
- n倍音: n倍の周波数のサイン波形
- 周波数: 振動回数/秒 [Hz]
- 波 = これらを適切な強さにして足しあわせた結果
- どの周波数のサイン波は強く, どの周波数のサイン波は弱いのか?
- 横軸を周波数, 縦軸を強度としてグラフを書く → **スペクトル**
- 通知表だってスペクトル!?

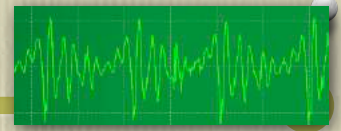


# 波形の分解と合成



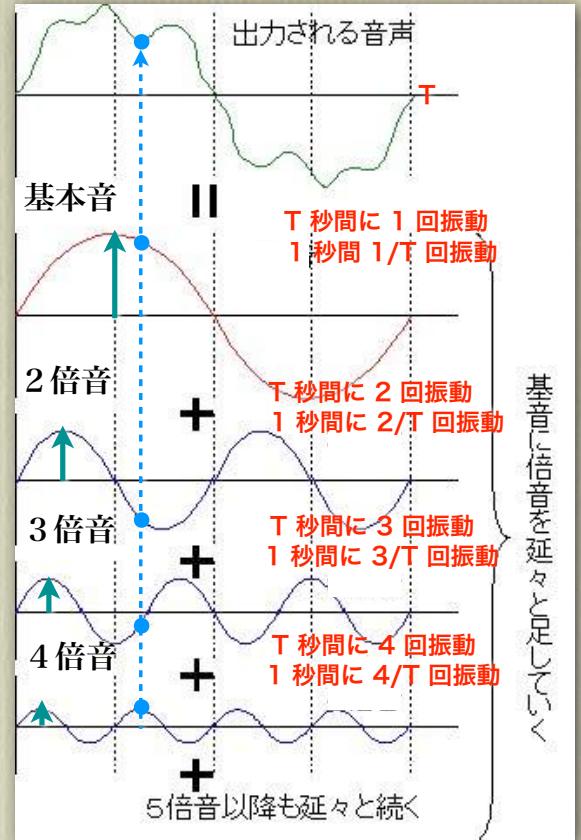
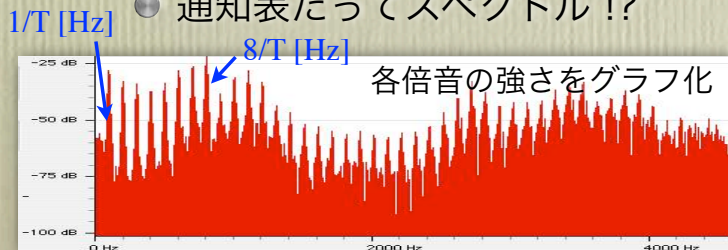


# 波形の分解と合成



## 基本音とその倍音の足合わせ

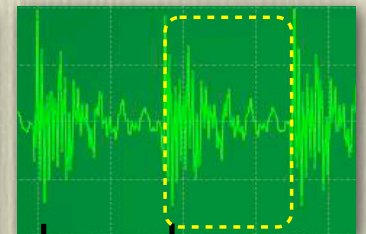
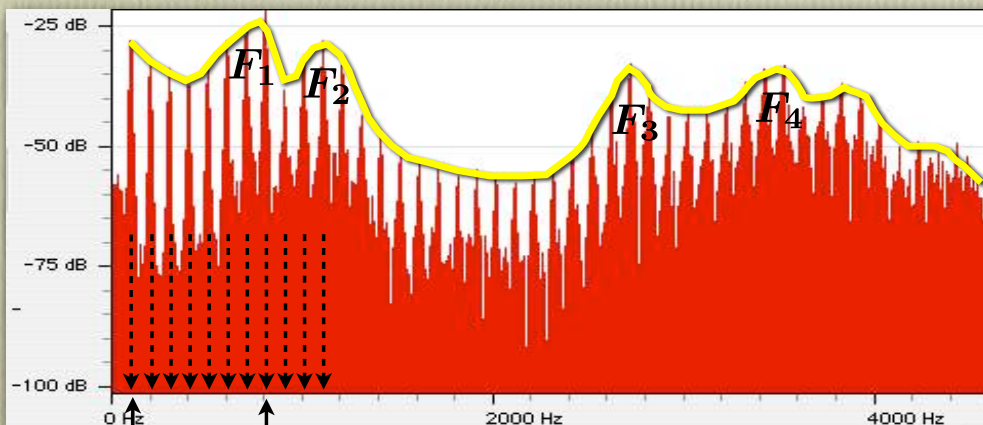
- 波 = 基本音 + 2倍音 + 3倍音 + ...
- n倍音：n倍の周波数のサイン波形
- 周波数：振動回数/秒 [Hz]
- 波 = これらを適切な強さにして足しあわせた結果
- どの周波数のサイン波は強く、どの周波数のサイン波は弱いのか？
- 横軸を周波数、縦軸を強度としてグラフを書く → **スペクトル**
- 通知表だってスペクトル!?



## スペクトルに見る高さや音色

### 高さ = ぎざぎざ, 音色 = 山

- 櫛状のぎざぎざ =  $F_0, 2F_0, 3F_0, 4F_0, \dots$  = 音の高さ (調波構造)
- 最初のぎざ：基本周波数, ピッチ (声帯振動の速さ)
- 山脈状の包絡 =  $F_1, F_2, \dots$  = 音の音色 (エネルギーの分布状況)
- ピーク位置：フォルマント周波数・共振周波数 (喉の形状)



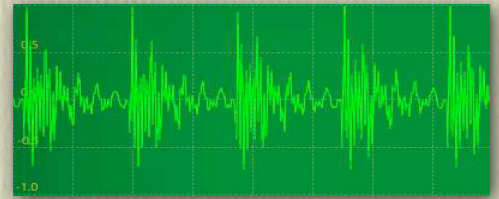
男声：広く  
女声：狭く

$1/T_0 = F_0$        $8/T_0 = 8F_0$       男声：狭く  
女声：広く

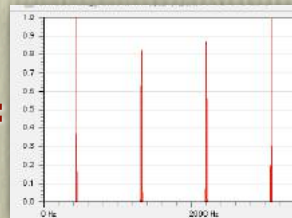


# 音声波形とスペクトル

ブザー（パルス列）→管→「あ〜」



=



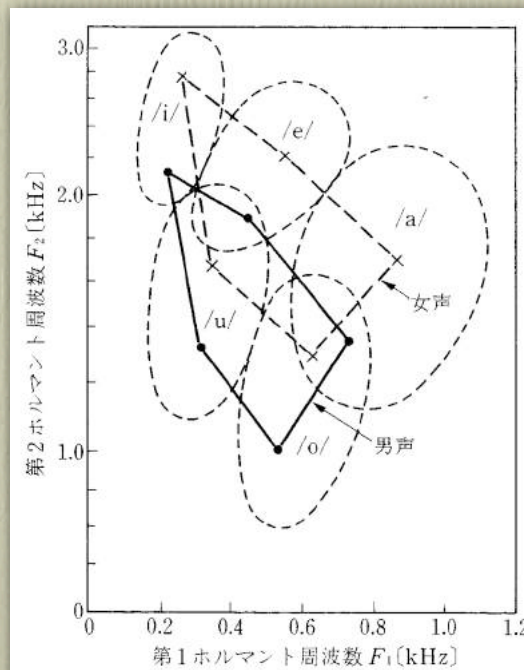
エネルギー配分に偏りが生じる  
この様子が管形状によって異なる  
エネルギーの局所的集中=共鳴

## 母音の違いと話者の違い

「母音の違い」も「話者の違い」もフォルマントの違い

● 日本語5母音の第一・第二フォルマント周波数

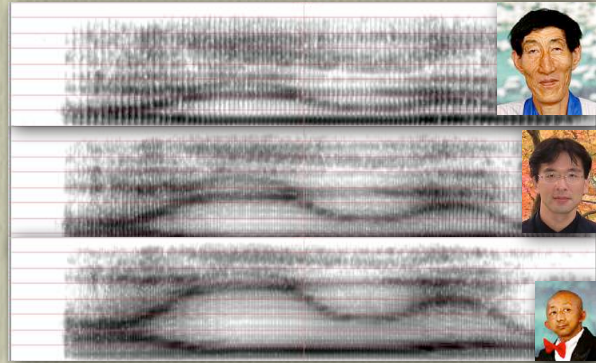
● 音色（エネルギーの分布状況）は喉の長さにも依存する





# 体格によって大きく変わる声

## 世界一の巨人と世界一の小人の会話



## 世界一小さな女子高生のインタビュー



# 音の四要素

## 様々な情報＝四要素の組み合わせ

### 音の高さ

- 高い音, 低い音

### 音の大きさ

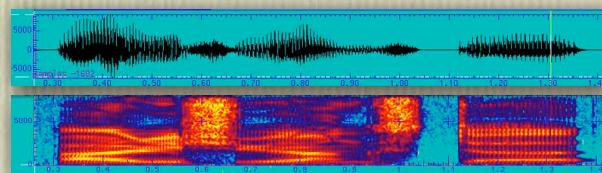
- 大きい音, 小さい音

### 音の長さ

- 長い音, 短い音

### 音の音色 (声色)

- 太い声, 細い声, 黄色い声, 甘い声, 渋い声, 色っぽい声...
- 高さ・大きさ・長さが同じ2音を「違う音」と認識した場合, その2音は音色が異なる。
- 「あ」と「い」      「あ」と「あ」



