

# 第四回音響・超音波ジョイント研究会

平成 24 年 3 月 11 日

金沢大学サテライトプラザ

主催 一般社団法人 日本音響学会 北陸支部

日本音響学会北陸支部第四回音響・超音波ジョイント研究会

日時： 平成 24 年 3 月 11 日 (日) 15:10 ~ 18:00

場所： 金沢大学サテライトプラザ 1 階交流サロン (金沢市西町三番丁 16 番地)

プログラム：

0. 開会の辞 森川 博由

< 15 : 10 ~ 16 : 25 > (前半)

座長： 山田 真司

1. PCに於けるGoogle音声認識の利用

山崎信久 (ウィンドクラフト), 船田哲男 (元金沢大)

2. JAIST・鶴木研究室の紹介

鶴木祐史 (北陸先端大)

3. 研究紹介 —口笛と骨伝導—

森 幹男 (福井大)

— 休憩 10分 —

< 16 : 35 ~ 17 : 50 > (後半)

座長： 森 幹男

4. 超音波を使う新しい教育研究領域の創成の試み

得永 嘉昭(金沢工大)

5. キャラクターボイスの印象について：萌える声、萌えない声

山田真司、塚本将成、加藤典、小柳拓巳 (金沢工大)

6. 2次元デジタルフィルタを用いる指向性制御方法の応用例

山川昭彦、伊藤栄太、齋藤毅、三好正人 (金沢大)

7. 閉会の辞 青木 茂明

# PCに於けるGoogle音声認識の利用

山崎信久 (ウィンドクラフト)  
船田哲男 (元金沢大)

# PCに於けるGoogle音声認識の利用



Google  
日本



D getrunken hatte,  
nich." Da segnete  
lle! Völker müsse  
ht sei, wer dir f  
afob hinausgegan

Wind-Craft (風工房)

- ・ 山崎信久 (ウィンドクラフト)
- ・ 船田哲男 (元金沢大)

# 音声認識の種類

方式	特定話者方式 (特定話者での学習が必要)		不特定話者方式 (学習不要)	
種類	連続認識	単語認識 (限定語認識)	連続認識	単語認識 (限定語認識)
例)	Windows搭載 ドラゴンスピー チ	音声コマンドTool	Julius AmiVoice Google音声認識	Julian(Julius統合)
用途例	口述筆記 翻訳	コマンド操作	口述筆記 翻訳	自動応答

Google音声認識: 単語認識に近いが限定語ではない

# 音声認識の状況

不特定話者認識ソフト例

・Julius（京都大学等）

○オープンソースで高速な認識

△辞書が固定

学習機能がないため認識率が向上しない



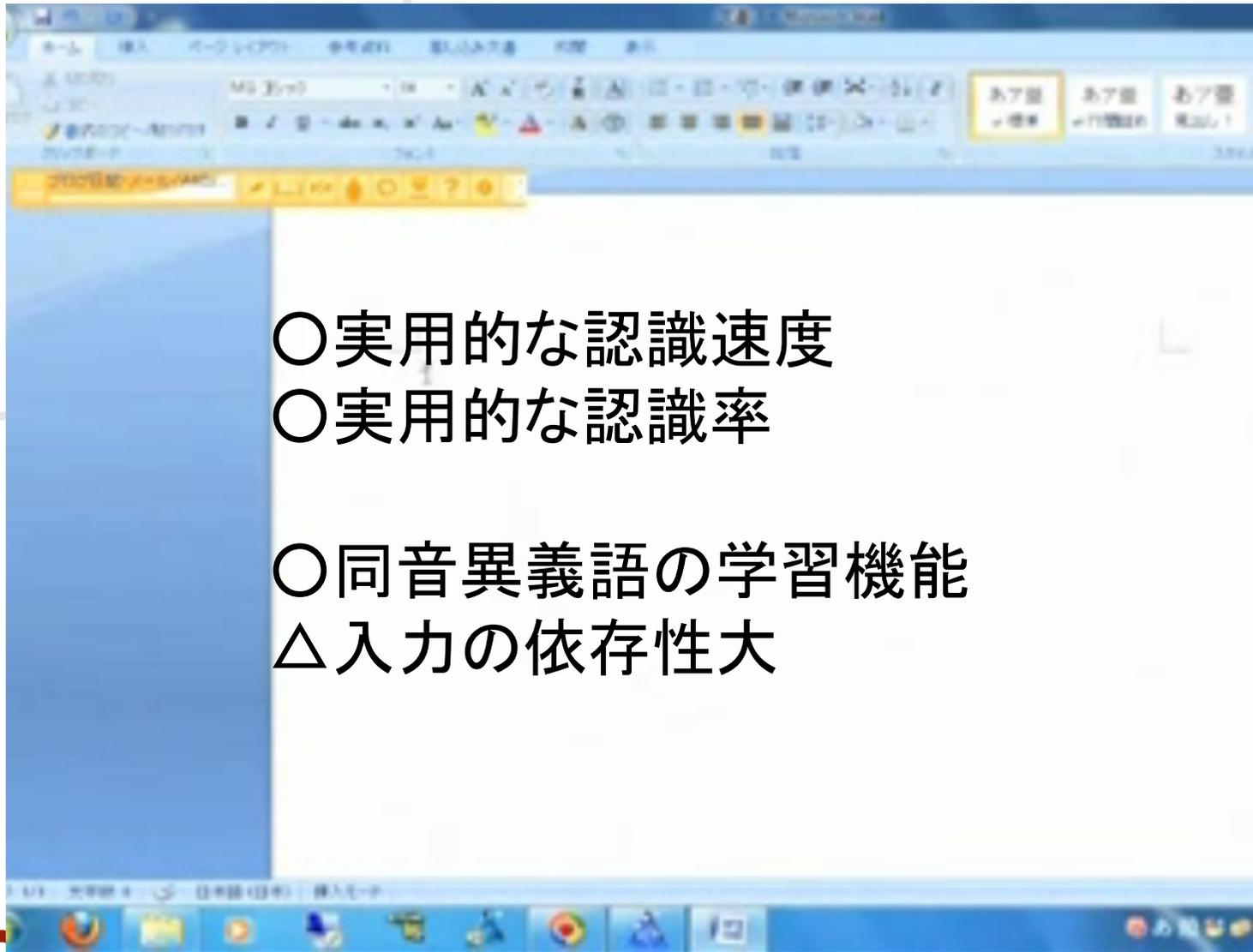
・AmiVoice（アドバンスド・メディア社製品）

○実用的な変換速度と認識率  
学習機能があり、認識率向上可能

△入力依存性が高い  
ヘッドセット以外では認識率低下  
開発者利用は困難

\* その他  
翻訳ソフトの「超速通訳 ツーシル」などもある。

# AmiVoiceのデモ(メーカー)



# Googleの音声認識の登場

- インフラの変化  
ネットワークの進歩(高速化)  
スマートホンの台頭(端末小型化)



- 使用者の変化  
通信による情報参照  
キーボードレス時の入力手段



- Googleが音声認識を発表  
2009年にスマートフォン向けから開始



# Google音声認識デモ

○実用的な認識率

キーボードより高速な入力

○入力の依存性小

通常環境で良好な認識

△語句レベルでの認識

長文に不向きだが検索には十分

△認識のタイムラグ

話しながらの認識には不向き

# Google音声認識(1)



## 【経緯】

- ・2009年末Googleが発表  
スマートフォン向けとしてiPhone・Android  
用に発表

## 【特徴】

- ・不特定話者認識(学習不要)
- ・実用的な認識率(検索で語彙の予測)

## 【条件】

- ・ネットワーク接続(要高速)

## 【ポイント】

サーバで認識する

- ・単語認識に近いが特定語ではない  
(良く検索する語句は認識率が高い)
- ・速度は話す速度には追いつかない  
(区切りが必要になる)



# Google音声認識(2)

## 【スマートホンの状況】

iPhone: 主に検索用のアプリとして提供されている

Android: 上記以外に

- ・「音声認識Intent(入力ダイアログ風)」として開発利用可能
- ・IME版の開発によりどのアプリでも音声入力可能(ベータ)

google-voice-typing-integration

<https://code.google.com/p/google-voice-typing-integration/source/browse/#git%2FVoicelmeDemo>



## 【Googleの提供内容】

### ・iPhone

アプリ内部での利用のみ

(Googleの提供アプリのみの利用になる)

### ・Android

開発者が利用可能

音声認識をIntent(ダイアログ風)として呼び出し可能

IME版が利用可能

これによってあらゆる入力部分で利用可能

メモでも入力可能・ショートメールにも入力可能

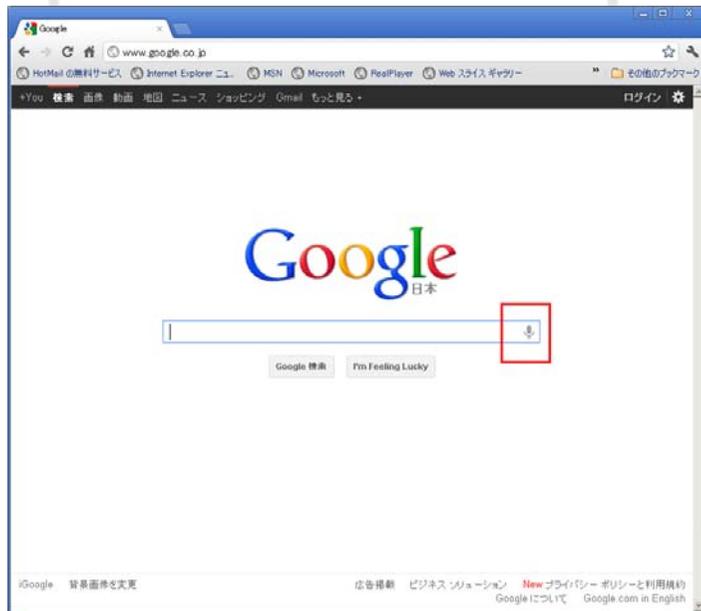


# Google音声認識(3)

## 【PCの状況】

- ・2011年春PC向けChromeに搭載

HTML5の音声入力拡張として動作



音声入力の手法)

HTMLで記述する。(HTML5)

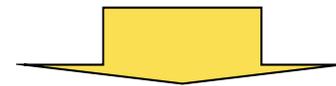
```
<input x-webkit-speech type="text" />
```

で入力が扱える。

条件)

