

# 『スピーチにおける感性情報』

(1) 「間(ま)」が意図や感情の伝達に果たす役割について

小森政嗣 (KOMORI Masashi)

長岡千賀 (NAGAOKA Chika)

..... 大阪大学人間科学研究科

(2) 自然音声のイメージと知覚特性について

芝崎朱美 (SHIBASAKI K. Akemi)

..... 生命工学工業技術研究所

(3) 音声イメージの製品開発への応用について

大本浩司 (DAIMOTO Hiroshi)

石田 勉 (ISHIDA Tsutomu)

..... オムロン株式会社

指定討論者

中村敏枝 (NAKAMURA Toshie)

..... 大阪大学大学院人間科学研究科

# 自然音声のイメージと知覚特性について

生命工学工業技術研究所 客員研究員

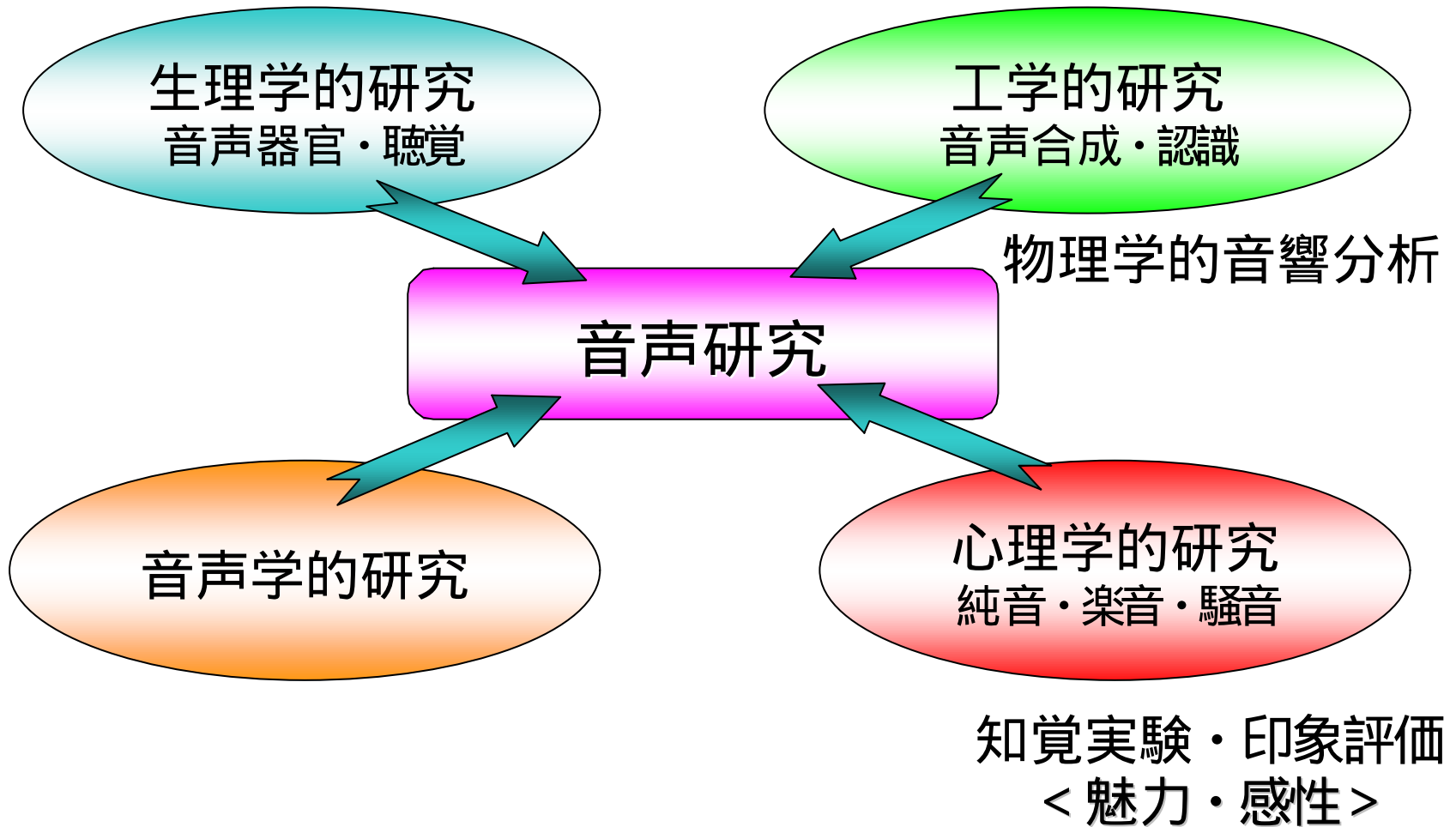
芝崎 朱美

SHIBASAKI K. Akemi

- 目次 -

1. 音声研究の流れ
2. 音声に対する高さ知覚特性
3. 高齢者の音声知覚と理解の問題

# 声質研究の流れと問題意識



# 『声』に関する生物学的背景

- **ヒトの発声**  
肺呼吸（生命維持）の副産物  
音声コミュニケーションとしての役割      文化として発展
- **その他の音声コミュニケーション**  
コウモリ・イルカ...超音波によるエコーロケーション  
サル...鳴き声による情報伝達（危険の伝達など）  
        コンタクトコール（社会性のある霊長類）  