- 1) 高さ, 2) 大きさ
- 3) 長さ, 4) 音色

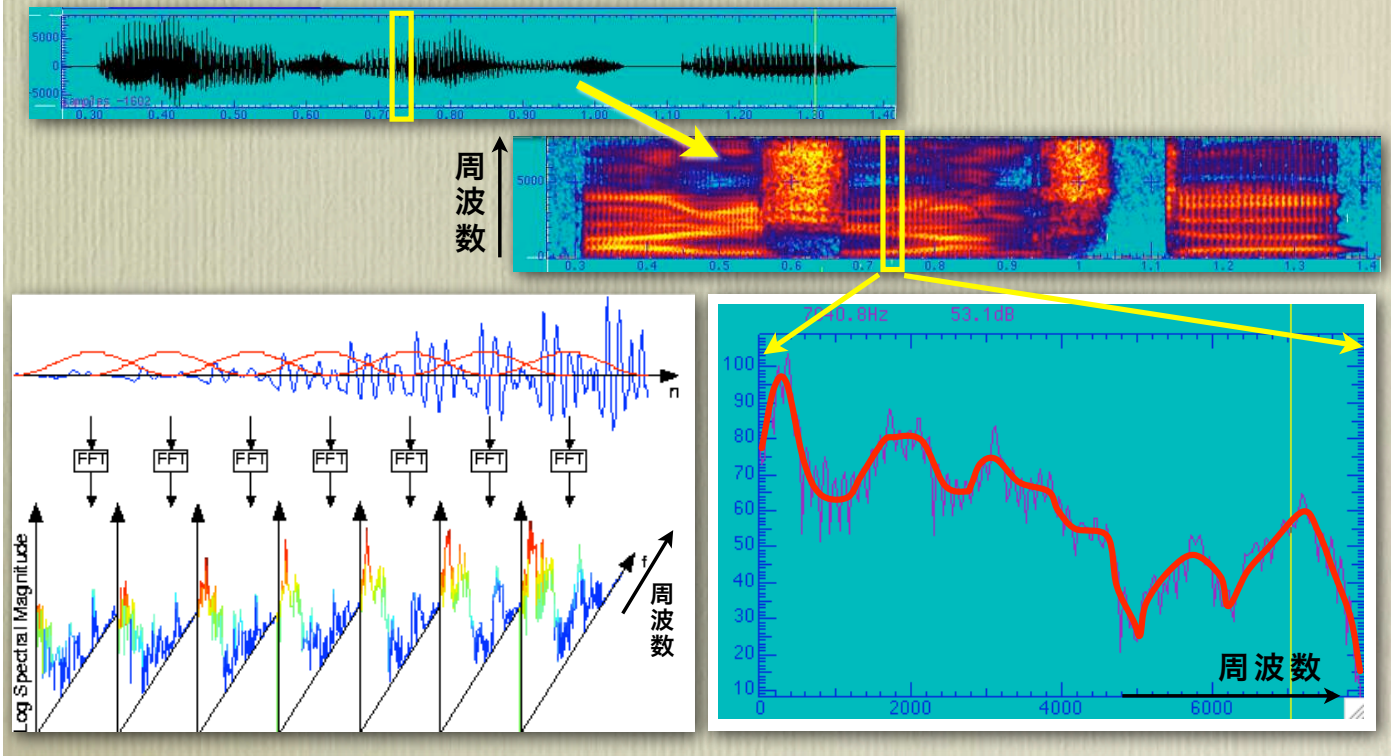
ある情報



# スペクトルからスペクトログラムへ

## 連続的に発声された音声の分析

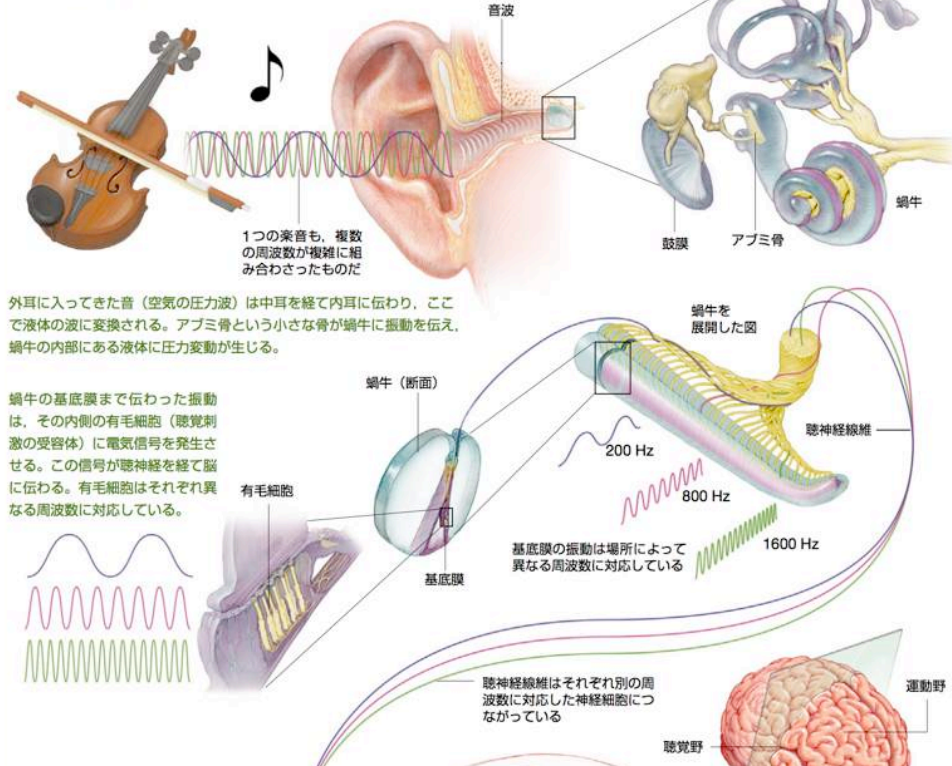
● 波形の区分化→波形の分解（スペクトル）→スペクトログラム



## 耳から脳へ ～音の通り道～

### 耳から脳へ 音楽の通り道

音楽を聴くと、脳では聴覚野だけでなく、さまざまな領域が反応する。通常なら音とは別の処理に関係している領域も加わり、視覚や触覚、感情といった体験すべてが脳の音楽処理に影響する。





## 質問 1

Q : 管を短くするとどんな (誰の) 声になるのか？



フォルマントが全体的に下がる



全体的に上がる

A → B → C

## 質問 2

Q : 管を「あ」から何の形に変えたのか当てなさい。

● この人の「あ」の管の形から変えていきます。

A =



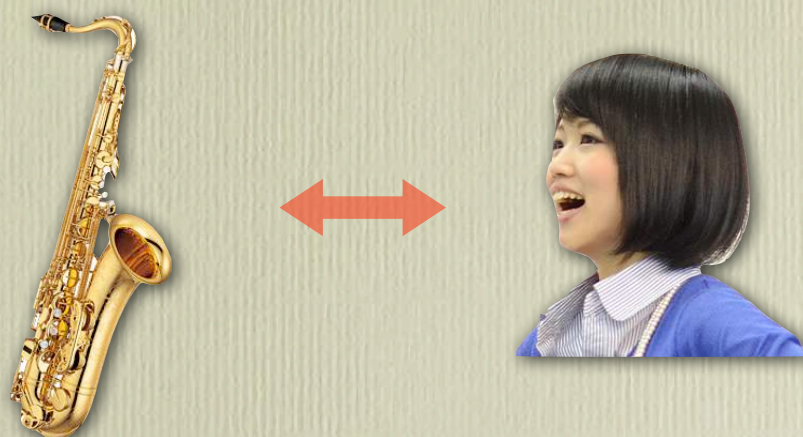
A → B B → A



## 質問 2

Q：管を「あ」から何の形に変えたのか当てなさい。

- この人の「あ」の管の形から変えていきます。



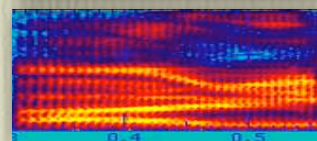
## 本日のメニュー

● 母音とその生成メカニズム

- 口の構え＝舌の位置，唇の丸め方

● 音響音声学の「いろは」

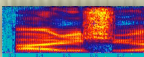
- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



● 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

● 鈍感さと敏感さのバランスが崩れるとどうなる？

-  -  = ?

- 刺激そのものに敏感に／刺激群の差異に鈍感になるとどうなる？
- 自閉症（＝関係性の病？）に関するいくつかの思考実験

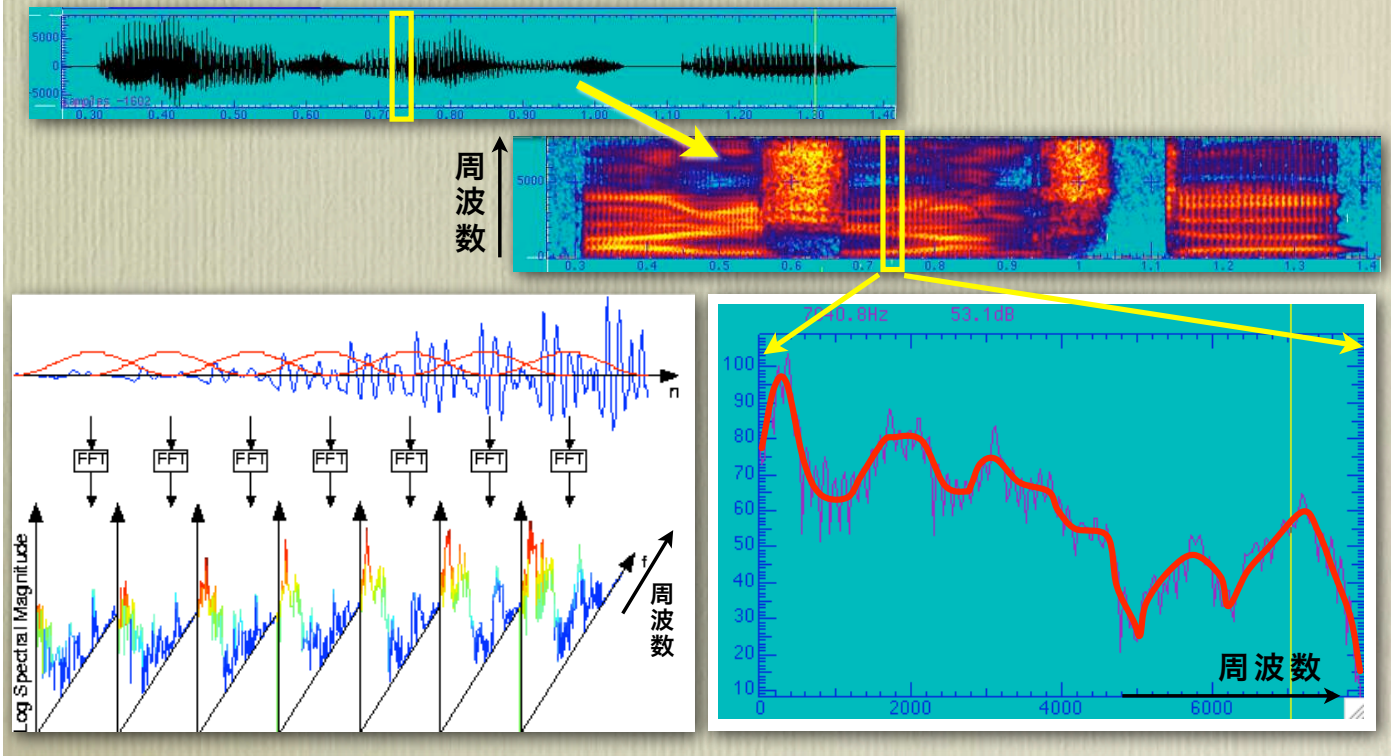
● まとめ



# スペクトルからスペクトログラムへ

## 連続的に発声された音声の分析

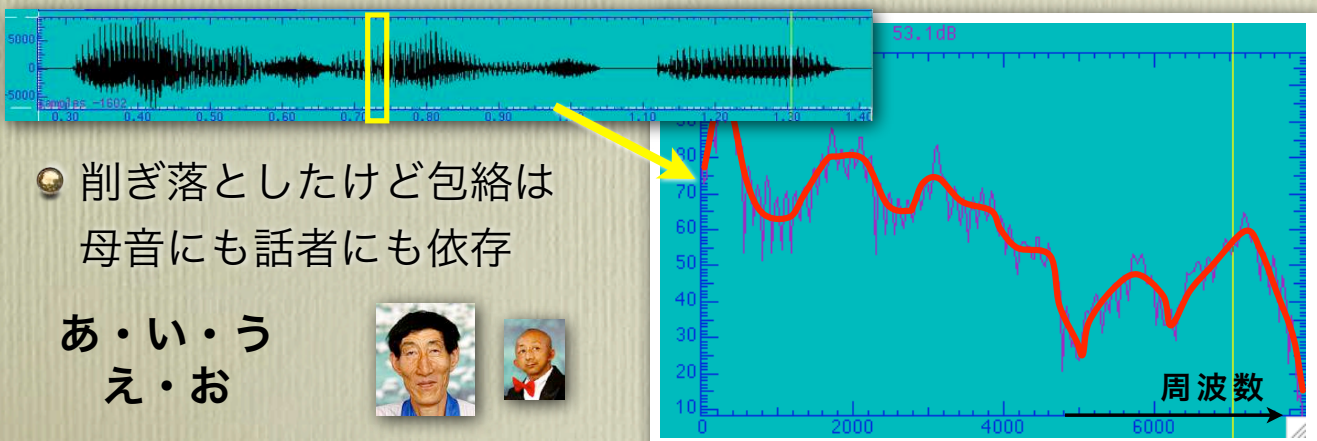
- 波形の区分化→波形の分解（スペクトル）→スペクトログラム



# 波形からスペクトル包絡（音色）へ

## 実は二段階の削ぎ落とし作業をしています

- 波形 → スペクトル（ギザギザ+包絡）
  - 人間の耳は波形の位相成分の違いに鈍感であることを利用
  - 波形から位相成分を削ぎ落とし
- スペクトル → スペクトル包絡（音色）
  - 人間の音韻（音素）知覚はピッチの違いに鈍感であることを利用
  - スペクトルからピッチ成分を削ぎ落とし



- 削ぎ落としたけど包絡は母音にも話者にも依存

あ・い・う  
え・お





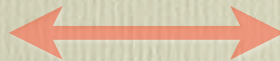
# 機械の音声処理と人間の音声処理

## 話者の違いに頑健な機械の音声認識

- 話者**バランス性**のとれた膨大な音声サンプルを集める
  - 人が一生で聞く音声サンプル数よりも多くの音声を集める [Moore'03]
- 新たな話者、環境に対して認識装置の調整が必要になることも
  - 想定した話者や環境以外の場所では、認識率は下落する

## 話者の違いに頑健な人間の音声認識

- 赤ちゃんが聞く声の多くは母親と父親（+ベビーシッター）
  - 話者**バランス性**の極めて偏った音声サンプルのみが提供される
- そんな環境でなぜ、頑健な音声認識能力が獲得できるのか？
  - 人間の持つ**汎化能力**、**抽象化能力**はどのように技術実装できるのか？
  - 峯松の声から「峯松であること」を消すことはできるのか？



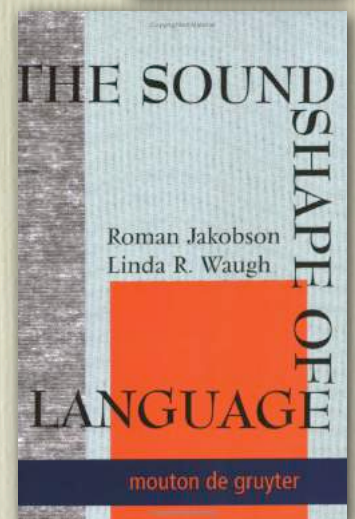
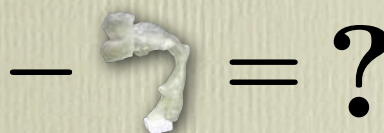
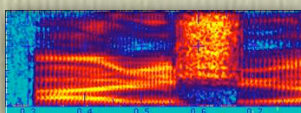
# 古典的音韻論より

## ヤコブソンの関係的不変性理論 [Jakobson+'79]

- 弁別素性理論
  - /x/ -- /y/ の関係は話者に依らず不変

We have to put aside the accidental properties of individual sounds and substitute a general expression that is the common denominator of these variables.

Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.

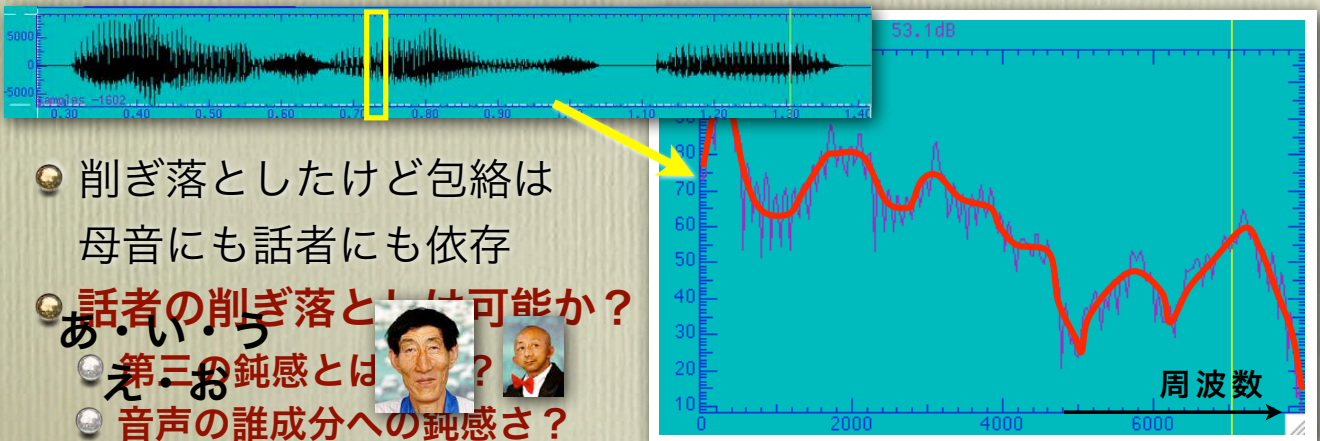




# 波形からスペクトル包絡（音色）へ

## 実は二段階の削ぎ落とし作業をしています

- 波形 → スペクトル（ギザギザ+包絡）
  - 人間の耳は波形の位相成分の違いに鈍感であることを利用
  - 波形から位相成分を削ぎ落とし
- スペクトル → スペクトル包絡（音色）
  - 人間の音韻（音素）知覚はピッチの違いに鈍感であることを利用
  - スペクトルからピッチ成分を削ぎ落とし



- 削ぎ落としたけど包絡は母音にも話者にも依存
- 話者の削ぎ落としは可能か？  
あ・い・う  
え・お  
第三の鈍感とは？  
音声の誰成分への鈍感さ？

## 音声模倣行為に見られる第三の鈍感

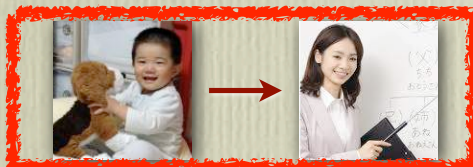
### 我々は相手の声（音）を真似ているのか？

- AがBの発声を真似る。例えば・・・
  - A：学生， B：語学教師
  - A：赤ちゃん， B：母親／父親
  - A：あなた， B：職業歌手（カラオケ）
- 我々は通常、音響的にずれた模倣をする
  - なぜ、話者の違いはスルーして音声模倣するのか？第三の鈍感？



### 動物に見られる音声模倣

- 動物学的には非常に稀な行為。霊長類では人間だけ[1]
- 他の動物では小鳥，クジラ，イルカくらいか[2]
- 基本的に、音響的な模倣，声帯模写的[2]





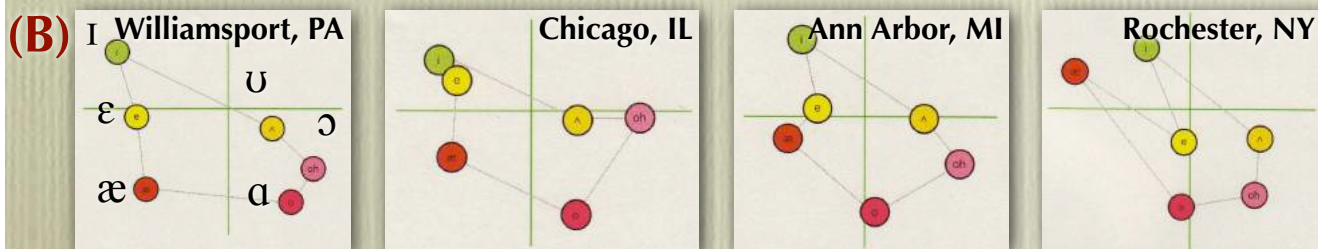
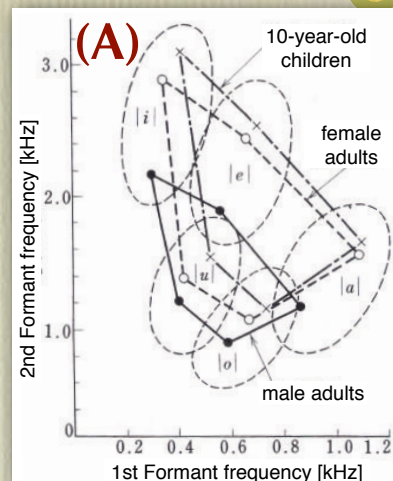
# 音声（母音）学習における鈍感・敏感

## 言語獲得における音的鈍感・敏感さ

- 年齢・性別に起因する差異 (A) には鈍感
- 方言に起因する差異 (B) には敏感

## 上記を少し言い換えてみると・・・

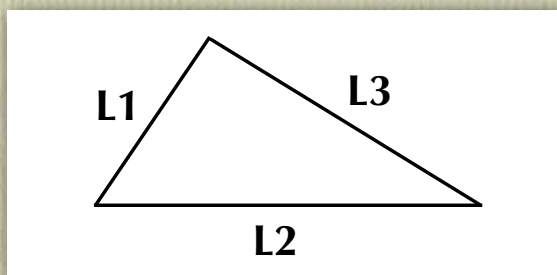
- 音声の表層的特性には鈍感で、音声の音的關係性（内部構造＝音的差異）には敏感
- (A) 表層的特性, (B) 音的關係性
- 音的關係性（母音の分布状況）は幾何学的には距離行列として表現できる。



Distribution of normalized formants among AE dialects [Labov et al.'05]

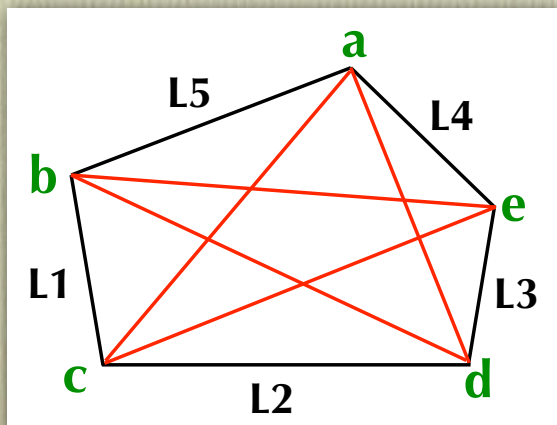
## 三角形と多角形の定義

### 三角形



(L1, L2, L3)

### 多角形



$$(L1, L2, L3, L4, L5) \begin{bmatrix} a & b & e \\ a & d_{11} & d_{12} & \dots & d_{1N} \\ b & d_{21} & d_{22} & \dots & d_{2N} \\ c & d_{31} & d_{32} & \dots & d_{3N} \\ d & : & : & \dots & : \\ e & d_{N1} & d_{N2} & \dots & d_{NN} \end{bmatrix}$$



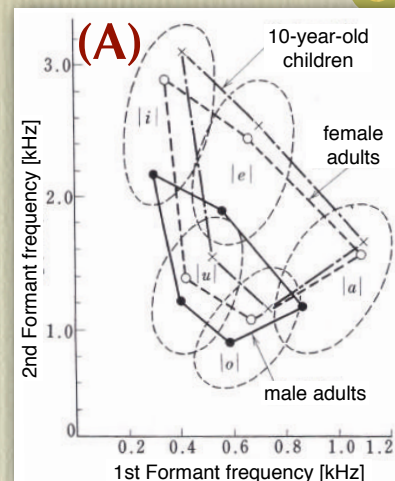
# 音声（母音）学習における鈍感・敏感

## 言語獲得における音的鈍感・敏感さ

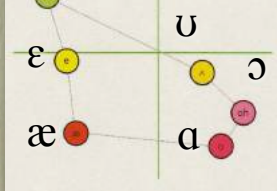
- 年齢・性別に起因する差異 (A) には鈍感
- 方言に起因する差異 (B) には敏感

## 上記を少し言い換えてみると・・・

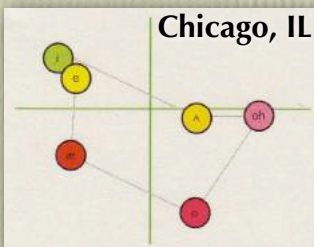
- 音声の表層的特性には鈍感で、音声の音的關係性（内部構造＝音的差異）には敏感
- (A) 表層的特性, (B) 音的關係性
- 音的關係性（母音の分布状況）は幾何学的には距離行列として表現できる。



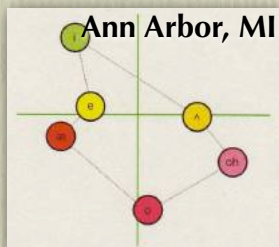
## (B) Williamsport, PA



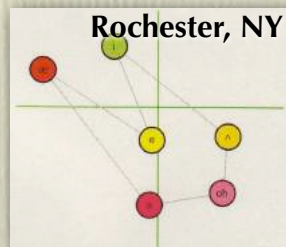
## Chicago, IL



## Ann Arbor, MI



## Rochester, NY



Distribution of normalized formants among AE dialects [Labov et al.'05]

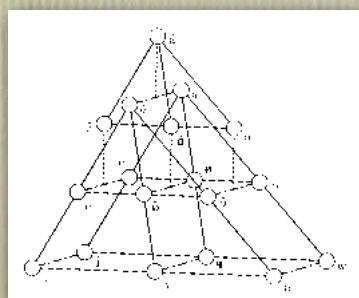
## 古典的音韻論より

## ヤコブソンの関係的不変性理論 [Jakobson+'79]

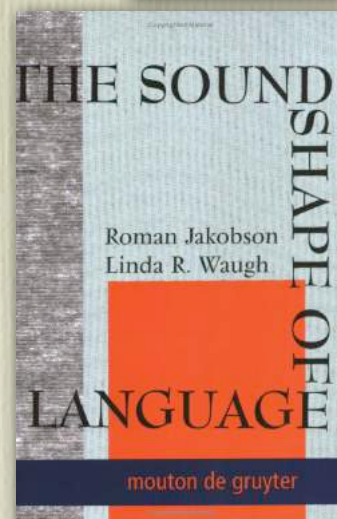
- 弁別素性理論
- /x/ -- /y/ の関係は話者に依らず不変

We have to put aside the accidental properties of individual sounds and substitute a general expression that is the common denominator of these variables.

Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.



$$\begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ d_{31} & & & \\ \vdots & & & \\ d_{N1} & d_{N2} & \dots & d_{NN} \end{bmatrix} = ?$$





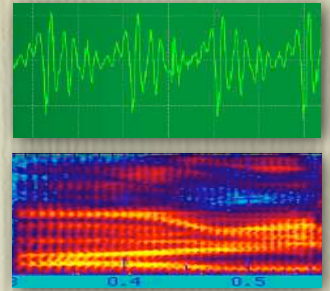
# 本日のメニュー

## 母音とその生成メカニズム

- 口の構え＝舌の位置，唇の丸め方

## 音響音声学の「いろは」

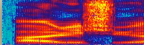
- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



## 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

## 鈍感さと敏感さのバランスが崩れるとどうなる？

-  -  = ?
- 刺激そのものに敏感に／刺激群の差異に鈍感になるとどうなる？
- 自閉症（＝関係性の病？）に関するいくつかの思考実験

## まとめ

# メロディー聴取における鈍感・敏感

## カラオケでキーを上げ下げして曲を聞く

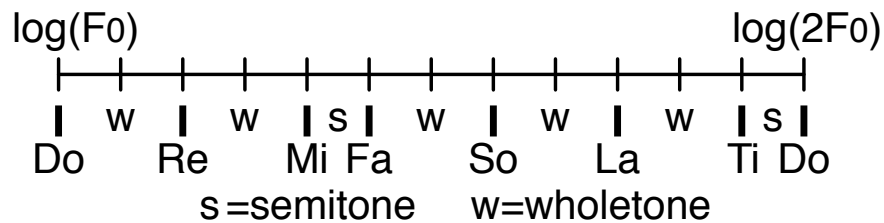


- 絶対音感者（ドレミは音名）
  - 1 = ソーミソドーラードドソー， 2 = レーシレソーミーソソレー
- 言語化可能な相対音感者（ドレミは階名）
  - 1 = ソーミソドーラードドソー， 2 = ソーミソドーラードドソー
    - 異なる音を同一と主張し，同一の音を異なると主張する。
    - 各音が他の音群に対して持つ音高差から階名は同定される。
    - 動物には相対音感がないため，移調すれば異なる音列となる[8]
- 言語化困難な相対音感者（ラーラ音感者）
  - 1 = ラーララーラーラーラー， 2 = ラーララーラーラーラー



# メロディー聴取における鈍感・敏感

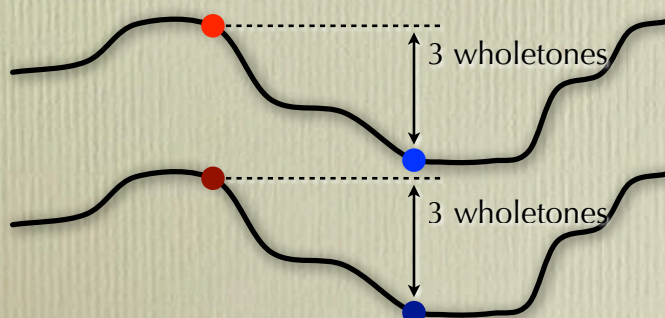
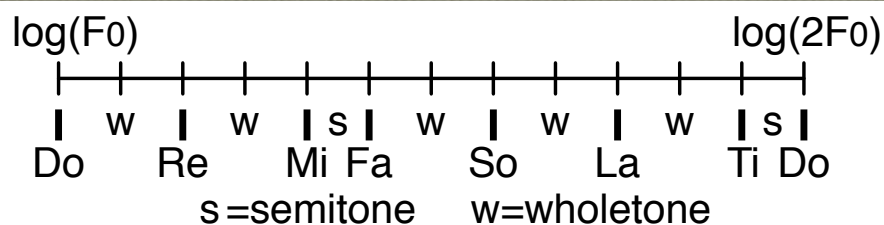
カラオケでキーを上げ下げして曲を聞く



- 階名：キーに依存しない音の同定・音高コントラストへの着眼
- しかし，孤立音の同定は不可能。コントラストがないから。

# メロディー聴取における鈍感・敏感

カラオケでキーを上げ下げして曲を聞く

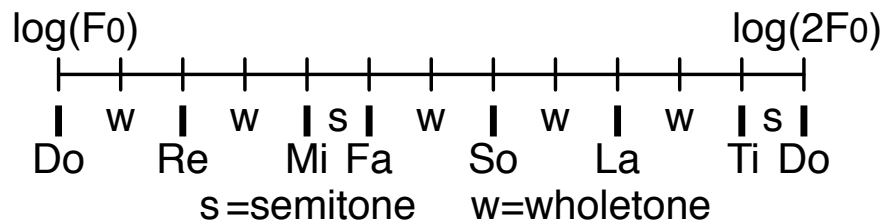


● ● and ● ● have to be Fa & Ti or Ti & Fa due to contrastive constraints.



# メロディー聴取における鈍感・敏感

カラオケでキーを上げ下げして曲を聞く



"I cannot identify."

# メロディー聴取における鈍感・敏感

カラオケでキーを上げ下げして曲を聞く



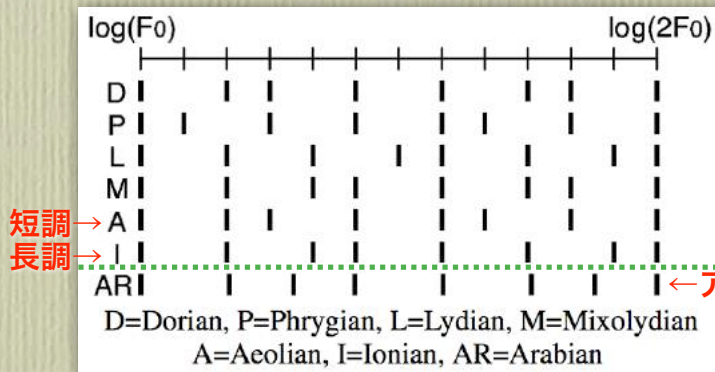
- 絶対音感者 (ドレミは音名)
  - 1 = ソーミソドーラードドソー, 2 = レーシレソーミーソソレー
- 言語化可能な相対音感者 (ドレミは階名)
  - 1 = ソーミソドーラードドソー, 2 = ソーミソドーラードドソー
    - 異なる音を同一と主張し, 同一の音を異なると主張する。
    - 各音が他の音群に対して持つ音高差から階名は同定される。
    - 動物には相対音感がないため, 移調すれば異なる音列となる。
- 言語化不可能な相対音感者
  - Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.





# メロディー聴取における鈍感・敏感

## 音楽における調不変の音配置とその変種



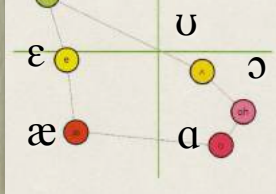
- 西洋音楽 = 5全音 + 2半音
- 種々の配置 = 教会音楽
- 民族音楽には半音以外の配置

أهلاً وسهلاً

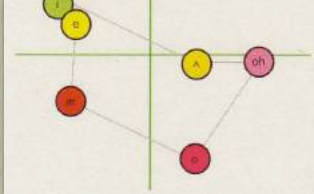


## 音声における話者不変の音配置とその変種 = 方言 [9]

(B) Williamsport, PA



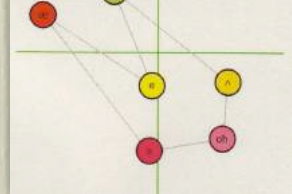
Chicago, IL



Ann Arbor, MI



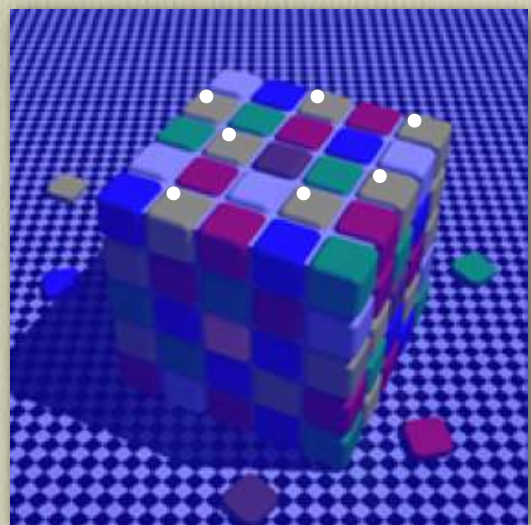
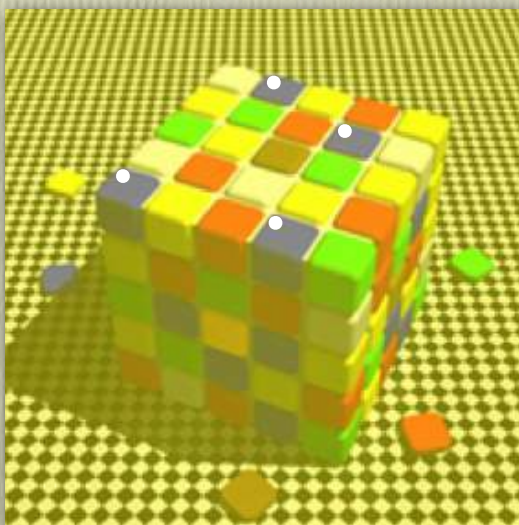
Rochester, NY



Distribution of normalized formants among AE dialects [Labov et al.'05]

# 色知覚における鈍感・敏感

## 黄・青眼鏡を通して眺めるルービックキューブ

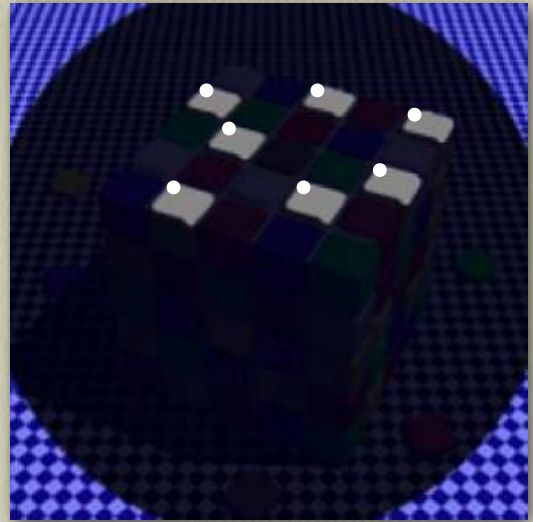
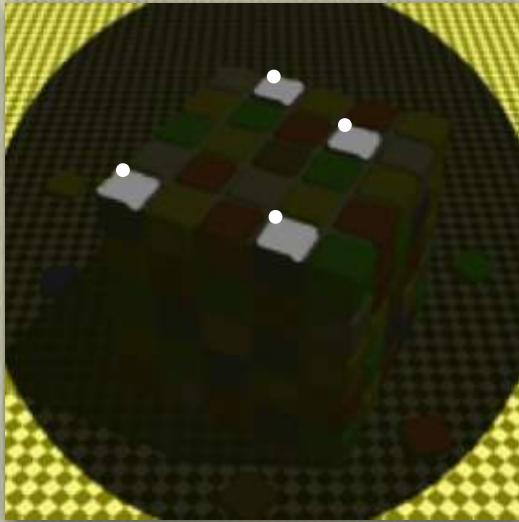


- 両者が同一のキューブであることは容易に認知可能
- 異なる色を同一と主張し，同一の色を異なると主張する。
- 各パッチが持つ波長（絶対量）だけではなく，各パッチが他のパッチ群とどのようなコントラストを持つのか，が非常に重要



## 色知覚における鈍感・敏感

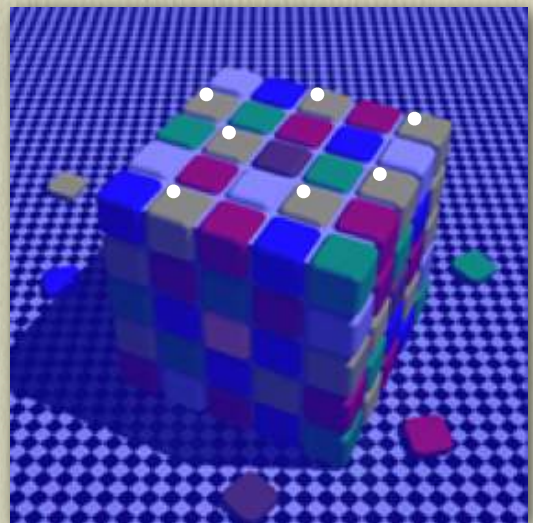
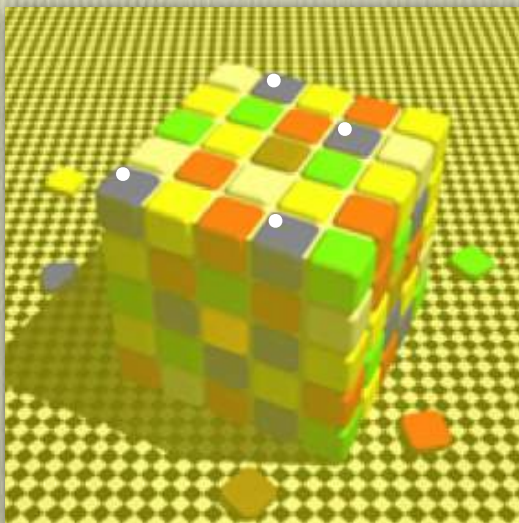
### 黄・青眼鏡を通して眺めるルービックキューブ