上のタグを解釈できるブラウザに限る。

現状はGoogleChromeのみ



そのままでは音声検索がしにくいので以下のエクステンションを入れる

**Speakable Textareas**

<http://userscripts.org/scripts/show/108011>

# スマートホンと同じ環境の実現

# GoogleChromeデモ

○スマートフォンと同じ環境

マイクの制限小

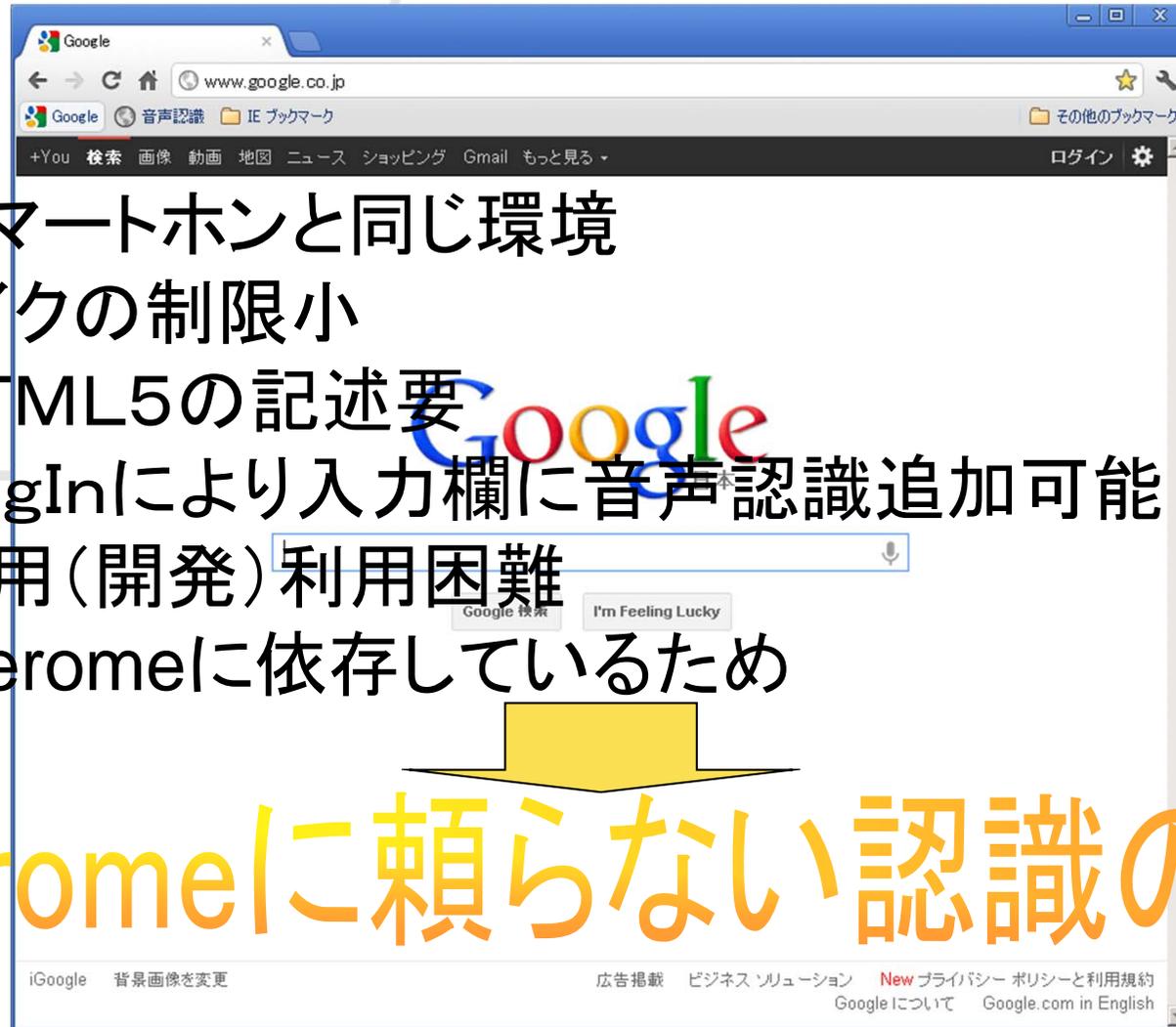
△HTML5の記述要

PlugInにより入力欄に音声認識追加可能

×応用(開発)利用困難

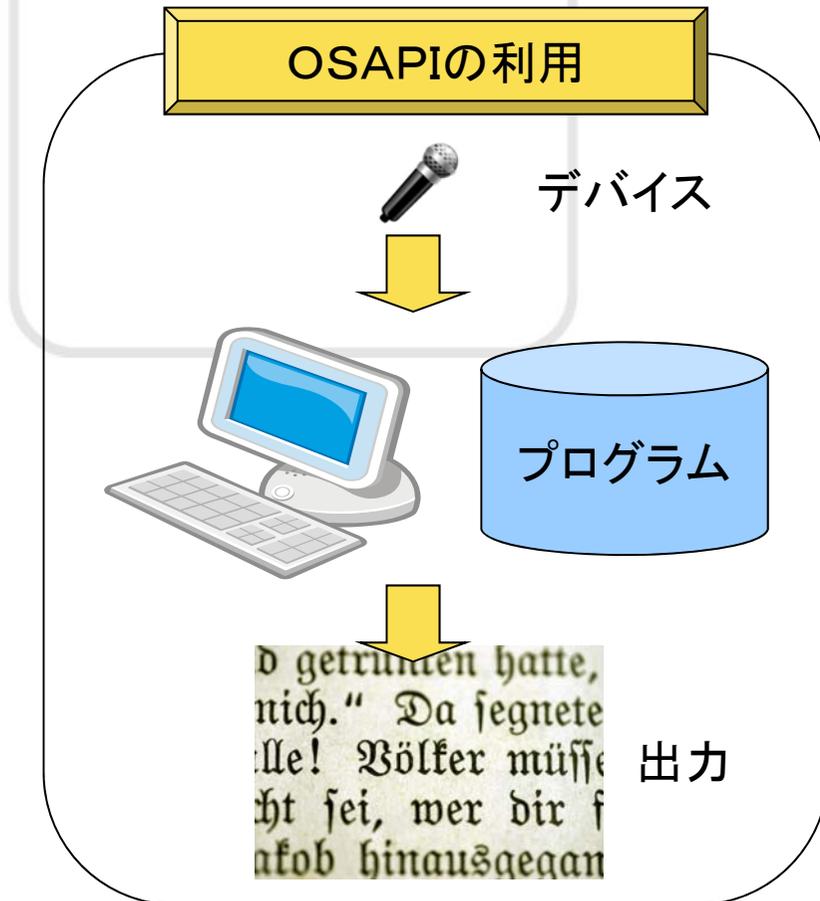
Chromeに依存しているため

Chromeに頼らない認識の実現



# アプリケーションの現状(1)

OSAPIからWebAPIへ



構成)  
処理はすべてOS上のプログラムで  
処理される。

長所)

ネットワーク接続が不要である

短所)

- ・処理能力がPCで決まる
- ・複雑なプログラムでは多くの知識を必要とする

開発分量が膨大になる

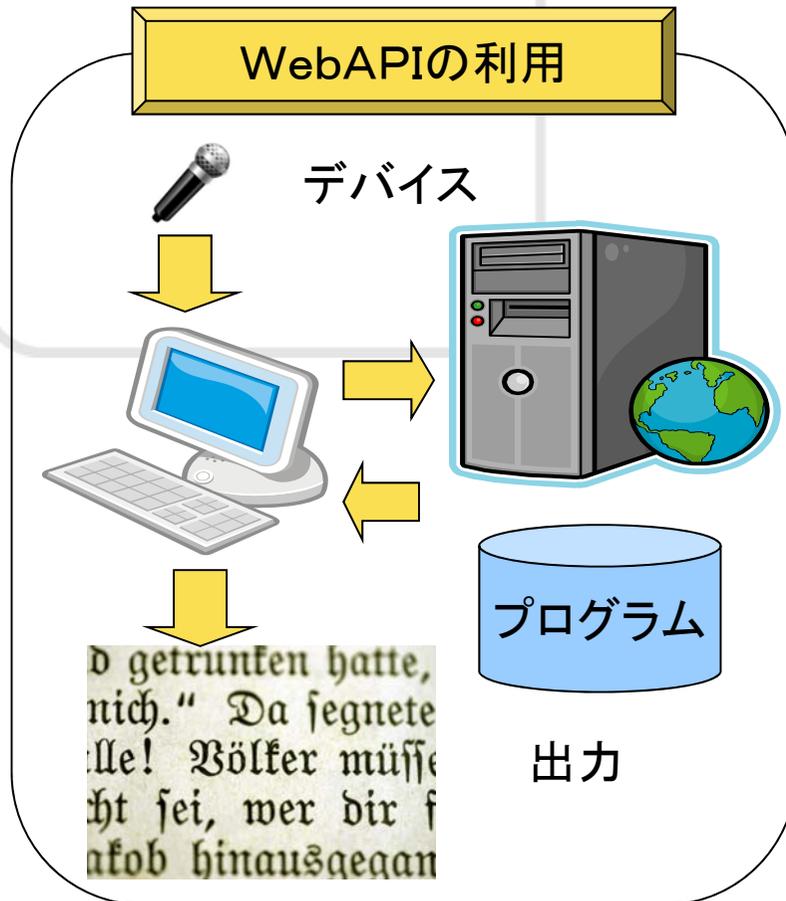
専門知識が必要になる

- ・OSによる依存性が高い

機種への移植が煩雑

# アプリケーションの現状(2)

OSAPIからWebAPIへ



構成)

処理の中核はサーバ上のプログラムで処理される。

長所)

- ・PCの処理能力に依存しない
- ・複雑なプログラムが不要  
パラメータを送信し結果を受け取る
- ・OS依存性が低い  
HTTP通信があれば可能

短所)

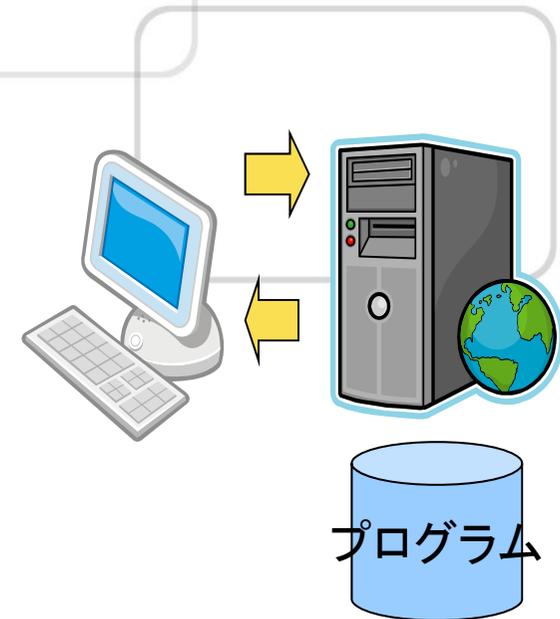
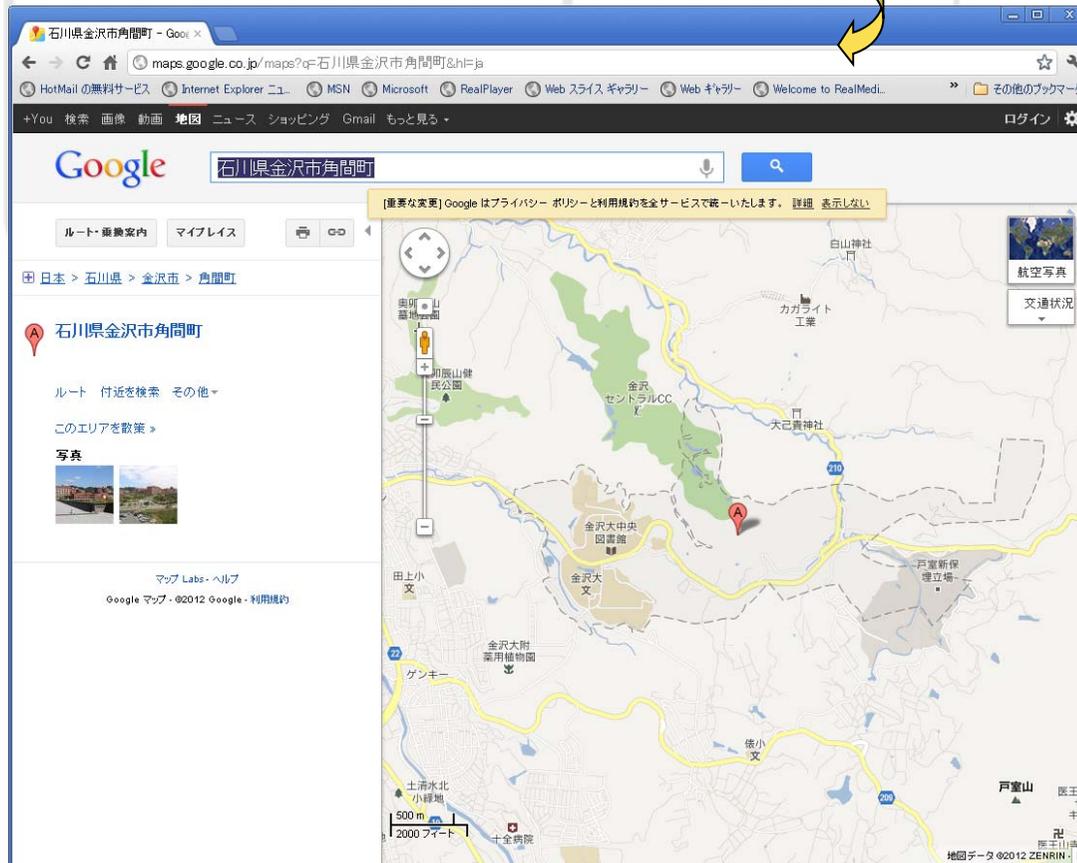
- ・ネットワークが必須になる  
パラメータによっては高速必須
- ・中身はブラックボックス  
プログラムは相手次第

# WebAPIの例(1)

例1)

住所の地図を表示する(金沢大学の住所)