鳥類...「歌」によるなわばりの主張・種族の識別により、無駄な争いをさける  
ヒトの新生児...クーイングへの反応により、言語行動が発達

# 『声』に関する文化的背景

## < 一次的声の文化 >



ものを書いたり、印刷したりということを全然知らない文化の『声』としての言葉に基づく文化 ... 宗教性・口頭伝承・権力

『文字の登場』による変化 < 文字が声をサポート >

## < 二次的声の文化 >

電話やラジオ、テレビ、エレクトロニクス装置によって  
ささえられた『声』としての言葉に基づく文化

### < 声が文字をサポート >

電話 ... 情報メディアとしての役割

ラジオ ... 娯楽性

テレビ ... 映像の登場



現代の声と文字  
の役割分担は？

# 文化による音声の好み

「その文化に適した発声様式にあわせて器官が発達」

- - A. トマトイス

- チベット人... よくとおる声      仲間や家畜を呼び寄せる
- 日本人... 高い音声を好む      自然な話し方には高低差が必要

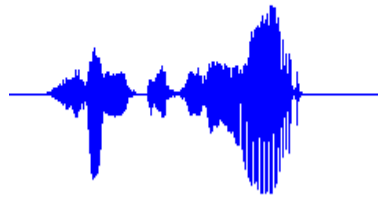
対人的印象形成において第一印象を形成するのは

- - A. Mehrabian

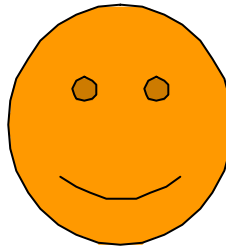
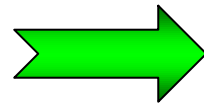
- 表情・態度...55%
- パラランゲージ(周辺言語) 38%
- 言葉の内容...7%