- 両者が同一のキューブであることは容易に認知可能
- 異なる色を同一と主張し，同一の色を異なると主張する。
- 各パッチが持つ波長（絶対量）だけではなく，**各パッチが他のパッチ群とどのようなコントラストを持つのか**，が非常に重要

## 色知覚における鈍感・敏感

### 黄・青眼鏡を通して眺めるルービックキューブ

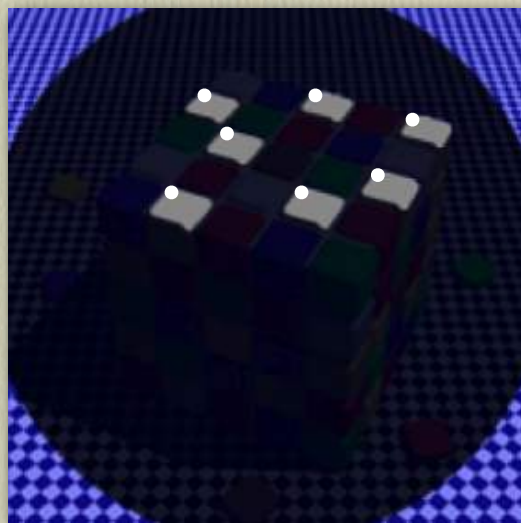
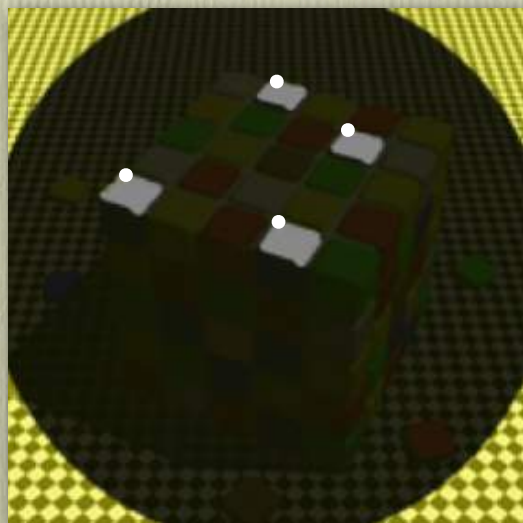


- 両者が同一のキューブであることは容易に認知可能
- 異なる色を同一と主張し，同一の色を異なると主張する。
- 各パッチが持つ波長（絶対量）だけではなく，**各パッチが他のパッチ群とどのようなコントラストを持つのか**，が非常に重要



## 色知覚における鈍感・敏感

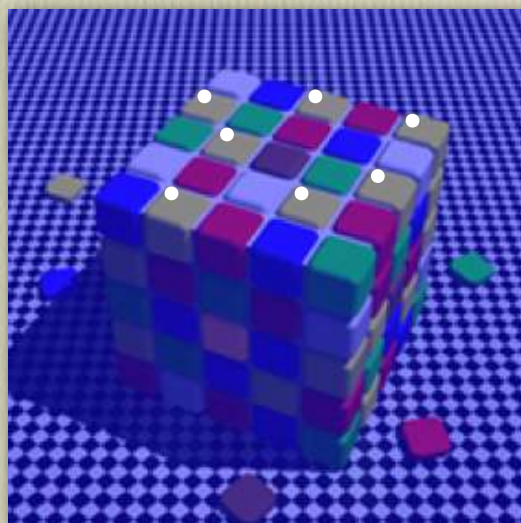
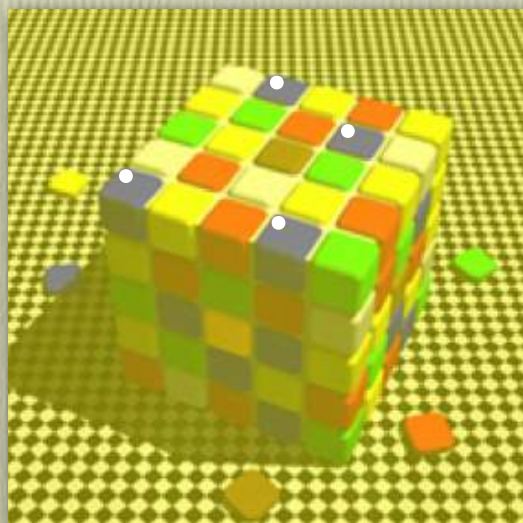
### 黄・青眼鏡を通して眺めるルービックキューブ



- 両者が同一のキューブであることは容易に認知可能
- 異なる色を同一と主張し，同一の色を異なると主張する。
- 各パッチが持つ波長（絶対量）だけではなく，**各パッチが他のパッチ群とどのようなコントラストを持つのか**，が非常に重要

## 色知覚における鈍感・敏感

### 黄・青眼鏡を通して眺めるルービックキューブ



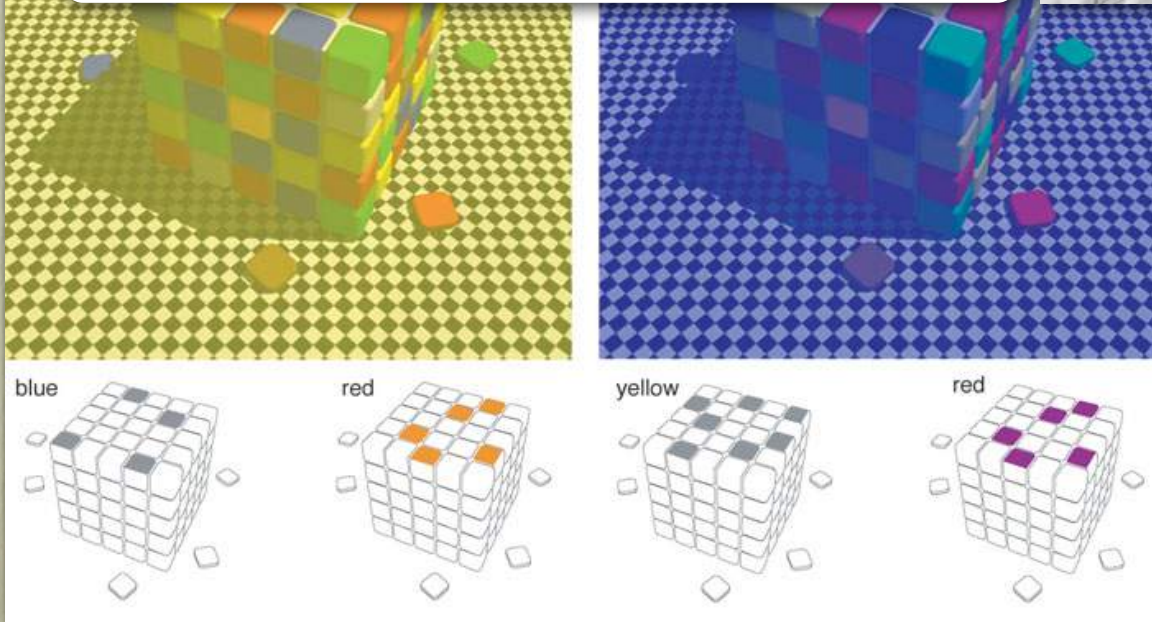
- 両者が同一のキューブであることは容易に認知可能
- 異なる色を同一と主張し，同一の色を異なると主張する。
- 各パッチが持つ波長（絶対量）だけではなく，**各パッチが他のパッチ群とどのようなコントラストを持つのか**，が非常に重要



# 色知覚における鈍感・敏感

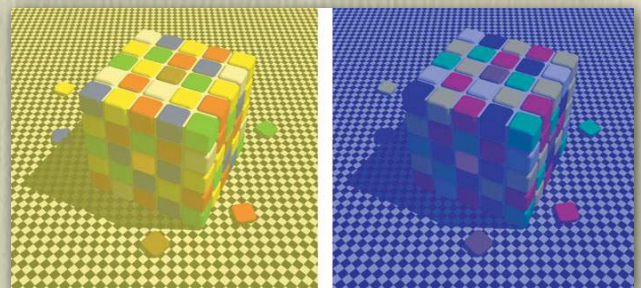
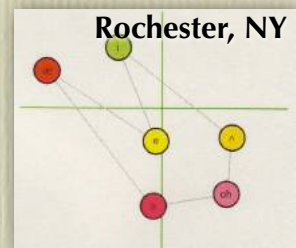
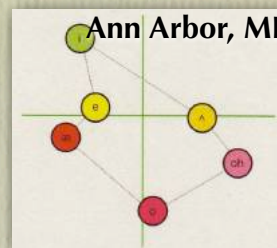
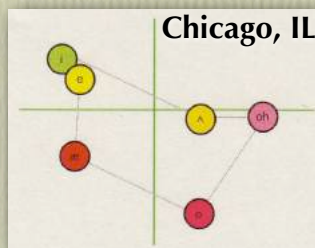
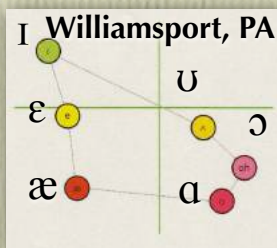


Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.



Reprinted from Dale Purves, R. Beau Lotto, Surajit Nundy, "Why We See What We Do," American Scientist, vol. 90, no. 3, page 236. [www.americanscientist.org/template/AssetDetail/assetid/14755](http://www.americanscientist.org/template/AssetDetail/assetid/14755).

# 様々なメディアに対する鈍感・敏感



要素（刺激）と要素（刺激）の間の差異・距離  
に対する敏感さが我々の知覚を支えている。

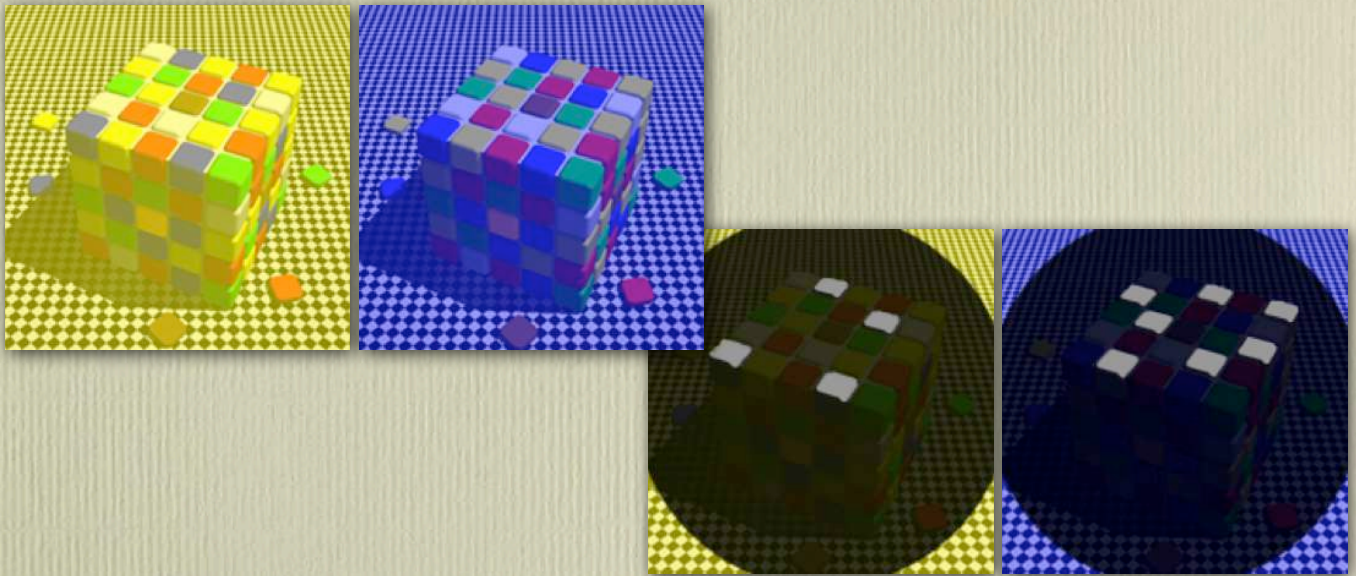
Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.