<http://maps.google.co.jp/maps?q=石川県金沢市角間町&hl=ja>

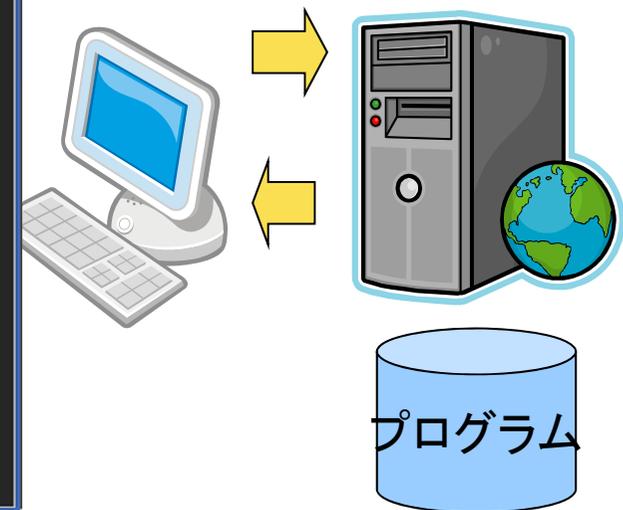
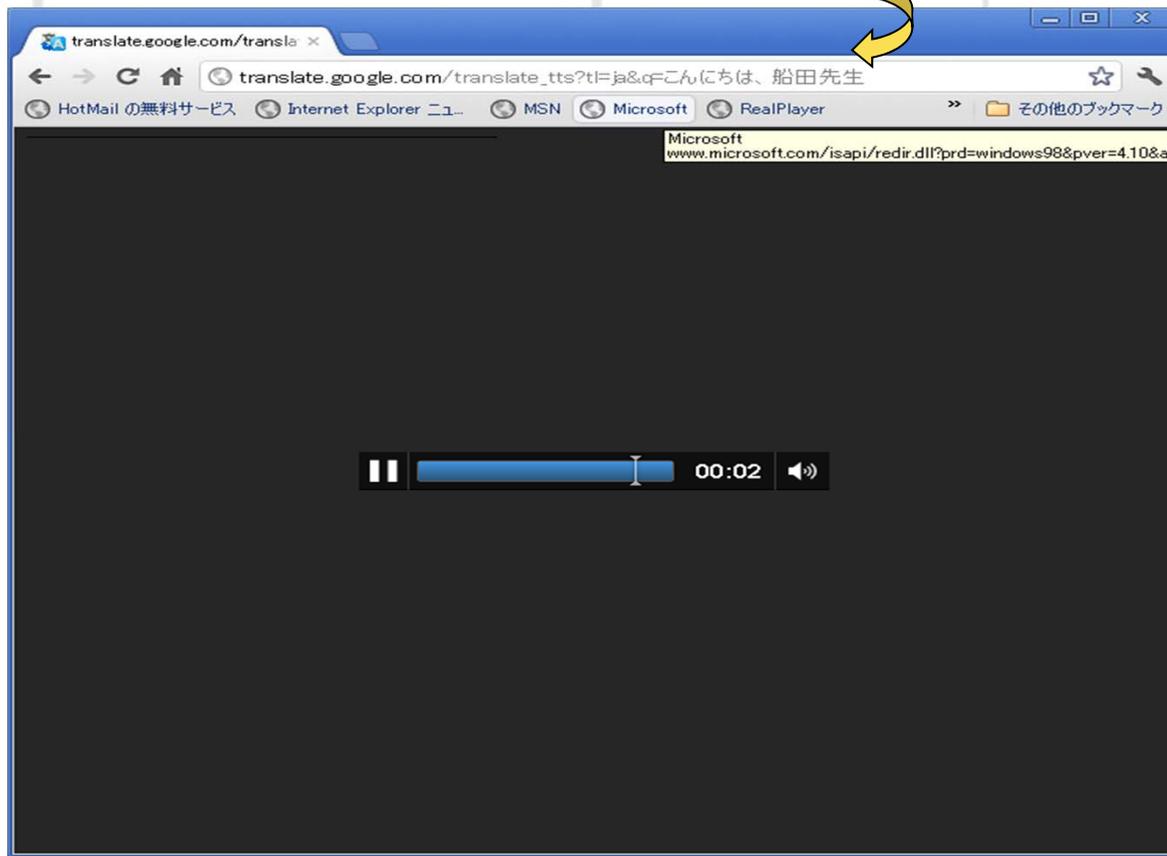


# WebAPIの例(2)

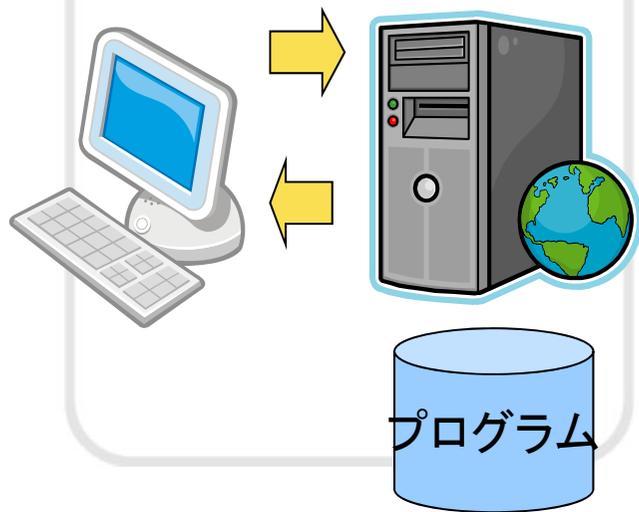
例2)

TTS(テキストの読み上げ)

[http://translate.google.com/translate\\_tts?tl=ja&q=こんにちは、船田先生](http://translate.google.com/translate_tts?tl=ja&q=こんにちは、船田先生)



# WebAPIの使い方まとめ



入力)

GETやPOSTで特定のURLに通知する

出力)

結果として定められた形式で出力される

例)

・地図

入力:住所や緯度経度(座標)など

出力:画像やマーカ

・TTS(読み上げ)

入力:テキスト(+言語)

出力:mp3形式のファイル

## 【WebAPIの特徴】

・基本はブラウザを前提としている

記述はHTMLやJavaScriptで行う。

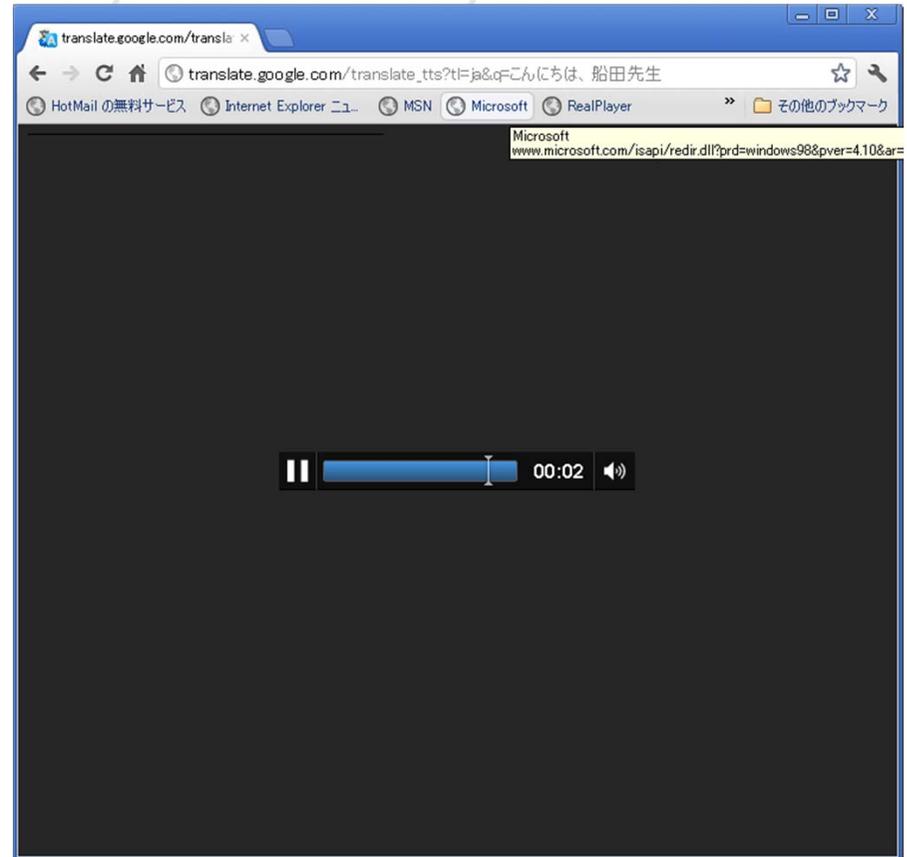
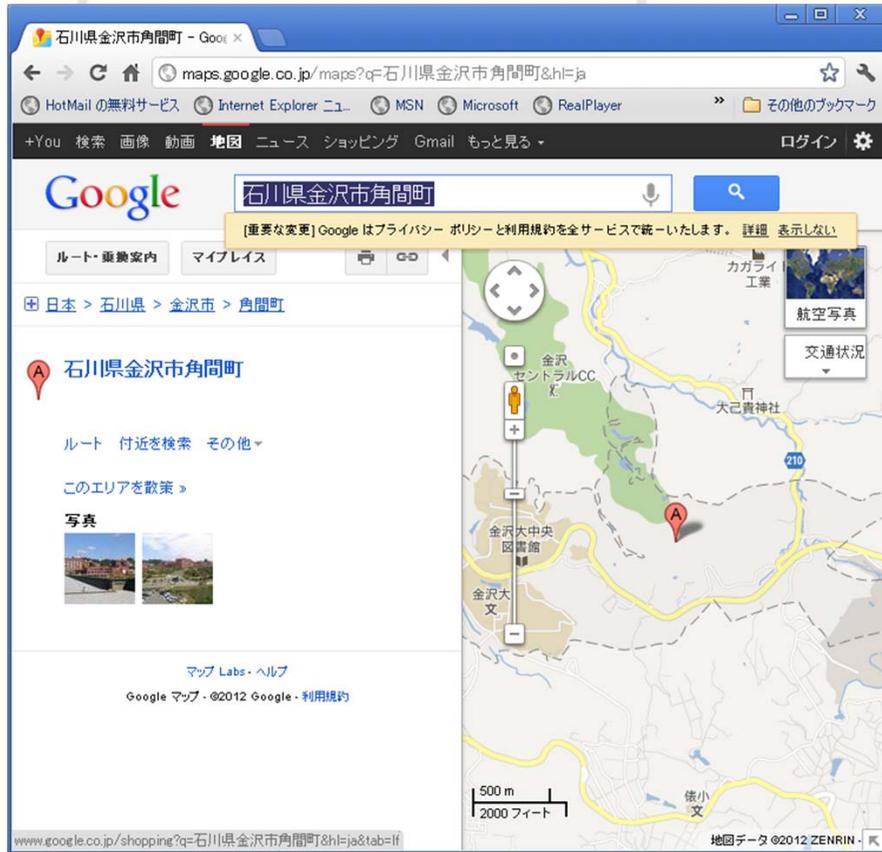
HTMLも時代に応じて拡張がなされている(HTML5)

・ローカルPCのプログラムでも

GET/POSTでHTTP通信を行うことでパラメータを送信できる

返答された結果を処理することで、プログラムとして動作できる

# WebAPIのデモ



# PCに於けるGoogle音声認識(1)

## 【Chromeの動作分析】

- ・HTML5の構文解釈
- 音声入力→音声をサーバに  
入力表示←サーバから結果受信

## 【具体的な内部動作】

- ・マイクから音声データを生成  
(音声ファイル化する)

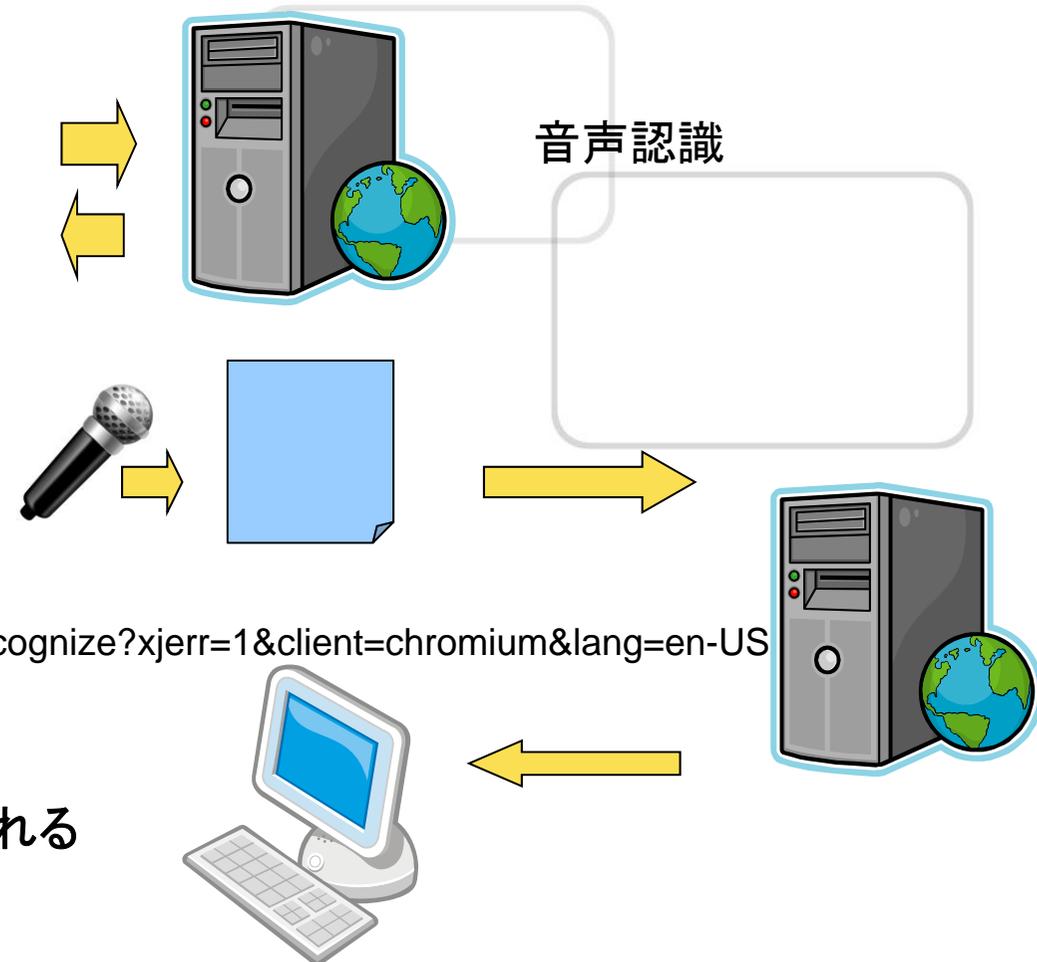
- ・それを以下にPOSTする

<https://www.google.com/speech-api/v1/recognize?xjerr=1&client=chromium&lang=en-US>  
(引数に関しては後述する)

- ・認識結果はJSON形式で返答される  
(結果の内容に関しては後述する)