『声』の印象

はじめに  
: 発話全部を均等に聞いている？



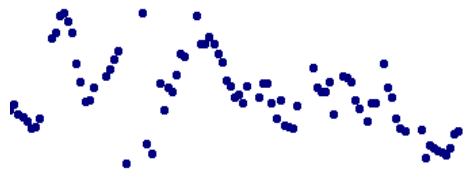
音声信号



話者属性

- 性別
- 年齢

.....etc



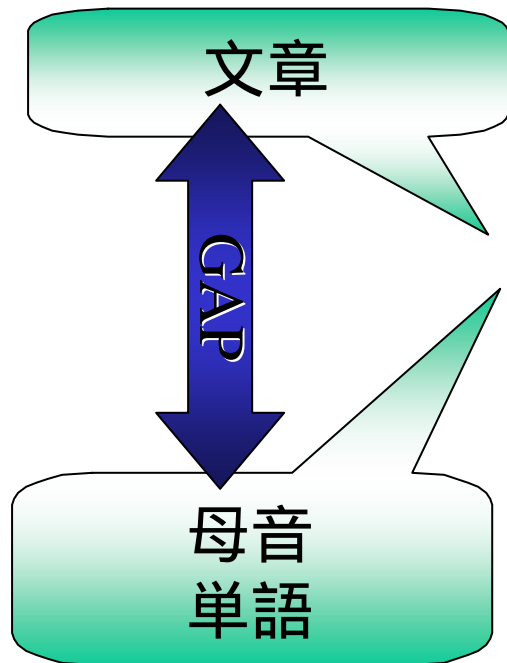
平均基本周波数

- 感情
- 意図
- 声質

.....etc

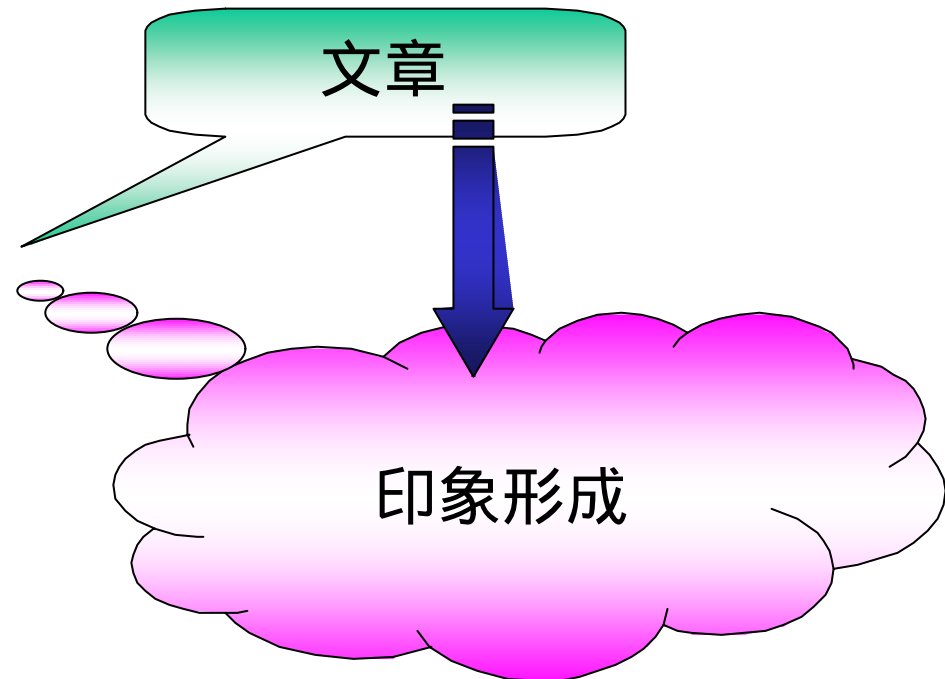
はじめに  
: 音素, 単語, or 文章?