# 我々だけなのか，それとも・・・

色の恒常的知覚はどこまで遡れるのか？ [Briscoe'01]

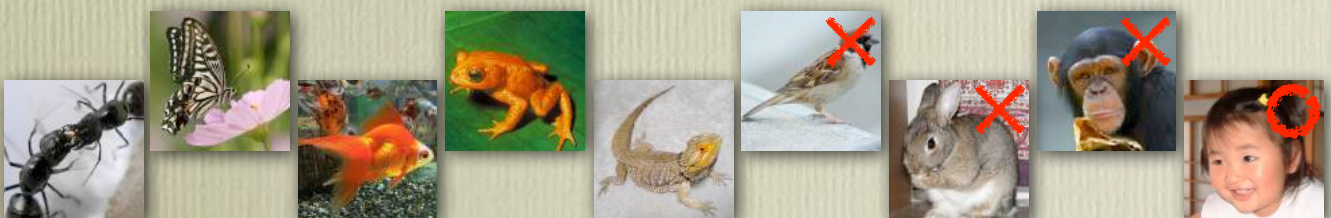


# 我々だけなのか，それとも・・・

ピッチの恒常的知覚はどこまで遡れるのか？ [Hauser'03]



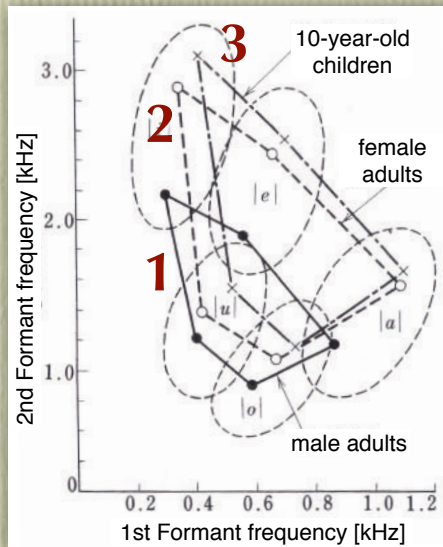
$$1 = 2$$





# 我々だけなのか，それとも・・・

母音の恒常的知覚はどこまで遡れるのか？



1=2=3



## 音声模倣行為に見られる第三の鈍感

我々は相手の声（音）を真似ているのか？

● AがBの発声を真似る。例えば・・・

- A：学生， B：語学教師
- A：赤ちゃん， B：母親／父親
- A：あなた， B：職業歌手（カラオケ）

● 我々は通常，音響的にずれた模倣をする

● なぜ，話者の違いはスルーして音声模倣するのか？第三の鈍感？



動物に見られる音声模倣

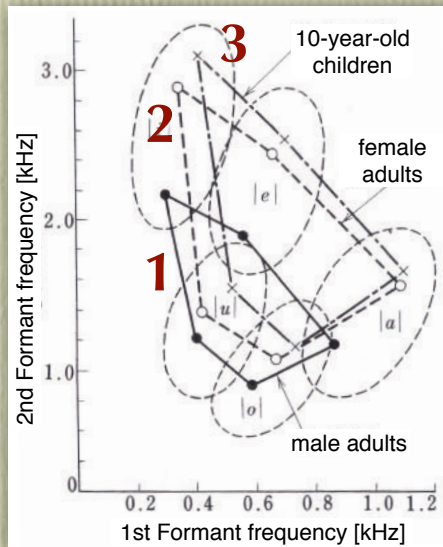
- 動物学的には非常に稀な行為。霊長類では人間だけ[1]
- 他の動物では小鳥，クジラ，イルカくらいか[2]
- 基本的に，音響的な模倣，声帯模写的[2]





# 我々だけなのか，それとも・・・

母音の恒常的知覚はどこまで遡れるのか？



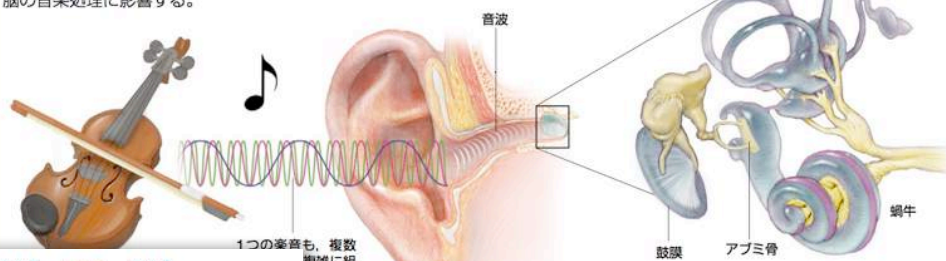
1=2=3



## 耳から脳へ ～音の通り道～

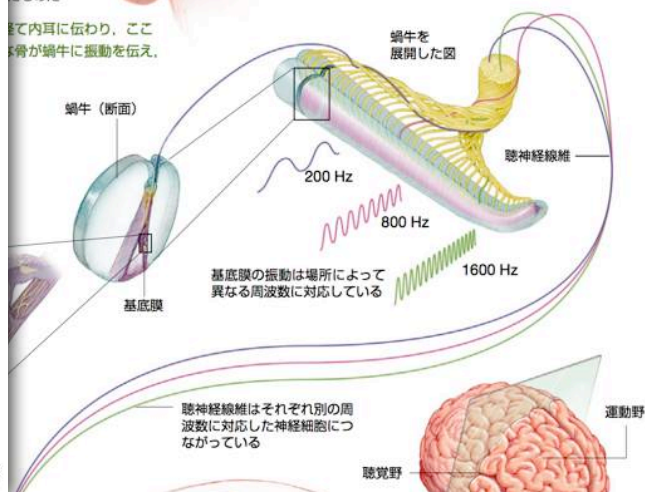
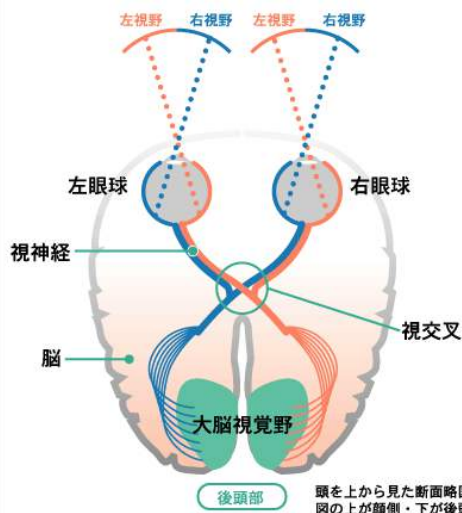
### 耳から脳へ 音楽の通り道

音楽を聴くと、脳では聴覚野だけでなく、さまざまな領域が反応する。通常なら音とは別の処理に関係している領域も加わり、視覚や触覚、感情といった体験すべてが脳の音楽処理に影響する。



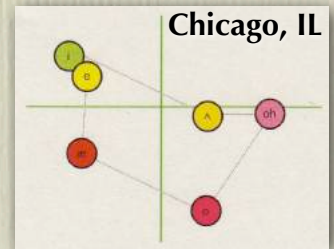
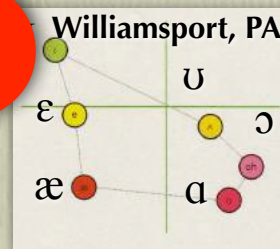
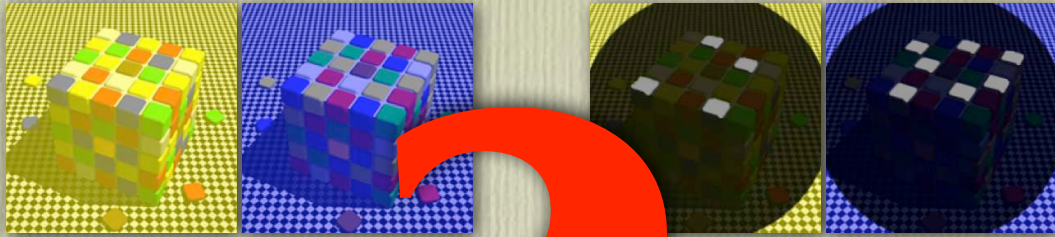
1つの楽音も、複数複雑に相たものだ

鼓膜が内耳に伝わり、ここが蝸牛に振動を伝える。





# 知覚の恒常性と進化的創発



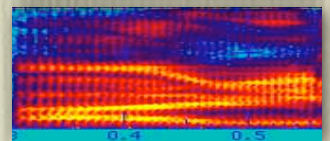
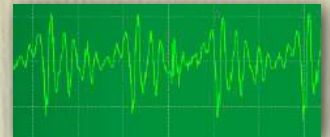
## 本日のメニュー

### 母音とその生成メカニズム

- 口の構え = 舌の位置, 唇の丸め方

### 音響音声学の「いろは」

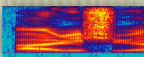
- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



### 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

### 鈍感さと敏感さのバランスが崩れるとどうなる？

●  -  = ?

- 刺激そのものに敏感に / 刺激群の差異に鈍感になるとどうなる？
- 自閉症 (= 関係性の病?) に関するいくつかの思考実験

### まとめ



# 音声模倣行為に見られる**第三の鈍感**

## 我々は相手の声（音）を真似ているのか？

- AがBの発声を真似る。例えば・・・

- A：学生， B：語学教師
- A：赤ちゃん， B：母親／父親
- A：あなた， B：職業歌手（カラオケ）

- **我々は通常，音響的にずれた模倣をする**

- なぜ，話者の違いはスルーして音声模倣するのか？第三の鈍感？



## 動物に見られる音声模倣

- 動物学的には非常に稀な行為。霊長類では人間だけ[1]
- 他の動物では小鳥，クジラ，イルカくらいか[2]
- 基本的に，音響的な模倣，声帯模写的[2]



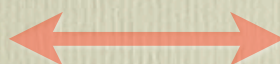
# 機械の音声処理と人間の音声処理

## 話者の違いに頑健な機械の音声認識

- **話者バランス性のとれた**膨大な音声サンプルを集める
  - 人が一生で聞く音声サンプル数よりも多くの音声を集める[Moore'03]
- 新たな話者，環境に対して認識装置の調整が必要になることも
  - 想定してた話者や環境以外の場所では，認識率は下落する

## 話者の違いに頑健な人間の音声認識

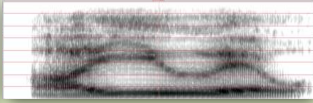
- 赤ちゃんが聞く声の多くは母親と父親（＋ベビーシッター）
  - **話者バランス性の極めて偏った**音声サンプルのみが提供される
- そんな環境でなぜ，頑健な音声認識能力が獲得できるのか？
  - 人間の持つ汎化能力，抽象化能力はどのように技術実装できるのか？
  - 峯松の声から「峯松であること」を消すことはできるのか？



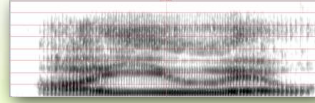


# 人の音声模倣は声の何を真似ているのか？

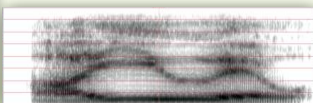
## 男子学生が女性英語教師の音声当真似る



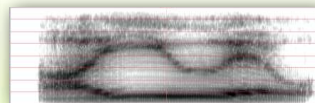
← /h e? ow/ →



## 幼児が母親の音声当真似る



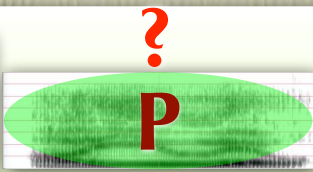
← /h e? ow/ →



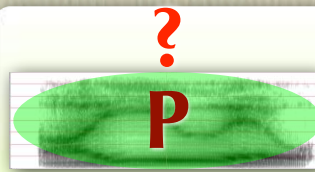
- 単語を音素の列として把握することは当然困難
- 日本語で言えば、「しりとり」なんてできない。

# 子供は声の何を真似ているのか？

親の発声 → 音韻同定 → 音韻列 → 個々の音韻を発声



✗  
← /おはよう/ →



- 音韻意識が希薄／しり取りも出来ない[4]。

## 発達心理学からの回答

- 幼児は語全体の語形・音形・ゲシュタルト・related spectral patternを獲得し、その後、個々の分節音を獲得する[4,5,6,7]
- 語ゲシュタルトには話者の情報は含まれない。二者間の共通項でないと、声帯模写をすることになる。
- 語ゲシュタルトの物理的・音響的定義は何か？
- 峯松が問いかけた心理学者で明確な定義を述べた方はまだいない。