原理サイト

<http://mikepultz.com/2011/03/accessing-google-speech-api-chrome-11/>



# PCに於けるGoogle音声認識(2)

## 【技術内容】

### 1) 音声入力をファイルにする

語句を区切るため「無音検出」で区切る

ファイル形式はflacまたはSpeex形式(flac形式が無難)

サンプリングレートは8KHzまたは16KHz。

16KHzのflacの時Content\_Typeは"audio/x-flac; rate=16000"を指定することでGoogleに通知。

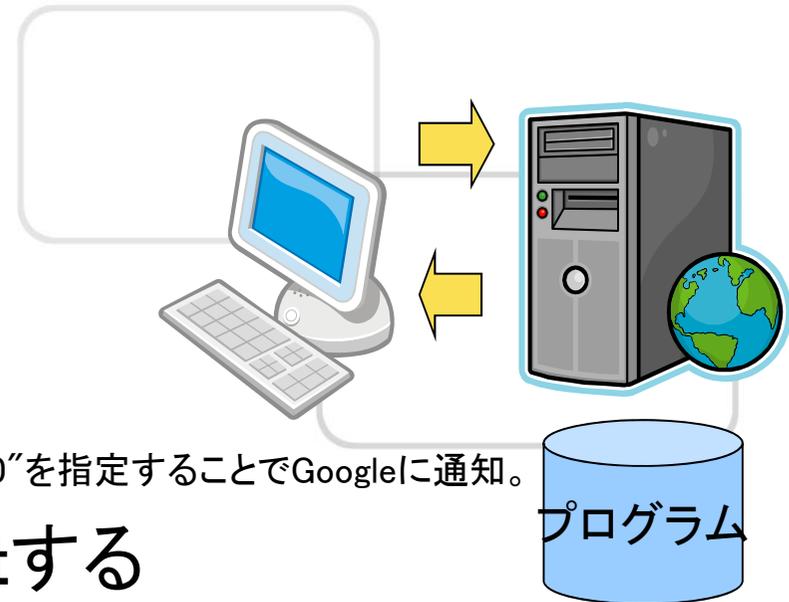
### 2) 音声ファイルをURLにPostする

引数は言語指定(英語はlang=en-US,日本語はlang=ja-JP)候補数指定(maxresult=10)などがある。

### 3) 結果をJSON形式で受け取る

結果はstatus(0以外は認識できず)id(順序判定用)に候補指定の分以下が付属する。

hypotheses(予測):utterance(発声)が認識結果、confidence(信頼性)が確からしさ。



解説サイト

<http://sebastian.germes.in/blog/2011/09/>

# PCに於けるGoogle音声認識(3)



## 【応用例】

### IEでの音声認識

#### 1) 音声入力ファイル作成

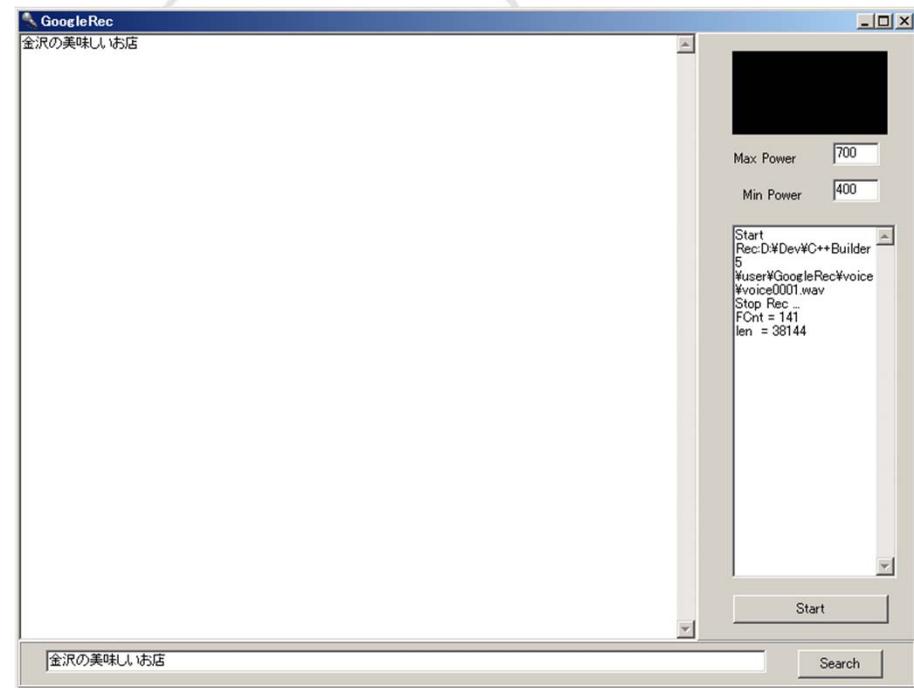
入力からFlashにてwavファイル作成  
中継サーバにてwav形式をflac形式に変換

#### 2) 音声ファイルのPost

中継サーバからPHPでGoogleにPost。

#### 3) 結果を受け取る

中継サーバで結果のトークンを受け取りFlashに通知。



作成ソフトのデモ

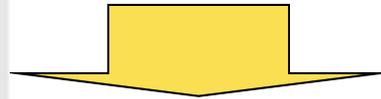
IEデモサイト

<http://select-items.net/pfu/google/speech/>

# 終わりに(まとめ)

スマートホンの発達(PC+モバイル)

ネットワークの発達(常時接続 & 高速化)



- 機種依存性の少ない開発

HTML+JavaScript

またはそれに類する通信+処理

- WebAPIの活用が加速

通信前提のアプリケーション開発

複雑な内容はサーバがサービス提供

例) 音声認識・音声合成・翻訳・OCR・地図...



# J A I S T ・ 鵜木研究室の紹介

鵜木祐史（北陸先端大）

# 情報科学研究科 人間情報処理領域 音情報処理分野

- ★講座HP: <http://ais.jaist.ac.jp/index-j.html>
- ★教官HP: <http://www.jaist.ac.jp/~unoki/>
- ★ライフスタイルデザイン研究センター  
<http://www.jaist.ac.jp/kscenter/index-j.html>

## 鵜木研究室



# 鵜木研究室 構成員・研究体制

- 准教授：鵜木祐史
- 助教：宮内良太
- 研究補助員（2名）
- 博士後期課程（3名）
- 博士前期課程（12名）（うち東京サテライト1名）
- 修了生：15名（2006年～） 赤木研（7名）
- 共同研究者：
  - Dr. Xugang Lu (NICT)
  - Prof. Roy D. Patterson (Univ. Cambridge)
  - Prof. Dr.-Ing. habil. Rüdiger Hoffmann (TUD)
  - Dr.-Ing. Rico Petrick (Tech. Univ. Dresden)

# Masashi Unoki, Ph.D

- Master/Ph.D @ JAIST (Prof. Akagi) (1994-1999)
- Visiting Researcher @ ATR-HIP (Prof. Irino) (1999-2000)



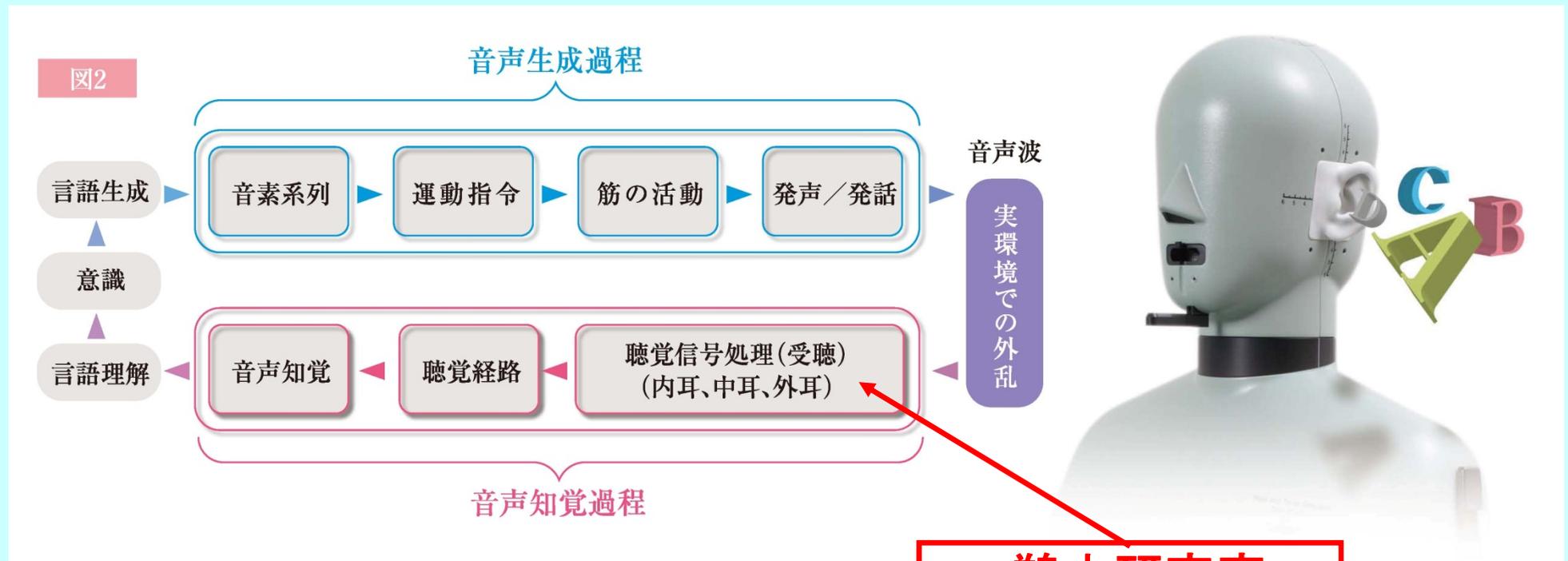
- Visiting Researcher @ Cambridge Univ. (2000-2001)
- Associate@JAIST (2001-2005)
- Associate Prof. @JAIST (2005- )

# 人間情報G／音情報処理の研究ターゲット

## ■ 基本路線

- **話す・聞く**は人間の営み → 人間を知り、そして、営みを記述することで、高度の音声処理システムの実現を目指す

## ■ 研究範囲：音声知覚，**聴覚信号処理**，音声生成



<赤木研究室と一緒に活動>

# 「耳」付き計算機の登場が、 音信号処理に革命を起こす!?

雑音や残響がある環境でも、人間は聞きたい音を聞くことができる。  
「もしも計算機上に、聴覚と同じ機能を持つ信号処理システムを構築できたら、さまざまな音信号処理に応用できるのでは？」  
という考えの下で研究を進めているのが、鷗木准教授の研究室だ。

## 耳に相当する機能を 計算機につけてみたい

ICレコーダーでインタビューを録音した音声を聞き直してみると、インタビュー時には気にならなかった雑音や残響で、話し手の声が聞き取りにくいことがある。それなのになぜインタビュー時には話し手の声がクリアに聞こえたのだろうか？ それは、私たち人間の聴覚（正確には、聴覚系+脳）が、雑音や残響を排除し、狙った音だけを聞き取るという、非常に優れた機能を有しているからだ。

けれども、「私たちの聴覚は非常に優れている」ということはわかっているのに、「どのような信号処理によってそれが実現されているのか」ということは、実ははまだ解明されていないことが多々ある。自分たちの体の一部であるにも関わらず、である。

「もしもその信号処理のしくみが全て解明できれば、音声認識や補聴システム等の音声信号処理に応用できると思うんです」と鷗木准教授は語る。そして、自らの研究スタンスを図解すると「 $\pi$ 型」(図1)になると言う。

「『聞く』『話す』といった音声コミュニケーションは、「ことばの鎖」(図2)の過程で行われています。情報科学研究科の音情報処理分

そこに行き着くためには、2本の柱による研究アプローチが必要です。1本目の柱は「人間の耳をまず良く知る」というアプローチです。学問的には、聴覚の各部位の機能解明を目指す『聴覚生理学』や、音の物理的変化に対する反応(知覚変化)を調べる『聴覚心理学』の研究領域です。もう1本の柱は「人間の耳を模擬して計算機上に記述する」というアプローチです。こちらは、学問的には『情報科学』あるいは『工学』の領域です。つまり、**聴覚生理学、聴覚心理学、情報科学の3つがオーバーラップする部分**が、私たちの研究領域です。JAISTの先生方や国内外の研究者に協力を仰ぎ、聴覚心理学・情報科学を本学で実施し、聴覚生理学に関しては海外の研究成果等を参考にしながら追求しています」

## 驚くほど快適な補聴器や 音声透かしの普及も間近

現在、鷗木研究室ではさまざまなプロジェクトが進行している。

そのひとつが、「**聴覚フィルタバンクの構築**」だ。聴覚器官では、蝸牛の基底膜で音を周波数成分に分解し、電気信号(パルス)に変換して神経に伝えている。この機能は、簡単にいうと帯域通過