日常会話



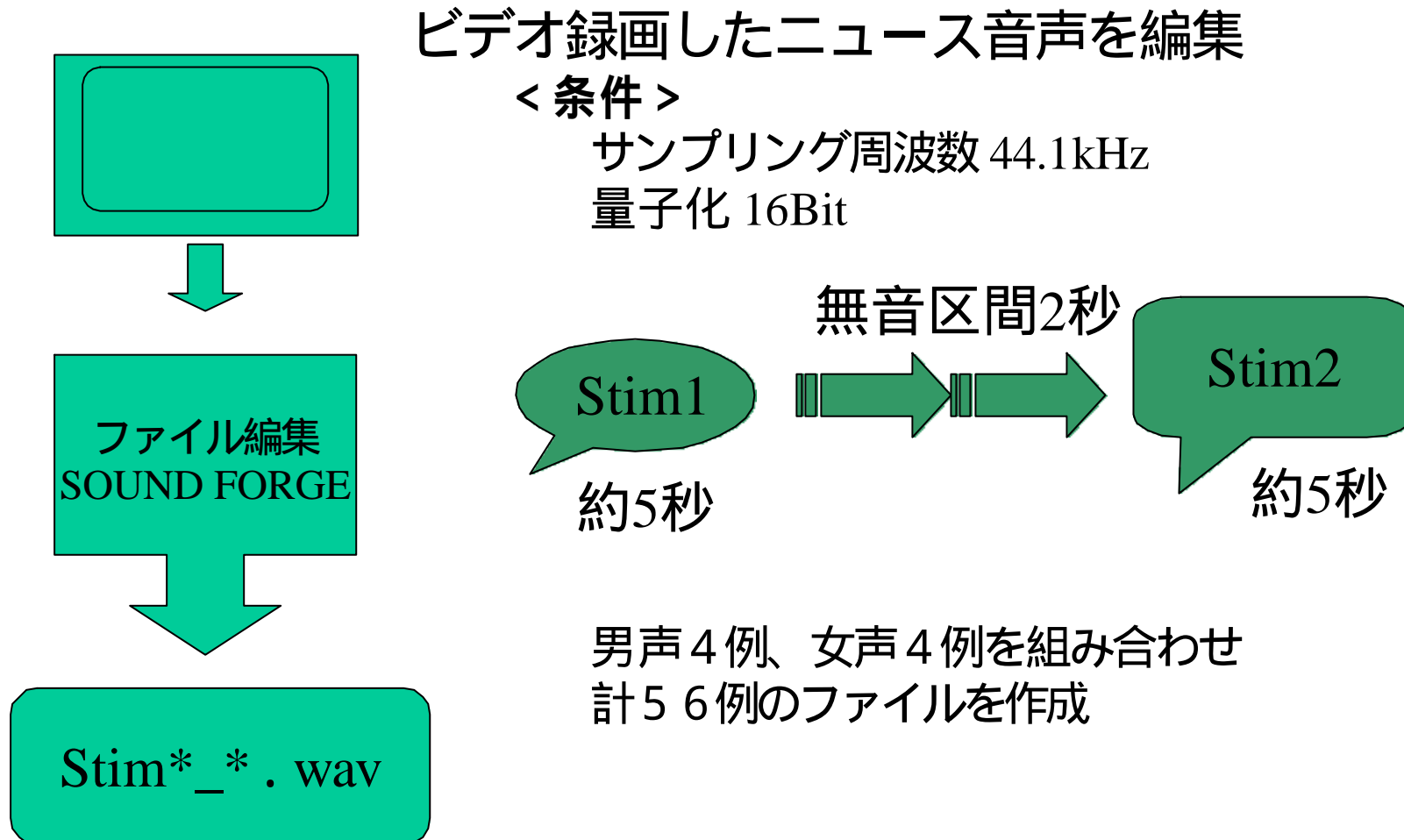
実験状況

本研究





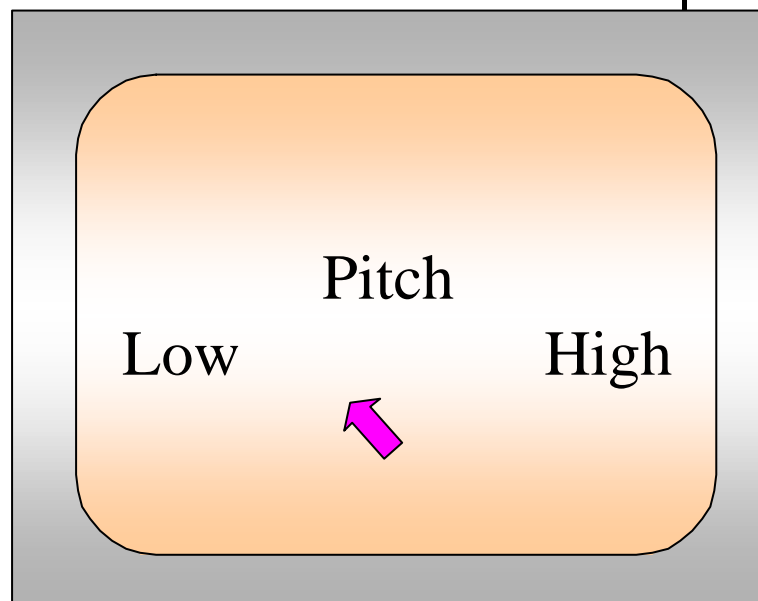
# 聴取実験：音声の高さ印象を求める



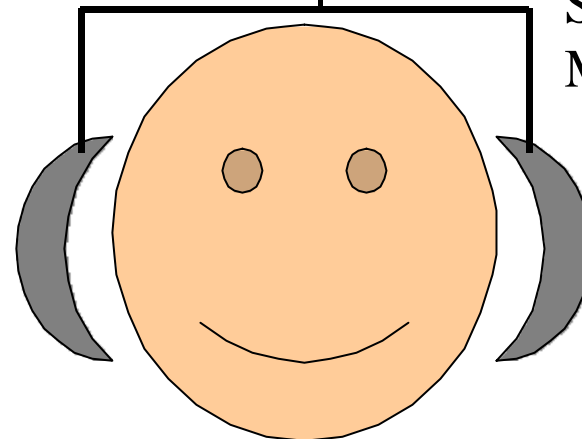
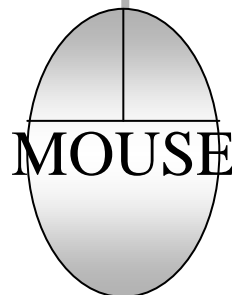
刺激音を快適に聞こえるレベルに調節