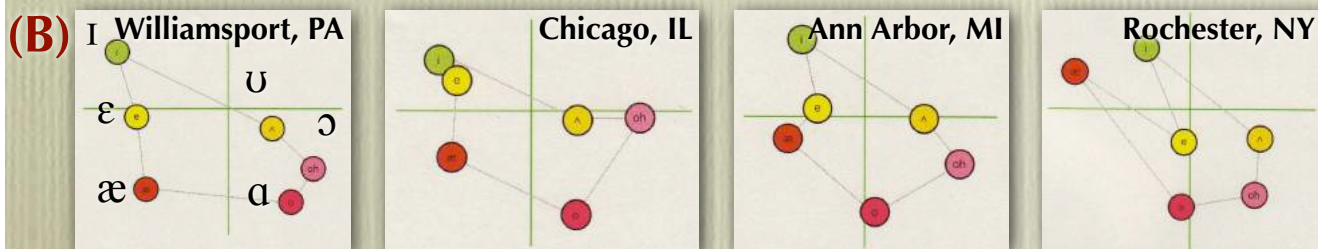
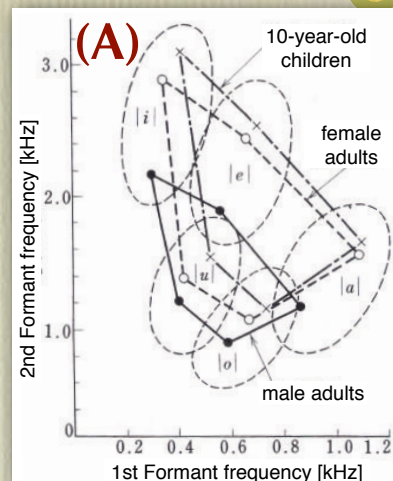
# 音声（母音）学習における鈍感・敏感

## 言語獲得における音的鈍感・敏感さ

- 年齢・性別に起因する差異 (A) には鈍感
- 方言に起因する差異 (B) には敏感

## 上記を少し言い換えてみると・・・

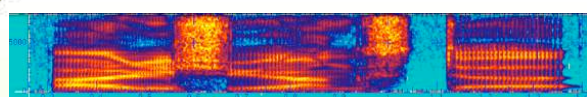
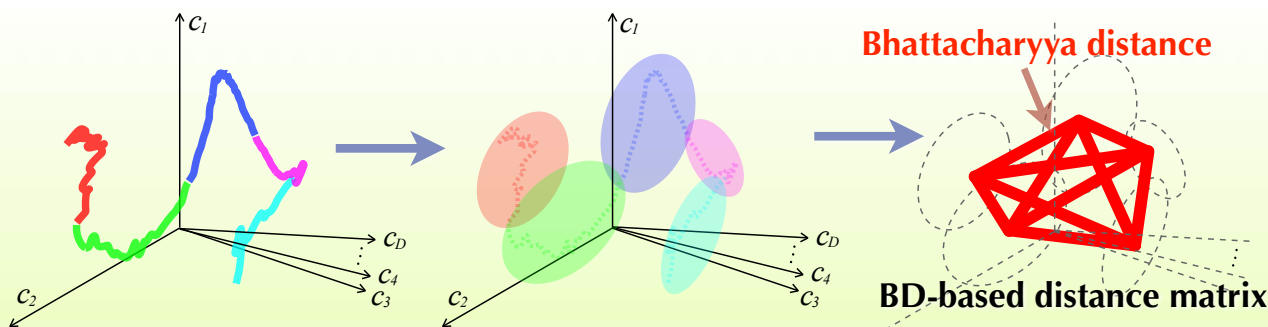
- 音声の表層的特性には鈍感で、音声の音的關係性（内部構造）には敏感
- (A) 表層的特性, (B) 音的關係性
- 音的關係性（母音の分布状況）は幾何学的には距離行列として表現できる。



Distribution of normalized formants among AE dialects [Labov et al.'05]

## ある発話を「形」として捉える

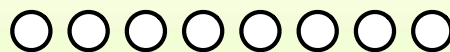
## Bhattacharyya distance で距離計測



spectrogram (spectrum slice sequence)



cepstrum vector sequence



distribution sequence

音色の相对音感 = 音声の構造的表象



# 古典的音韻論より

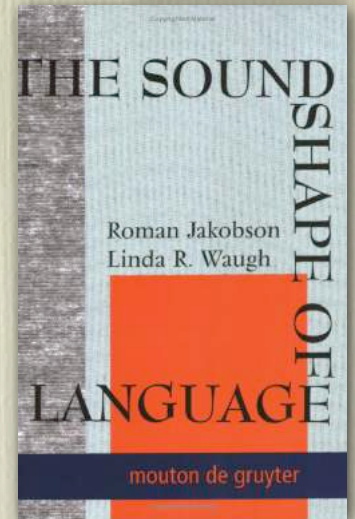
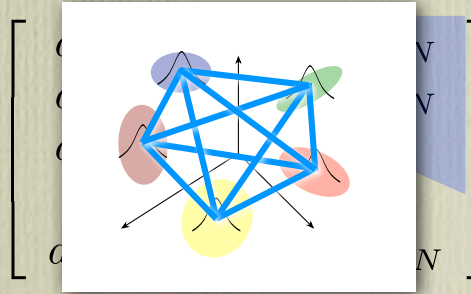
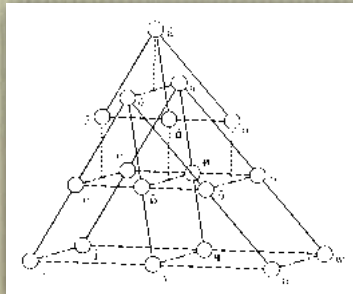
## ヤコブソンの関係的不変性理論 [Jakobson+'79]

### 弁別素性理論

- /x/ -- /y/ の関係は話者に依らず不変

We have to put aside the accidental properties of individual sounds and substitute a general expression that is the common denominator of these variables.

Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.



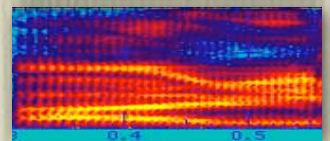
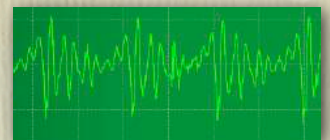
## 本日のメニュー

### 母音とその生成メカニズム

- 口の構え = 舌の位置, 唇の丸め方

### 音響音声学の「いろは」

- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



### 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

### 鈍感さと敏感さのバランスが崩れるとどうなる？



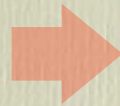
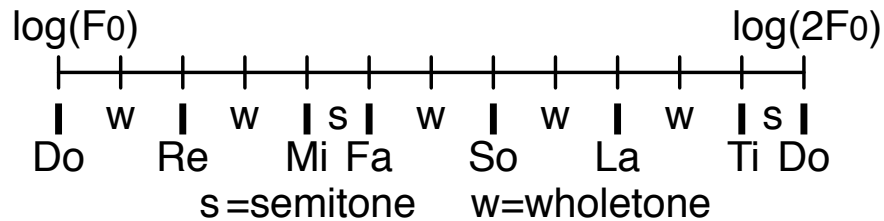
- 刺激そのものに**敏感**に / 刺激群の差異に**鈍感**になるとどうなる？
- 自閉症 (= 関係性の病?) に関するいくつかの思考実験

### まとめ



# メロディー聴取における鈍感・敏感

カラオケでキーを上げ下げして曲を聞く



"I cannot identify."

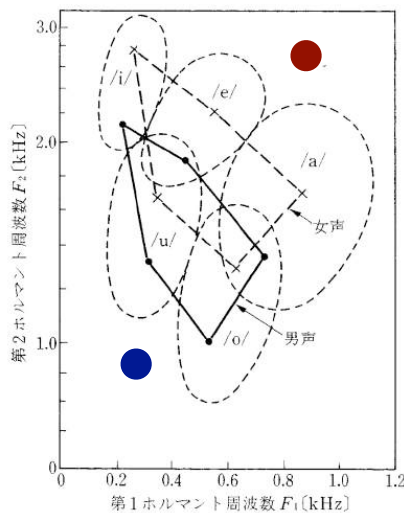
## 音色の相

言語化できる相対

- 孤立的に提示された
- 孤立的に提示された

巨人&小人の音声

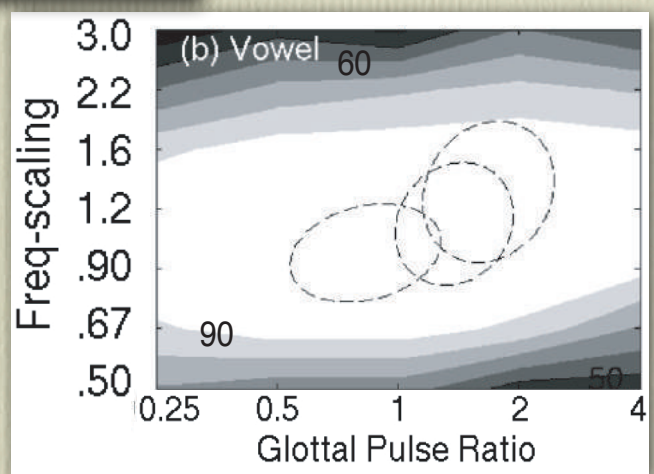
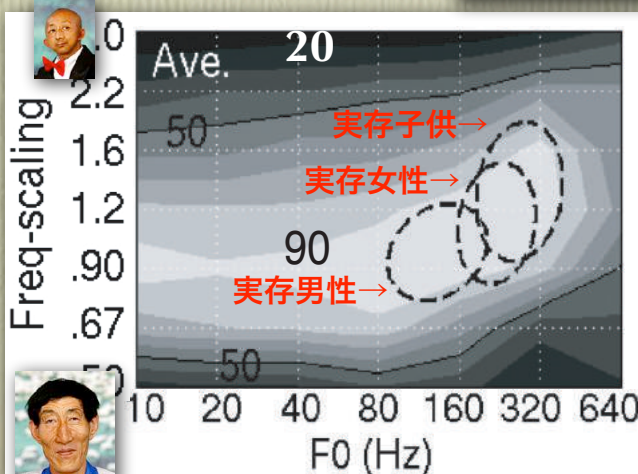
- 孤立母音の同定は困
- でも、無意味語で



## する問い



ると書き起こせる[12]





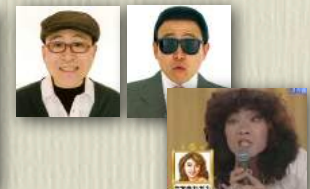
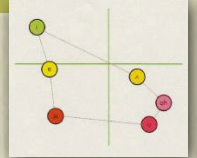
# 音色の絶対音感に対する問い

## 音声の相対的特徴（音間距離）に鈍感になると

- おそらく方言の獲得に苦勞するはず

## 音声の絶対的特徴（音そのもの）に敏感になると

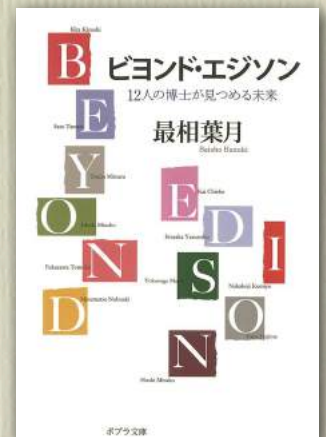
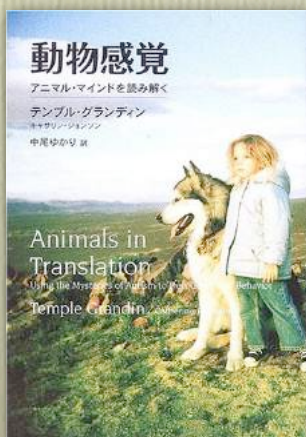
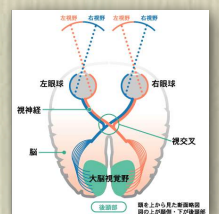
- 過去に聞いた声と「同一と知覚できる」声との遭遇が難しくなる
  - 呼びかけられても、話者、部屋、呼びかけ方が異なると違う声
  - 特定の人声は意味になる（理解可能）が、他の人だと難しくなる
- 音響的に同一の声（録音&再生声）を好む・反応する
  - アニメの決め台詞、CM音声、ニュースキャスターの定型的文言
  - 録音&再生技術が一般化される前は、彼らは何を真似たのか？
- 音声模倣が音響的な模倣（声帯模写）になる
  - 決め台詞、カラオケ、外国語学習など、提示された声（音）を模倣