耳はそもそもなぜ必要なのか？それが知りたくて研究を続けているんだと思う。

情報科学研究科 准教授

鵜木 祐史 Associate Professor Masashi Unoki

**【趣味】** 学生時代はラグビーをやっていました。根気よく研究を続けるには、体力や強い精神力が必要ですが、ラグビーを通じて研究に向かっていく力が身についたと思います。

**【この研究を続ける理由】**

聴覚はなぜ必要なのか？ どういう機能を持っているのか？ 科学を志している人間として、その一端を垣間見たいという純粋な理由で研究を続けているんだと思います。でも、生きている人間の耳を解剖するなんてことは不可能ですよね。そんな中で人間の耳の機能を解明するためには「人間の耳の各部位の機能は何か？」「人間の耳と等価なものを作るにはどうすればいいか？」ということを丹念に調べて、データを積み重ねていくしかない。遠回りに見えるけれど、実はそれが本質的でいちばんの近道なのかもしれないと思うんです。

**学位**

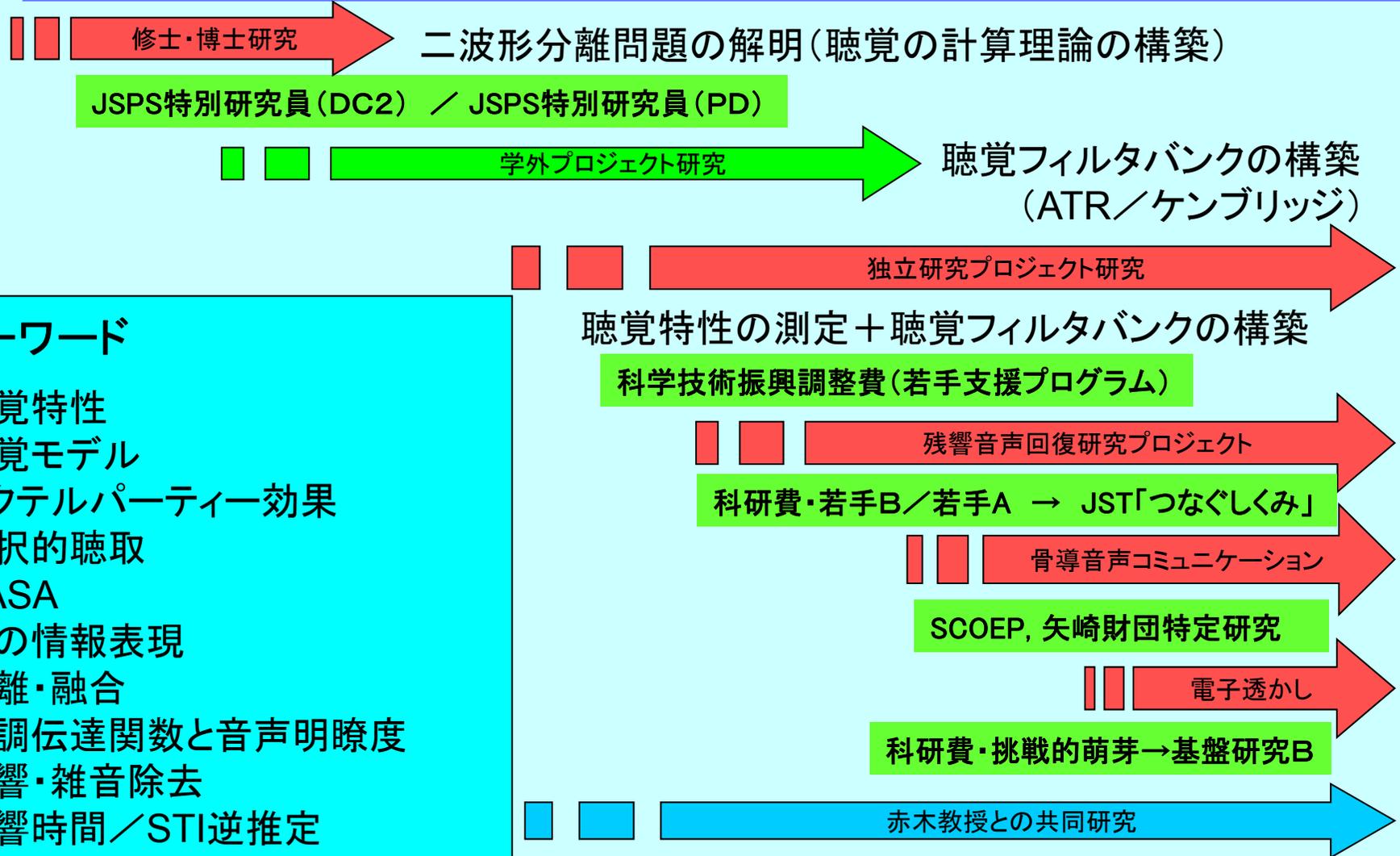
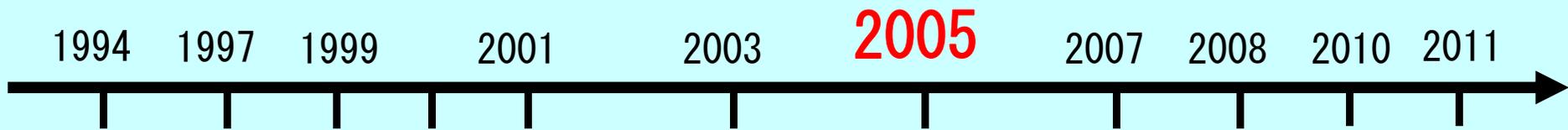
北陸先端科学技術大学院大学 情報科学修士(1996) / 北陸先端科学技術大学院大学 情報科学博士(1999)

**経歴**

日本学術振興会 特別研究員(DC2)(1998) / ATR人間情報通信研究所 客員研究員(1999) / ケンブリッジ大学 CNBH客員研究員(2000-2001) / 日本学術振興会 特別研究員(PD,北陸先端科学技術大学院大学 情報科学研究科)(1999-2001) / 北陸先端科学技術大学院大学 情報科学研究科助手(2001-2005)



音環境ディバイドレス・パンフレットより抜粋



- キーワード**
- ・聴覚特性
  - ・聴覚モデル
  - ・カクテルパーティー効果
  - ・選択的聴取
  - ・CASA
  - ・音の情報表現
  - ・分離・融合
  - ・変調伝達関数と音声明瞭度
  - ・残響・雑音除去
  - ・残響時間/STI逆推定
  - ・電子音響透かし
  - ・マルチメディア情報ハイディング

歌声合成/報知音知覚/音源分離/認識  
 非言語音声(SCOPE)/聴覚フィードバック

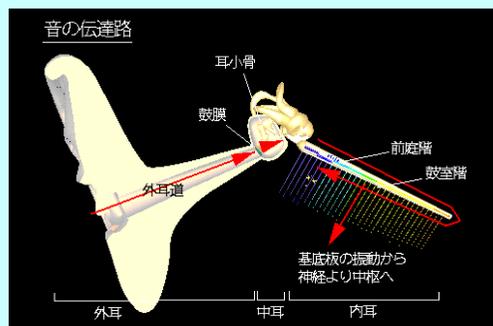
# プロジェクト

## 聴覚的信号処理

聴覚モデル化

音（音声）  
信号処理

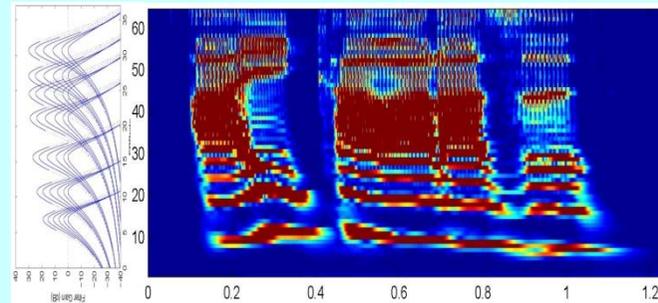
### \* 蝸牛モデル



### \* 周波数選択性



### \* 聴覚フィルタバンク



### \* 骨導音声回復



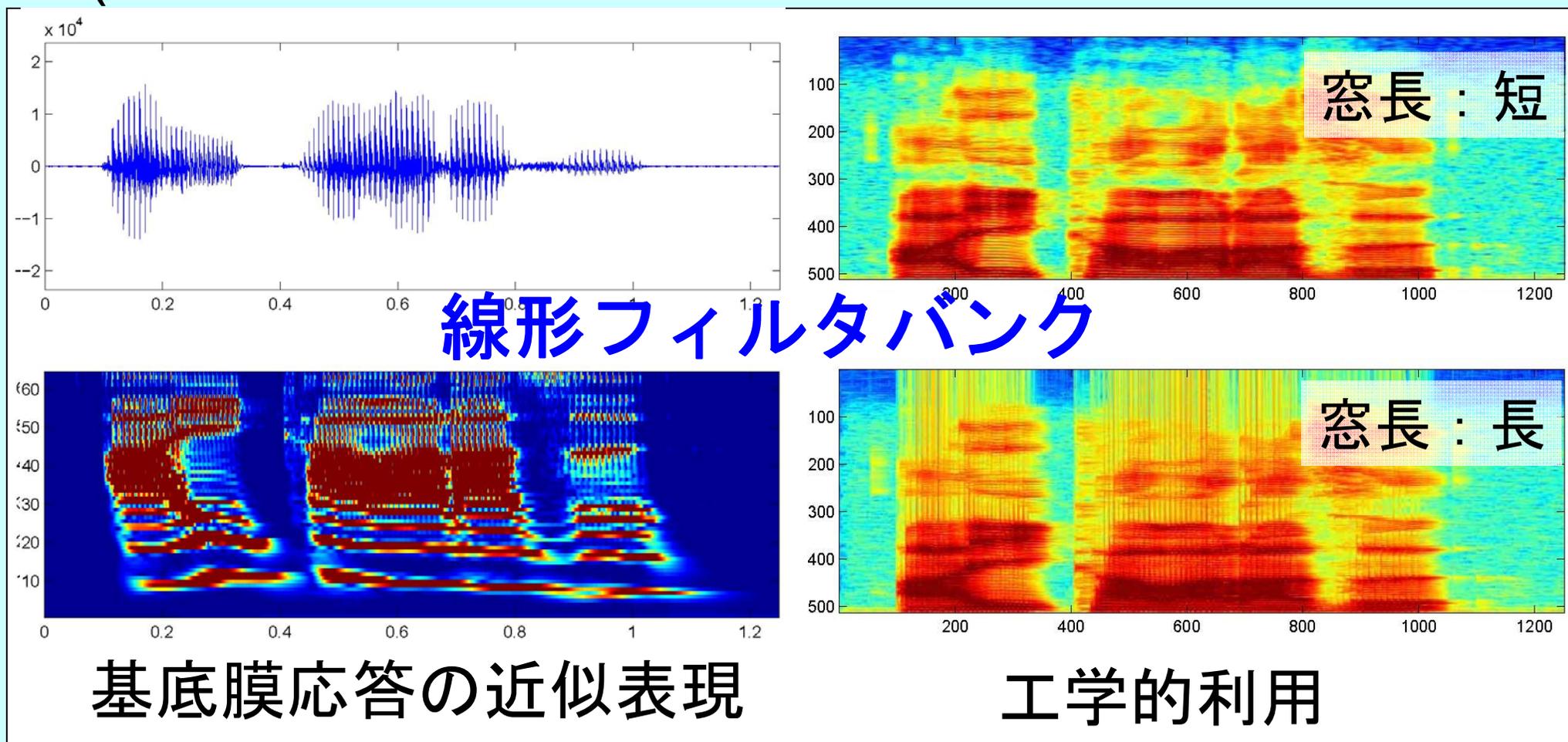
### \* 電子音響透かし



# 聴覚的信号分析技術の確立

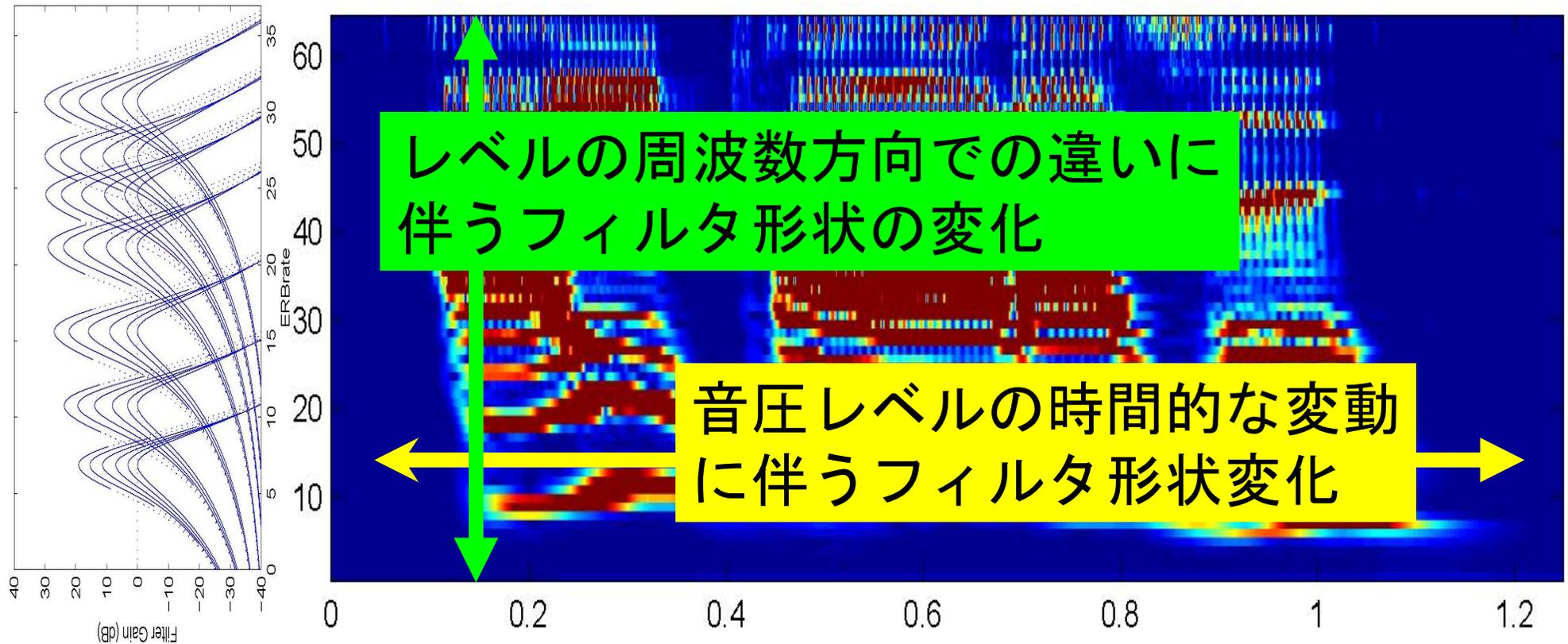
定Qフィルタバンク  
(ウェーブレット変換)

vs. 短時間フーリエ変換  
(窓関数による切取り)