【一対比較法 (Scheffe's Test)】

刺激音声に対する高さ印象を7段階評価



NEC NX VC23/3  
SOUND BLASTER LIVE!

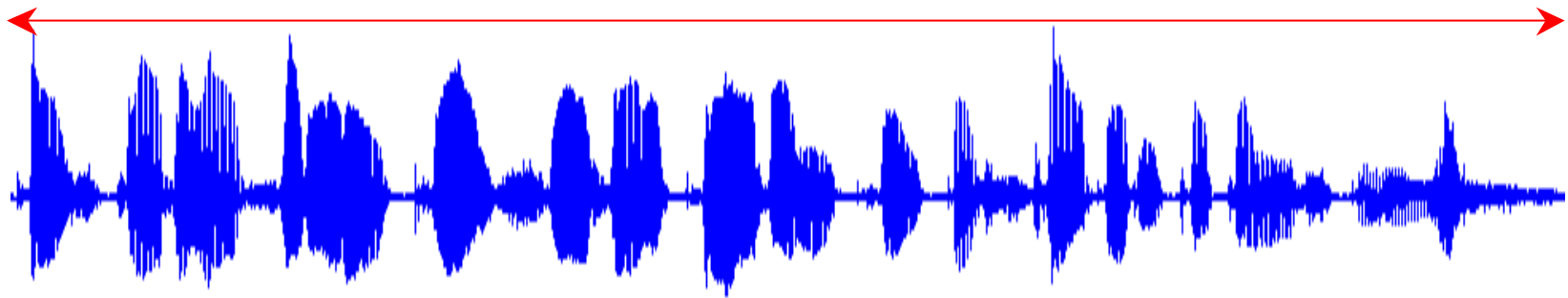


SONY  
MDR-CD1700

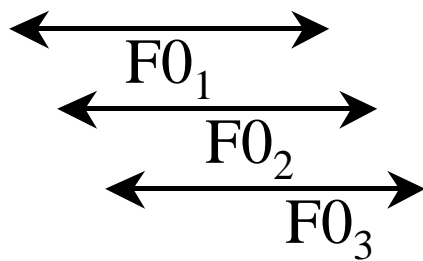
正常聴力の被験者27名  
22歳 ~ 43歳 (mean 35)