# 動物的感觉と自閉症

## 動物の情報処理と自閉症者の情報処理

- Dr. Temple Grandin（アスペルガー＆動物学者）
- 「動物感觉」
  - 動物と自閉症者の情報処理的類似性を主張
  - 局所的／具体的／実体的 ↔ 全体的／抽象的／概念的
- 第一次視覚野からの結線が異常に太い・多い脳を持つ彼女

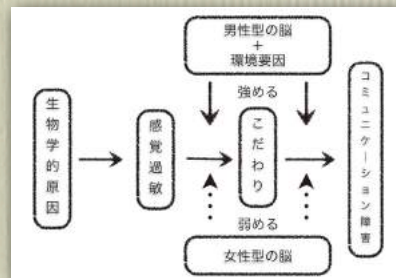




# 感覚の問題から捉える自閉症

## 社会性／心の問題より感覚問題へ焦点を当ててみる

- 当事者からの主張（感覚・知覚が健常者とは違う！）
  - T. Grandin 「自閉症の脳を読み解く」
- 抽象化・汎化処理及び規則抽出処理の不備から捉える自閉症
  - 藤居学「自閉症 ～からだとせかいをつなぐ新しい理解と療育～」
- 意識化できる範囲が健常者に比べて限定される
  - 神田進「脳内ディスプレイ」
- 感覚過敏・愚鈍から捉える自閉症
  - 熊谷高幸「自閉症と感覚過敏」
  - 感覚過敏・愚鈍＋男性脳を2大要因として描写

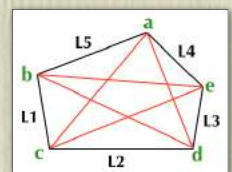
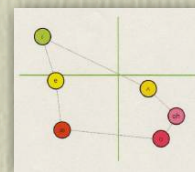
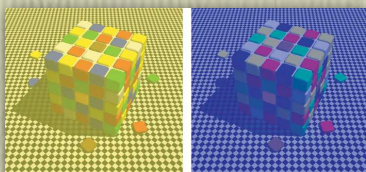


## XはAにもBにもCにもなる

### 物理的

Physiologically identical sounds may possess different values in conformity with the whole sound system, i.e. in their relations to the other sounds.

- 他者との差異・距離関係に基づいて、Xの機能は決まる。



### 社会的環境によって解釈が異なるX

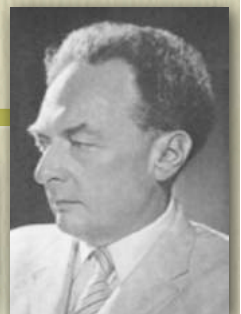
- 父@家, 夫@家, 次男@実家, 教授@大学, 理事@学会・・・

### 言語的環境（文脈）によって解釈が異なるX

- 頭をたたくな！ 頭を使いなさい（比喩表現）

### これらは重ねて議論すべきものなのか？

- 自閉症＝関係性の病？





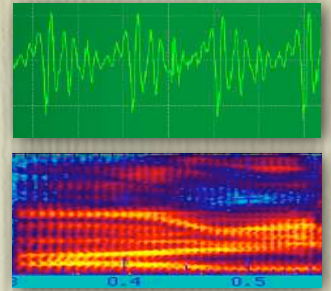
# 本日のメニュー

## 母音とその生成メカニズム

- 口の構え = 舌の位置, 唇の丸め方

## 音響音声学の「いろは」

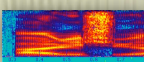
- 波形とスペクトルにみる「高さ」と「音色」
- 「あいうえお」も「大人・子供」も音色の違い



## 音声・メロディー・色を知覚するメカニズム

- 異なるものを同じと感じるメカニズム
- 刺激そのものに対する鈍感さと刺激群の差異に対する敏感さ

## 鈍感さと敏感さのバランスが崩れるとどうなる？

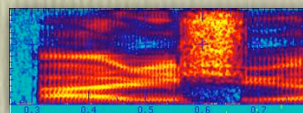
●  -  = ?

- 刺激そのものに**敏感**に / 刺激群の差異に**鈍感**になるとどうなる？

- 自閉症 (= 関係性の病?) に関するいくつかの思考実験

## まとめ

# 最初の一步, 現状, そして今後

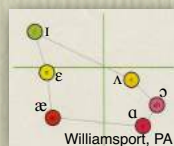


-  = ?

鈍感



音声学  
物理学  
心理学  
数学



敏感／鈍感  
相対感覚  
抽象化／汎化





- [1] W. Gruhn, Proc. Int. Conf. on language and music as cognitive systems (2006)
- [2] 岡ノ谷, 音響学会春音講論, 1-7-15, 1555-1556 (2008, 但し, 質疑応答を含む)
- [3] 宮本「声を作る・声を見る ～九宮鳥からヒトへ～」森北出版 (1995)
- [4] 原, コミュニケーション障害学, 20, 2, 98-102 (2003)
- [5] 加藤, コミュニケーション障害学, 20, 2, 84-85 (2003)
- [6] 早川, 月刊言語, 35, 9, 62-67 (2006)
- [7] P. Lieberman, Child Phonology, 1, 7, Academic Press, 1981.
- [8] M. D. Hauser and J. McDermott, Nature neurosciences, 6, 663-668 (2003)
- [9] W. Labov, Atlas of North American English, Mouton and Gruyter (2005)
- [10] R. Jakobson and L. R. Waugh, The sound shape of language, Mouton De Gruyter (1987)
- [11] 青木他, 音響学会秋音講論, 2-P-6, 373-374 (2004)
- [12] 林他, 音響学会秋音講論, 2-Q-27, 473-474 (2007)
- [13] Briscoe, A. D. and Chittka, L., Annual review of entomology, 46, pp.471-510, 2001.
- 峯松信明「あ」という声を聞いて母音「あ」と同定する能力は音言語運用に必要なか?, 日本語学4月号, 187-197, 明治書院 (2008)
- 東田直樹「自閉症の僕が飛び跳ねる理由」エスコアール (2007)
- T. Grandin「動物感覚」日本放送出版協会 (2006)
- 最相葉月「ビヨンド・エジソン」ポプラ社 (2012)
- 峯松他「音声に含まれる言語的情報を非言語情報から音響的に分離して抽出する手法の提案～人間らしい音声情報処理の実現に向けた一検討～」電子情報通信学会論文誌, J94-D, 1, 12-16 (2011, 招待論文)
- 峯松他「音声の物理特性を通して考える失読症の音声認知」発達性ディスレクシア研究予稿集 (2006)
- 峯松他「音声の物理特性を通して考える失読症・自閉症の音声認知」認知神経心理学研究会予稿集 (2006)
- N. Minematsu, et al., "Development of a spoken word recognizer without phonemic awareness -- Is this machine a Dyslexia simulator? --," Proc. The 2nd Riken Brain Science Institute and Oxford-Kobe Joint International Symposium, 49-51 (2007)



ご清聴, ありがとうございました





# 声はことばなのか音なのか？

## 自閉関連の書籍より

- 言葉（音声言語）が出てこない。話ことばの無い世界。
- 呼びかけても答えない，声に対する反応が鈍い， などなど・・・

## 横田敬一郎先生@日本自閉症協会顧問より

- 自閉っ子を指導を（あれこれ考えながら）50年続けた結論
  - 彼らの音の処理， 声の処理をもっと調査すべし！
  - 乳幼児を取り巻く音環境が「うるさすぎる！」
  - 音環境を整えれば言葉が出てきたはずの子供たち！
- 自閉っ子の聞こえ・ことばの検査
  - 聴覚テスト（純音をヘッドホン提示して行う聞き取り検査）
  - 知能テスト（話し言葉の理解）
- 聴覚テストと知能テストの間にあるべきテストがない
  - 音でしかない声を「特殊な音， ことば」として聞けているかどうか
  - 音がシンボルに／対象物を指す音（言葉）になっているかどうか



# 言語音の聴取と非言語音の聴取

## 音声科学

- 人間の音の処理には二つのモードがある。
  - 音声モード／非音声モード。音声は非言語音とは異なる処理。

## 脳科学

- 音声＝左脳， 非音声＝右脳
  - ただしメロディーも絶対音感者は左脳で聞く。
  - 音ストリームをシンボル列として捉えているかどうか？

## 声を音として聞くのか， 言語として聞くのか？

- fMRI, PET, EEG, NIRS などの脳活動イメージングで可能
  - 声を聞いた被験者の反応を観察することで， もっと簡便な方法で。
  - N歳児検診で使えるような簡便な方法で。
- （Nの制約を無視して）模倣行為を考えてみる
  - 与えられた音を， 音として真似る／言葉として真似る





# 音真似・声真似・言真似



## 動物の声には、人の声に聞こえるものがある

- <https://www.youtube.com/watch?v=prrrMaLrkc5U>

## 声模倣被験者実験をやってみる

### ● ある音源を**動物の声**として真似させる



- ジブリの次回作は動物の鳴き声を全部人の声で作るそうだ
- 「風立ちぬ」で飛行機のエンジン音は全て人の声
  - <https://www.cinematoday.jp/news/N0055623>
  - <https://www.youtube.com/watch?v=l42ZHkKALd4>
- 色々な動物の声の真似上手を探しているそうだ
  - 犬, 猫, カラス, きつね, たぬき, きりん, **アザラシ**...

### ● その音源を**人の声（言葉）**として真似させる

- OJAD 講習会アフリカツアーを計画している。
  - 南米横断ツアーは2019/2に敢行！次はアフリカ横断ツアー
  - 折角だから、現地の言葉（挨拶とか）を少し勉強していこう。

# 音真似・声真似・言真似

## 結局やりたいのは

- 脳活動を眺めずに、聴取者の声に対する反応を観察することで、音／言葉として聞いているのか、を判断したい。

## 松本先生@ガジュマルつがるからの助言

- 言葉が理解できる幼児相手なら同様のタスクは可能
- 乳児を対象に（こちらの統制下で）真似させるのは困難
- 聞こえについては、聴覚障害支援関係者に聞くとよい？

## 馬塚先生ら@理化学研究所からの助言

- 動物の鳴き声のオノマトペに追従できるか否か？
  - 本当の犬の鳴き声が「犬」であることは分かる。
  - 日本語の「わんわん」が「犬（の声）」であることが分かるか？
  - 鳴声を「日本語の音韻体系」という制約でデフォルメ＝オノマトペ？
    - 鳴声とオノマトペは、コーギー富田とコロッケの違い？
    - 言葉が遅れる子供に対してオノマトペを導入を促す例がある。





# 音真似・声真似・言真似

## 神田先生（八戸病院）からの助言

- 自閉っ子は発話のタイミング制御が健常者とは異なることがある
- 母親との声のやりとりにおいても、独特のタイミング制御に？
  - 母親と子供のやりとりでは、母親の方が真似してくる。
  - 母親の真似音声（+意図的な変形音声）に上手に答えられるのか？

## とある有名なYouTube動画 ～Twin Talk～

- [https://www.youtube.com/watch?v=\\_JmA2CIUvUY&t=11s](https://www.youtube.com/watch?v=_JmA2CIUvUY&t=11s)
  - 声はまだ言葉として機能していないが、言葉が生まれる様々な必要条件が満たされている（共同注意、操作模倣、表情表出・・・）

## 私の娘もはまった、ものまねぬいぐるみ

- <https://www.youtube.com/watch?v=uz0xfNeVUxA&t=46>
  - ものまねぬいぐるみ2匹使って、交互に音声模倣させる
  - 声を変形して（ヘリウム声的）しているので、こうになってしまう(^\_^;)

# これはありか？なしかな？

## 人の声と一般的な音の機能的差異は乳児は分らない

- 母親の声は絶対的（馬塚）
- でも、一般の声に対する選好性は初めから高いのか？

## 他人の声、母親の声、自分の声、そして、自己聴取音

- 人が聞く自身の声は気導音と骨導音との混合
- 録音音声（話者の）自己聴取音へ変形
  - 峯松他1998, 森田他2019
  - **自身が真似られていることを、音響的同一性より、確実に気づかせる。**
- **声のかけあいを誘発する「ものまねぬいぐるみ」の開発**
  - 入力音の選択：幼児の声だけに反応。ドアの音などは真似ない。
  - 声色制御：自己聴取音 → 自己録音音 → 他人の声へ
  - 韻律（発話スタイル）制御：長さ、高さ、強さをランダムに変形
    - ジャズのセッション的に、偶発的なスタイル変形を埋め込む
  - 自己聴取音から始めて、様々な声と遊ぶことを覚えさせる。