# マスキング特性⇒フィルタバンクの構築

- ある音（マスキュー）が妨害音（マスキュー）によって聞き取り難くなること

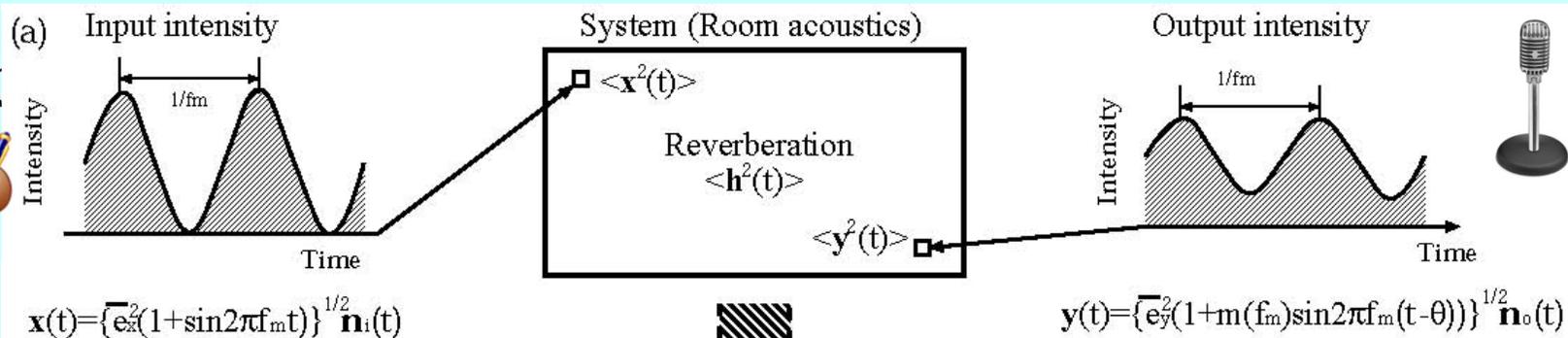


# マスキング実験

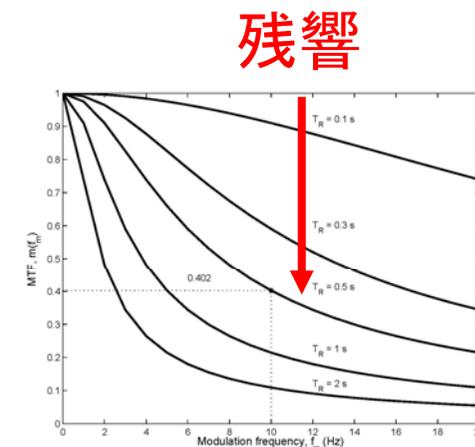
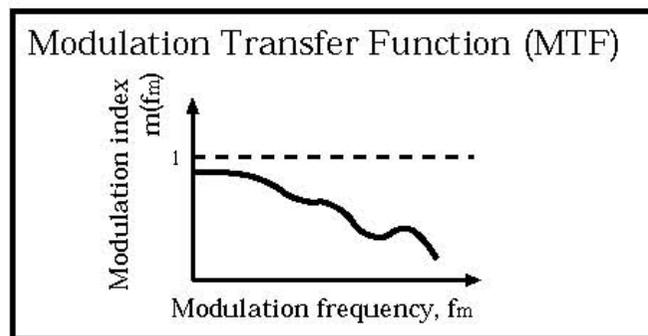
## ■ 実験環境の整備



# 変調伝達関数に基づいた信号処理



(b)

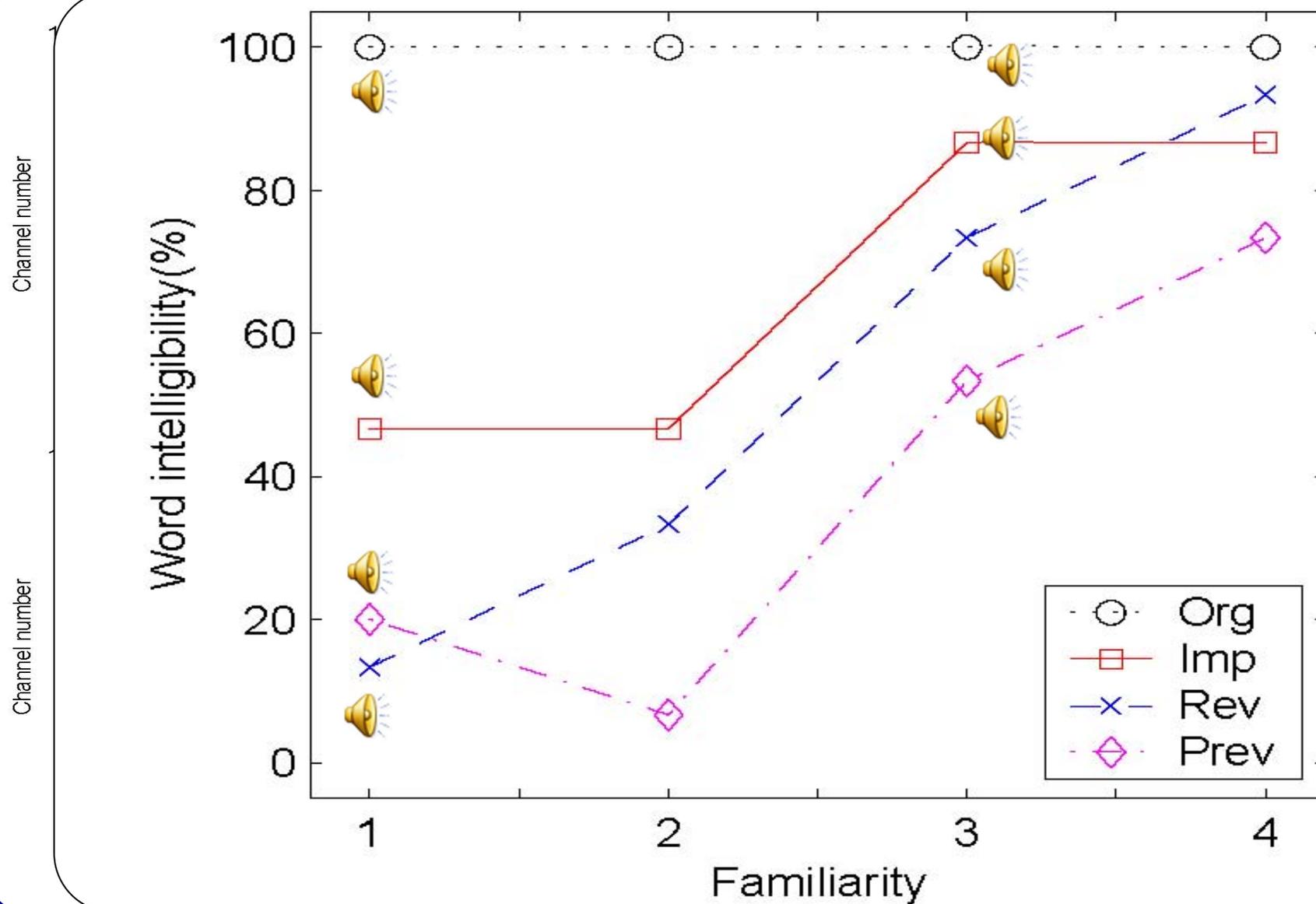


(c)

Speech Transmission Index (STI)