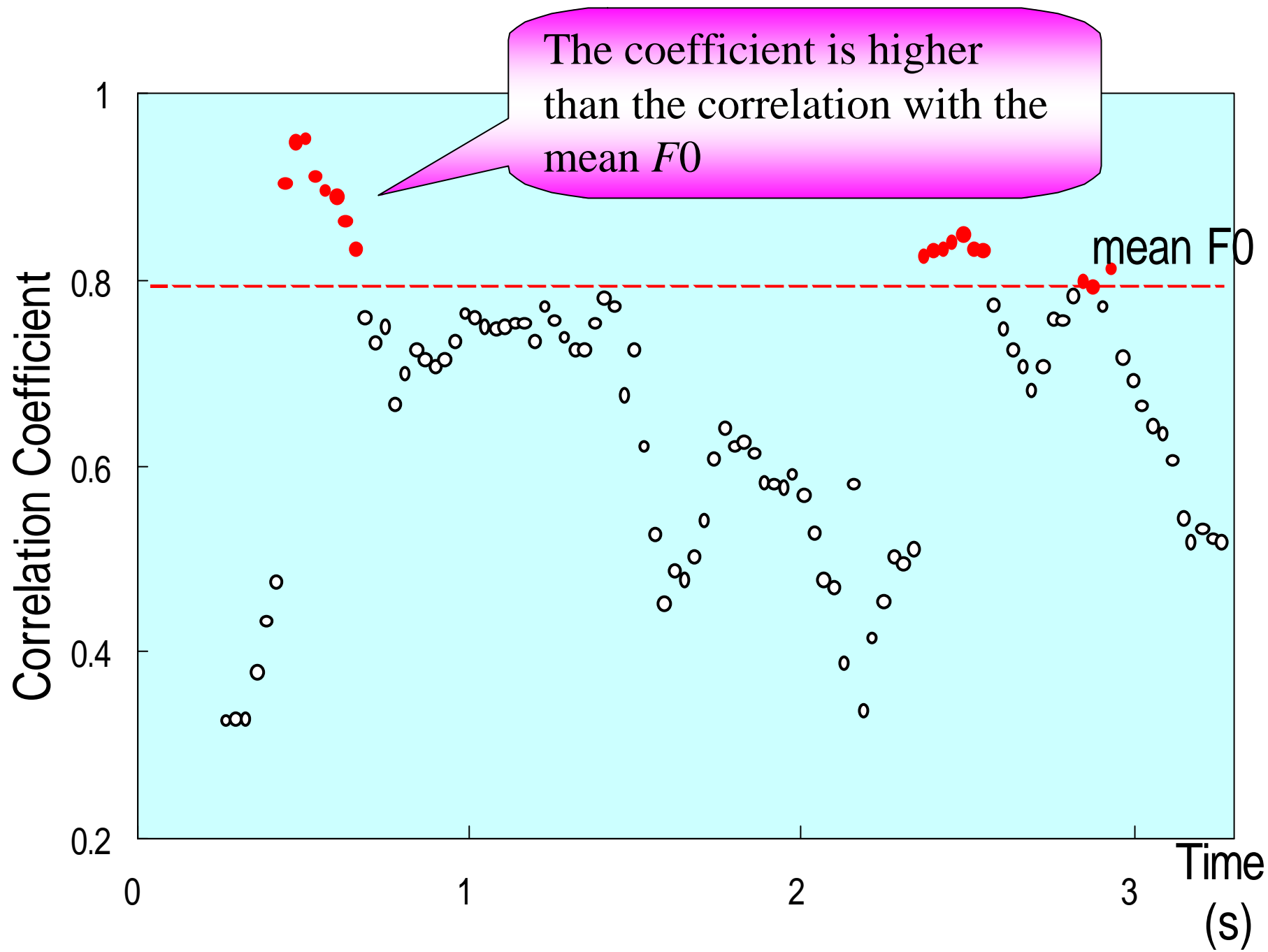
# 結果(分析手続き)

発話全体の平均F0



30ms ごとの平均F0





## 結論

1. 高さ印象は、発話全体の平均F0よりも短時間平均F0とより相関が高い。
2. 聴取者は、とくに発話開始部分に注目し、印象を形成していると考えられる。

高さ印象は、  
文章の開始部分で決定される

## “ Why the subjects focused on the beginnings of the sentences?”

### (1) Linguistic reason:

“Japanese speech tends to have a high frequency at the beginning and a low frequency at the end of a sentence.”

### (2) Psychological reason:

“Human infants have a preference for high-frequency sounds.”

The high frequency segments at the beginning of the sentences attracted the subjects' attention.

Human memory      People focus on certain parts in an utterance when listening to speech.

### (3) Methodological reason:

The method of paired comparison      Subject compared the ending of one stimulus item to the beginning of the next?

A preliminary experiment supported the present results.

People pay more attention to the beginning segment of speech than other segments.



## possible methods of improving speech recognition in elderly people

By slowing down the beginning segments of speech, it might be possible to extract more information.

こんな応用ができるかも！



合成音声の経済効率を良くする

注目される部分の精度を上げ  
注目の低い部分の精度を下げる

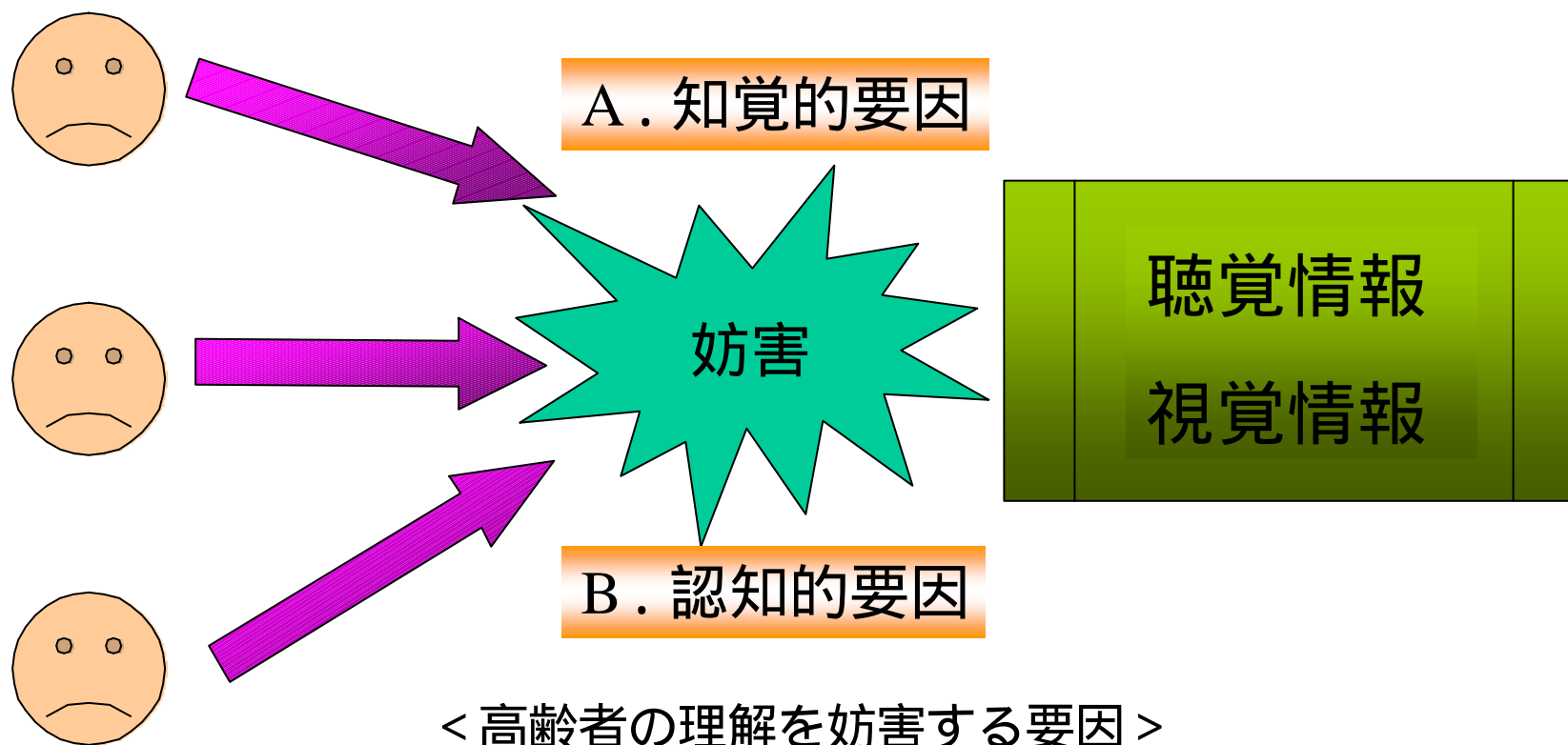




NHK放送技術研究所

<http://www.str1.nhk.or.jp/toppage.cgi>

# 【研究の背景】 高齢者が情報を得るときに起こる困難



# 音声の聞き取りが困難になる条件

特に高齢者で影響が大きい

1. 妨害となる音が存在する場合
  - A) 背景騒音が強い
  - B) 部屋の残響が大きい
  - C) 他の話者の声と競合する
2. 音声がひずんでいる場合
  - A) 周波数帯域が狭い(電話など)
  - B) 音声が途切れる(受信状態の悪いラジオなど)
  - C) 音響機器でのひずみ(マイクへの過大入力など)
3. 発話者側に問題がある場合
  - A) 発話速度が速い
  - B) 発音が不明瞭
  - C) 質の低い合成音

## 高齢者の言語理解を妨害する要因

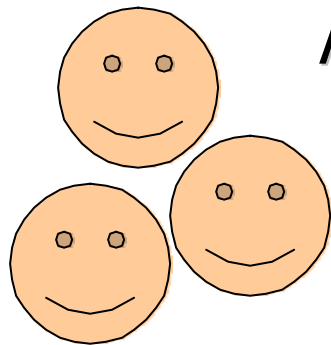
### A. 聴覚能力低下に伴う知覚的要因

- ・ 高周波数の聞き取りが困難に  
子音の聞き取りが悪くなり、言語理解能力が低下
- ・ 周波数選択性の低下  
騒音からターゲットを聞き分けられない

### B. 記憶容量や処理能力低下に伴う認知的要因

- ・ 意味的プライミング  
若齢者との差は、視覚ではあり、聴覚ではなし  
反応時間に差はあるが、質的な差はない
- ・ 視聴覚は相補的？

# 高齢者による情報アクセスを容易に



Accessibilityを高く！



情報発信源  
Webページなど

あらゆる人が特別な措置なし  
で利用できるように最初から  
配慮して作ることで、さらに多く  
の人が利用可能にする情報設計

# 展望と技術応用

Accessibility

視覚

聴覚

情報機器の小型化・携帯化により  
ディスプレイ依存度が増加

視覚の負担を軽減  
音声による情報接触要求の高まり

高齢者の利用を容易に

健常者の利用を拡大  
(視覚・聴覚の切り替え)  
次世代の移動体通信機器、  
ネットワーク端末などへの  
応用

# Just Suggestive (JS) 原則 を応用した場合の 問題点を検証する

## 【視覚】

Just Suggestive (JS) 原則 を守ることが理解促進  
たとえば、「見出し 要約 本文」のような3段階構造

ページ内デザインを統一

リンクの張り方を統一

## 【聴覚】

リズム感を生み、  
位置感覚を与える

声質の変化を統一  
(本文は男声、リンクは女声)