Prediction of speech intelligibility

# 残響音声回復のデモ



# 室内音響特性のブラインド推定



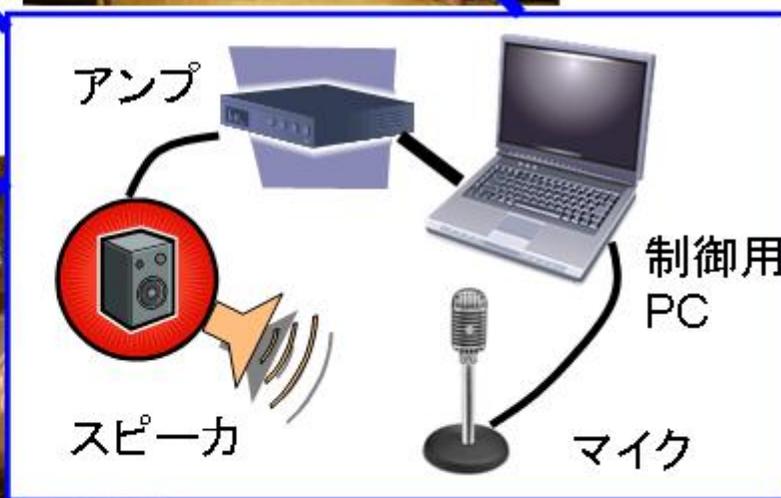
空港内



講義室



駅構内



## 問題点:

人を排除しなければいけない  
装置を持ち込みしなければならない



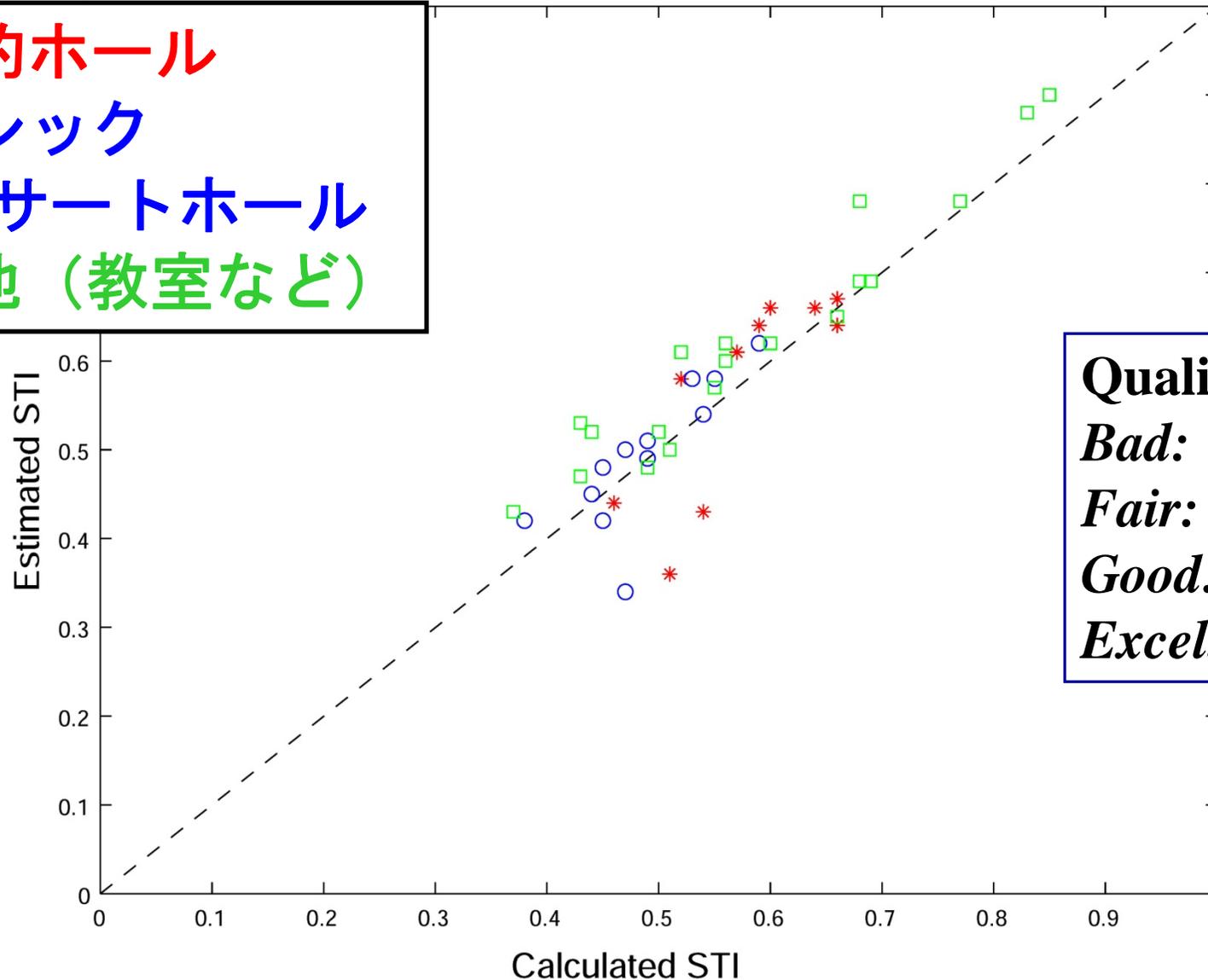
- ・実時間測定
- ・コンパクト
- ・持運び自由
- ・音声を利用



提案法

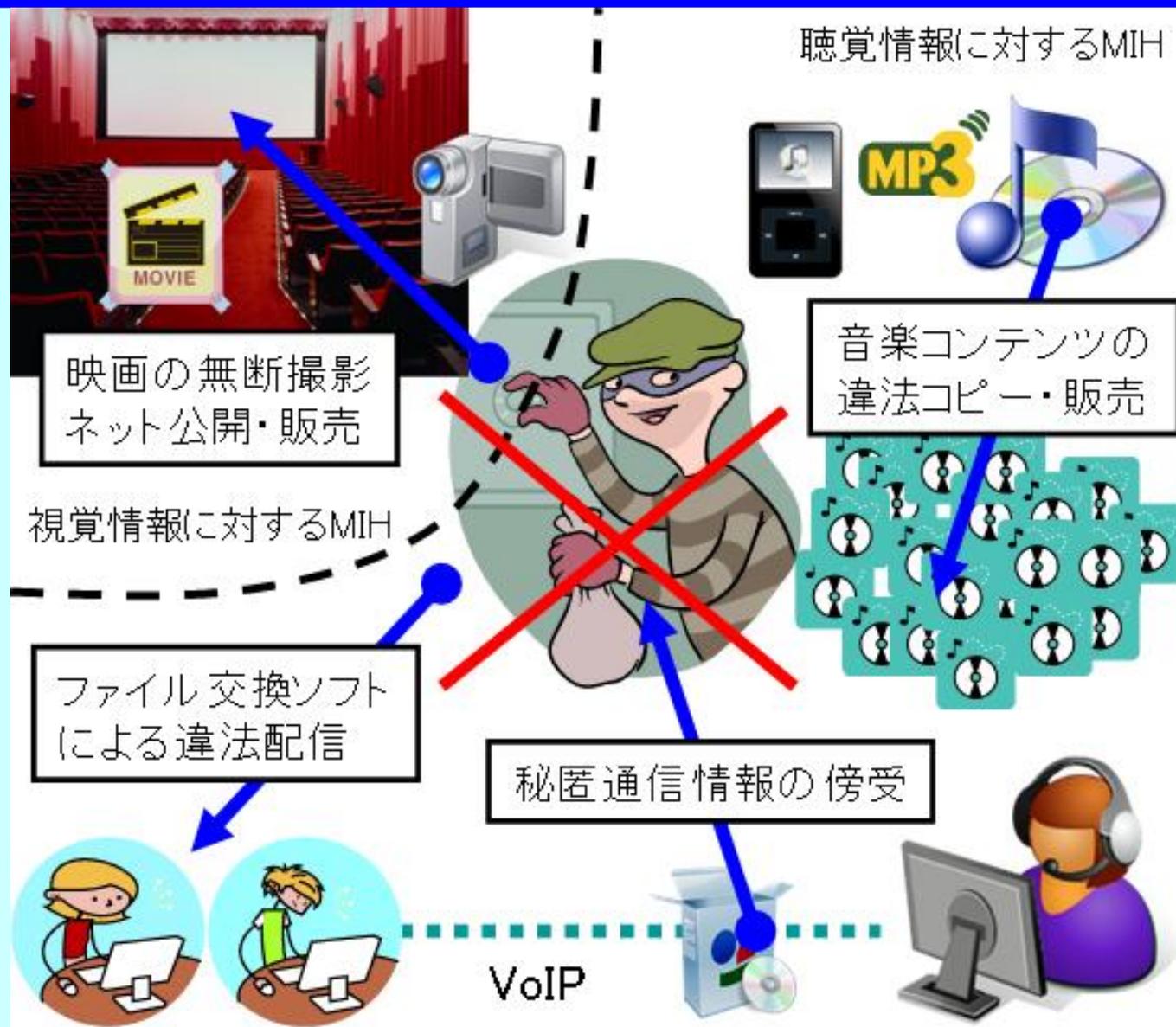
# 結果 (STI)

- ★多目的ホール
- クラシック  
コンサートホール
- その他 (教室など)

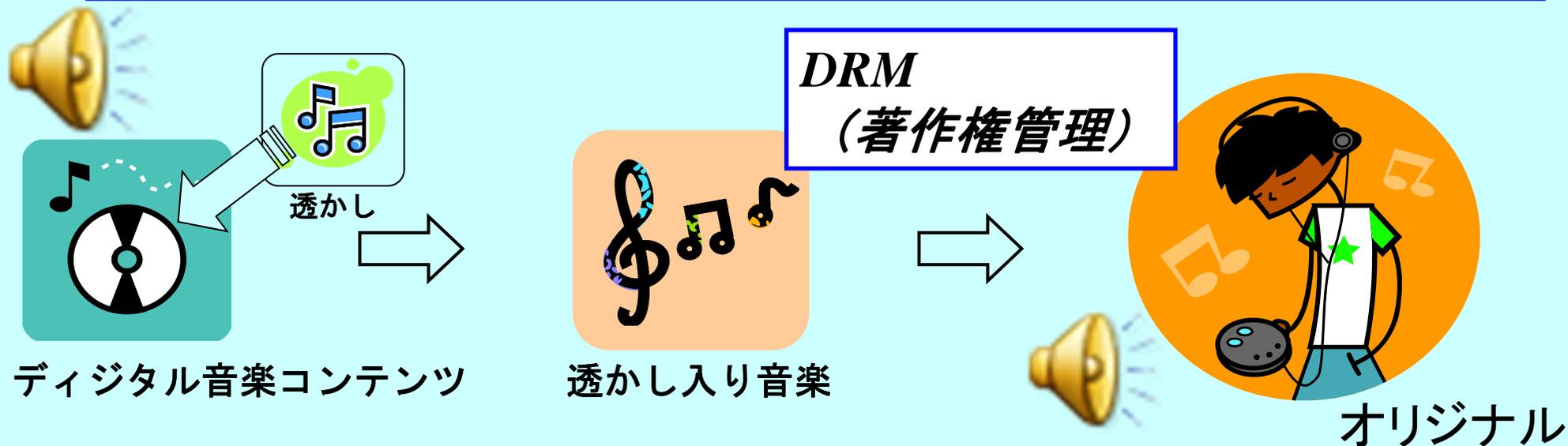


**Quality SIT**  
*Bad:* 0.0-0.45  
*Fair:* 0.45-0.6  
*Good:* 0.6-0.75  
*Excel.:* 0.75-1.0

# マルチメディア情報ハイディング



# 電子音響透かし

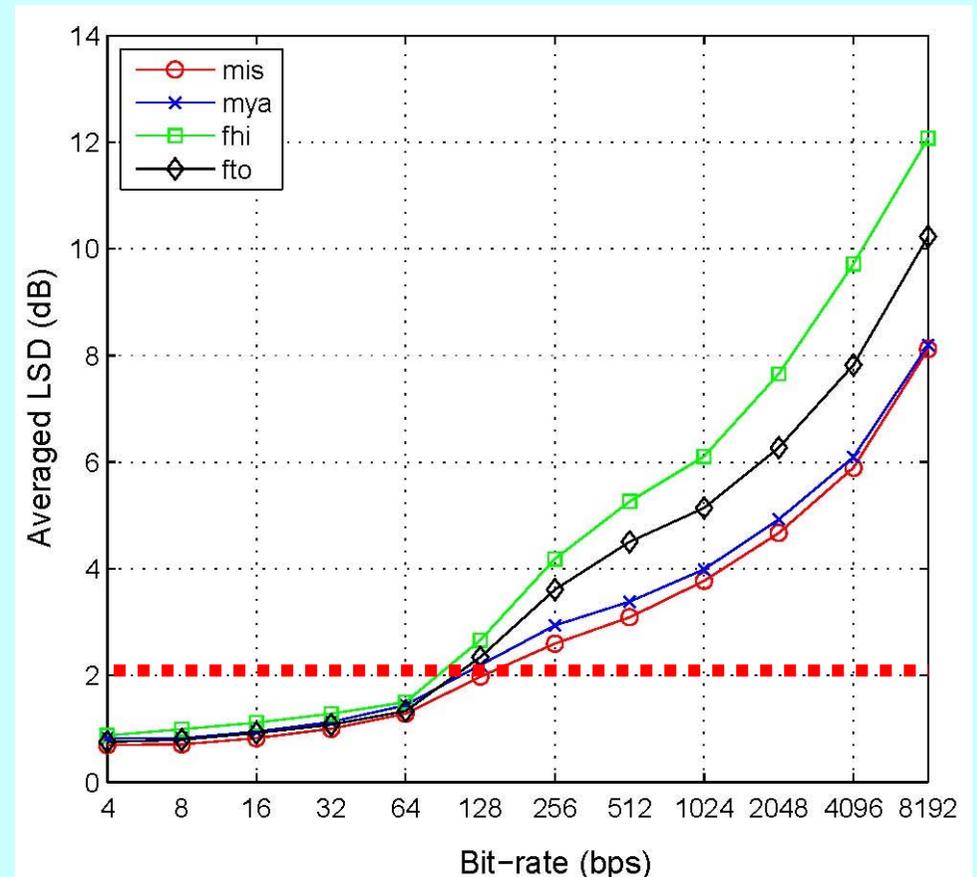
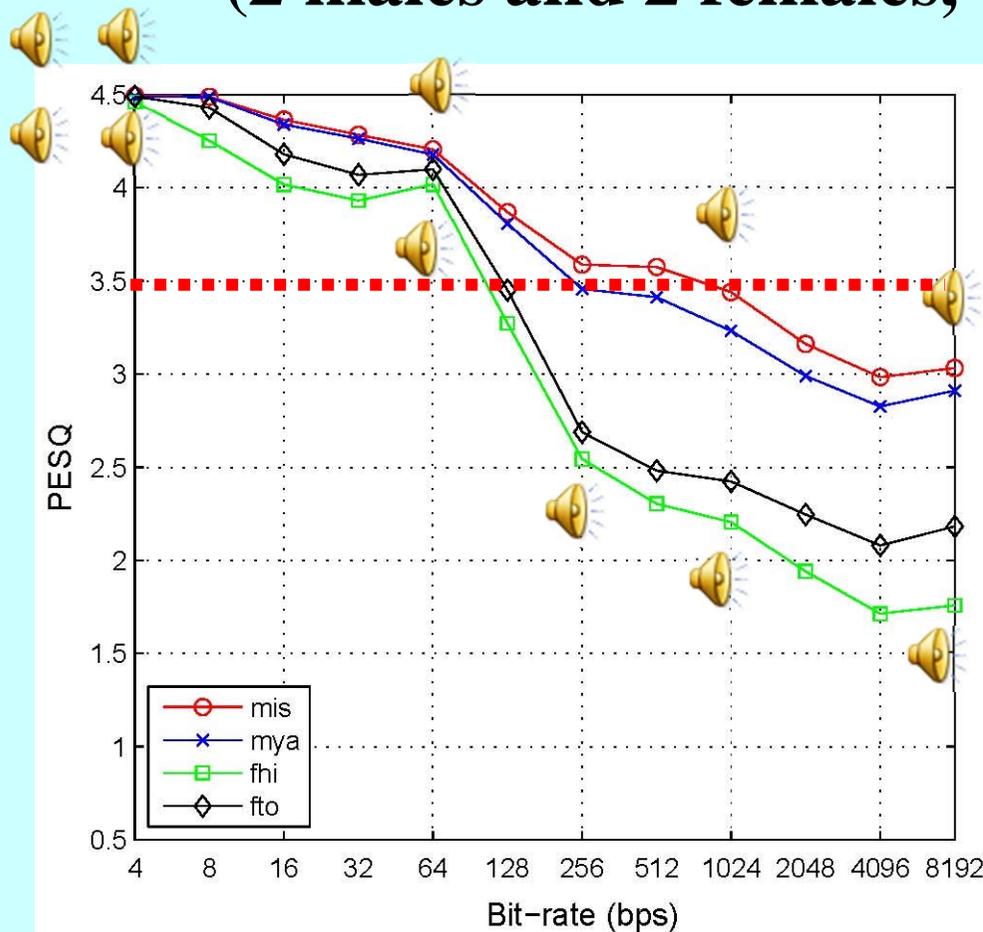


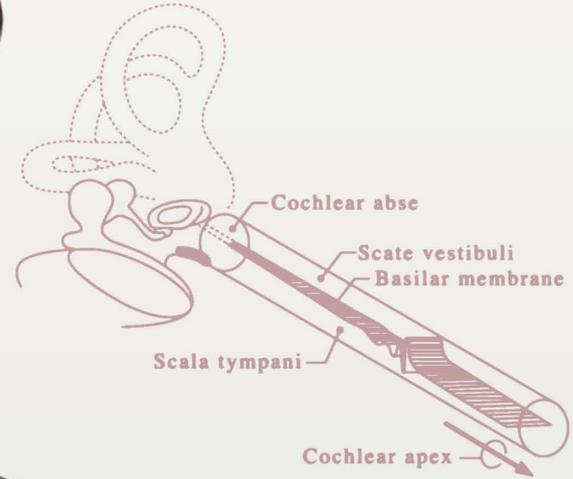
**蝸牛遅延特性 → 知覚不可能**

	Echo hiding	Periodic PM	Proposed
Original			
			

# 音声信号への情報ハイディング

Speech database for intelligibility tests: 320 stimuli  
(2 males and 2 females, 4 familiarities, 20 sets)





なぜ、人間の聴覚は優れているのか？  
その知られざる機能を全て解明し、  
人間の聴覚を超える聴覚システムを作りたい。

# 「耳」付き計算機の登場が、音信号処理に革命を起こす!?

雑音や残響がある環境でも、人間は聞きたい音を聞くことができる。

「もしも計算機上に、聴覚と同じ機能を持つ信号処理システムを

構築できたら、さまざまな音信号処理に応用できるのでは？」

という考えの下で研究を進めているのが、鶴木准教授の研究室だ。

## 耳に相当する機能を 計算機につけてみたい

ICレコーダーでインタビューを録音した音声を聞き直してみると、インタビュー時には気にならなかった雑音や残響で、話し手の声が聞き取りにくいことがある。それなのになぜインタビュー時には話し手の声がクリアに聞こえたのだろうか？ それは、私たち人間の聴覚（正確には、聴覚系+脳）が、雑音や残響を排除し、狙った音だけを聞き取るという、非常に優れた機能を有しているからだ。

けれども、「私たちの聴覚は非常に優れている」ということはわかっているのに、「どのような信号処理によってそれが実現されているのか」ということは、実はいまだに解明されていないことが多々ある。自分たちの体の一部であるにも関わらず、である。

「もしもその信号処理のしくみが全て解明できれば、音声認識や補聴システム等の音声信号処理に応用できると思うんです」と鶴木准教授は語る。そして、自らの研究スタンスを図解すると「**π型**」（図1）になると言う。

「『聞く』『話す』といった音声コミュニケーションは、「ことばの鎖」（図2）の過程で行われています。情報科学研究科の音情報処理分野では、党教授が音声生成を、赤木教授が音声知覚を、徳田准教授が非線形力学に基づいた音声分析を、そして私が聴覚信号処理を専門にし、この4人で音声の生成から知覚情報処理までをすべてフォローしながら連携して研究を行っています。

そして、私の研究室が目指しているのは「耳の機能とまったく同等な信号処理を計算機上に作りたい」——さらにわかりやすく表現すれば「**耳に相当する機能を計算機に付けたい**」ということなんです。

そこに行き着くためには、2本の柱による研究アプローチが必要です。1本目の柱は「人間の耳をまず良く知る」というアプローチです。学問的には、聴覚の各部位の機能解明を目指す『聴覚生理学』や、音の物理的変化に対する反応（知覚変化）を調べる『聴覚心理学』の研究領域です。もう1本の柱は「人間の耳を模擬して計算機上に記述する」というアプローチです。こちらは、学問的には『情報科学』あるいは『工学』の領域です。つまり、**聴覚生理学、聴覚心理学、情報科学の3つがオーバーラップする部分**が、私たちの研究領域です。JAISTの先生方や国内外の研究者に協力を仰ぎ、聴覚心理学・情報科学を本学で実施し、聴覚生理学に関しては海外の研究成果等を参考にしながら追求しています」

## 驚くほど快適な補聴器や 音声透かしの普及も間近

現在、鶴木研究室ではさまざまなプロジェクトが進行している。

そのひとつが、「**聴覚フィルタバンクの構築**」だ。聴覚器官では、蝸牛の基底膜で音を周波数成分に分解し、電気信号（パルス）に変換して神経に伝えている。この機能は、簡単にいうと帯域通過フィルタがたくさん集まった周波数分析器とみることができる。最近では、この分析機能が音の音圧レベルや周波数に応じて能動的に変わることがわかってきている。鶴木研究室では、聞きたい音と妨害音を組み合わせたさまざまなマスキング実験を行い、能動的な周波数分析の能力を調べることで聴覚的信号分析技術（聴覚フィルタバンク）を確立しようとしている。

さらに、この聴覚フィルタバンクを利用した応用として「**聴覚特性に基づいた信号処理**」の研究も進め、残響音声や骨導音声の明瞭度を回復したり、雑音や残響

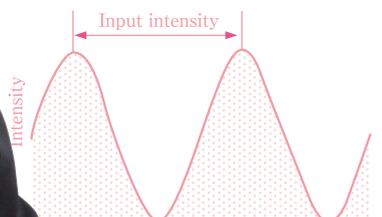
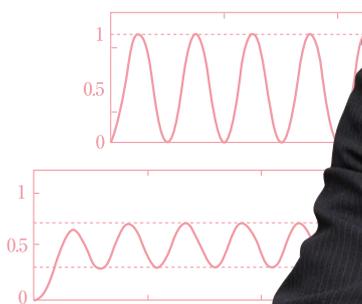
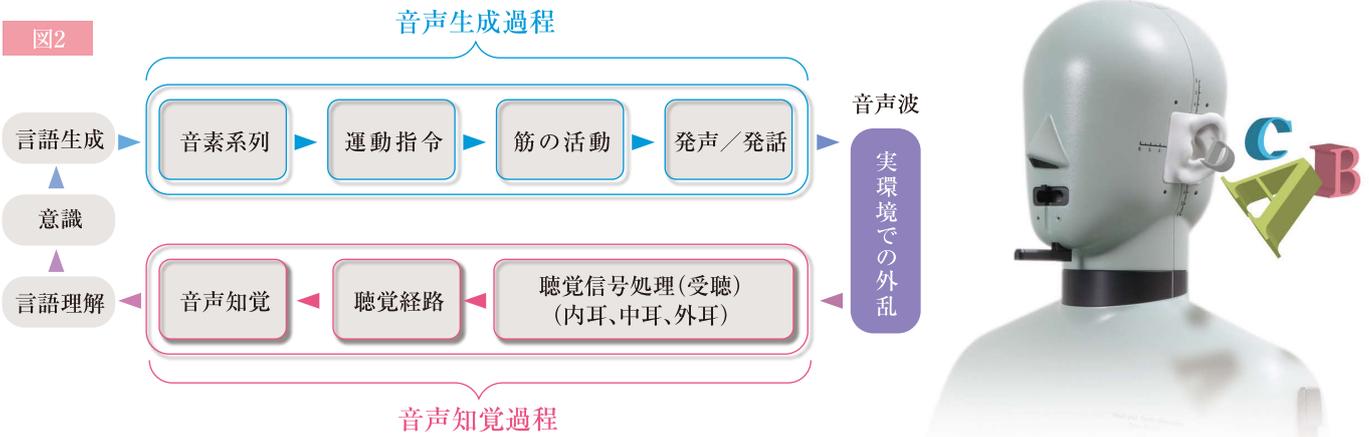


図2



の影響を取り除くといった研究に取り組んでいる。「補聴器をつけると、音全般を増幅するため人の話し声も聞こえるようになるが、周囲の雑音も一緒に増幅されてしまうので、うるさくて仕方がない」という話をよく耳にするが、これらの研究が進めば補聴器の性能は飛躍的に進化するだろう。

また、この信号処理研究の延長線上で、「電子音響透かし」という技術も実現した。蝸牛の基底膜が音を周波数分解する際、低い周波数成分のほうが高い周波数成分よりも少しだけ遅れて渡される。これを「蝸牛遅延」と呼ぶが、この時間差を利用して、音声ファイルを聞いているリスナーにはまったく聞こえないけれども、管理者側ではある条件下でのみ検出できる情報を埋め込むのだ。偽造を防ぐために紙幣に透かしが入っているように、音楽データの違法コピーを発見する場合などに大きな力を発揮する。鶴木准教授いわく「すでに基本特許は出願済みで、音楽機器メーカー等と実用化に向けた話し合いが持たれている段階です」とのこと。この技術が、さまざまなコンテンツの保護に寄与する日も近いだろう。

## 音が好きであることと「考える」意志が重要

耳と同等あるいはそれ以上の機能を作ってみたいという純粋な探究心を持ちつつ、技術の実用化もしっかりと行う。鶴木准教授のエネルギッシュな研究姿勢は、JAISTの修了生であることと大きな関係がありそうだ。

「大学時代はニューラルネットワークを利用した連想記憶の研究をしていました。そのころから、通信ではノイズの影響を受けてしまうような状況でも人間は影響を受けずにちゃんと聞こえるのはなぜだろう、ニューラルネットワークを利用した雑音除去フィルタのような処理はできないものかと興味を持ちました。その後、この分野で研究してみた

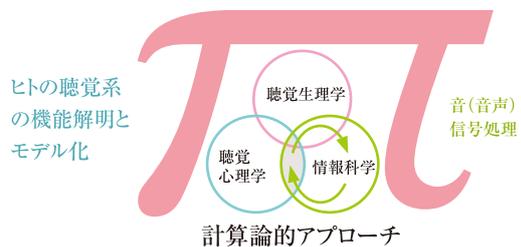
と思った頃に、大学の恩師からJAISTの存在を教えてもらったんです。音声知覚を研究していらっしゃる赤木教授の存在が入学の決め手だったんですが、**フロンティア精神を持って研究をしよう**という建学精神にもすごく惹かれましたね。JAISTが縁でたくさんの素晴らしい人に出会えましたし、本当に良かったと思っています」

そんな鶴木准教授が大事にしているのは、「考える」という言葉。「マスターなら『与えられた課題を自分の力でどう解くか?』、ドクターなら『自分自身で問題を発見し、それをどう解くか?』が求められますよね。**研究者は、常に考えなければいけない**。逆に言えば、学生たちに考える環境を用意することが私たちの仕事です。

この研究室の学生たちに共通しているのは、『**音が好き**』ってことだと思いますね。例えば、楽器が好きで自分でバンドを組んでいて、そういう関連の研究をしたいなあとあってここを選んだっていう学生もいます。でも、大学時代に音声の研究をしていた学生はほとんどいないんですよ。実際、私自身もそれまでは違う分野を学んでいたわけですし。だから、今まで違う研究をしていても、音に興味があって一生懸命考えてみたい、根気強く研究してみたいという意志がある人なら、ここで一緒に学び、苦楽を味わい、切磋琢磨することで、自分自身の可能性を無限に広げることができると思います」

鶴木研究室の「考える」力が、音信号処理の世界を大きく変えようとしている。

図1 研究スタンス(π型アプローチ) 聴覚的信号処理



聴覚フィルタバンクの構築のために簡易型防音室で行う、マスキング実験の様子。



防音室内からの音信号をチェックしたり、音声プログラムを整理する際に使用。



ゼミの風景。「音好き」のメンバーによって、常に活発な意見が交わされる。

# 耳はそもそもなぜ必要なのか？それが知りたくて研究を続けているんだと思う。

情報科学研究科 准教授

**鵜木 祐史** Associate Professor Masashi Unoki

**【趣味】**学生時代はラグビーをやっていました。根気よく研究を続けるには、体力や強い精神力が必要ですが、ラグビーを通じて研究に向かっていく力が身についたと思います。

## 【この研究を続ける理由】

聴覚はなぜ必要なのか？ どのような機能を持っているのか？ 科学を志している人間として、その一端を垣間見たいという純粋な理由で研究を続けているんだと思います。でも、生きている人間の耳を解剖するなんてことは不可能ですよ。そんな中で人間の耳の機能を解明するためには「人間の耳の各部位の機能は何か？」「人間の耳と等価なものを作るにはどうすればいいか？」ということを丹念に調べて、データを積み重ねていくしかない。遠回りに見えるけれど、実はそれが本質的でいちばんの近道なのかもしれないと思うんです。

学位	北陸先端科学技術大学院大学 情報科学修士(1996) / 北陸先端科学技術大学院大学 情報科学博士(1999)
経歴	日本学術振興会 特別研究員(DC2)(1998) / ATR人間情報通信研究所 客員研究員(1999) / ケンブリッジ大学 CNBH客員研究員(2000-2001) / 日本学術振興会 特別研究員(PD, 北陸先端科学技術大学院大学 情報科学研究科)(1999-2001) / 北陸先端科学技術大学院大学 情報科学研究科助手(2001-2005)



エクセレントコア形成支援プロジェクト研究テーマ

## 音環境ディバイドレス化

残音、残響などの音環境ディバイド(異なる音環境間の音声コミュニケーションで生じる、分断された障壁・格差)がどれだけ音声コミュニケーションに悪影響を与えるのか、また音環境ディバイドをどのように取り除けばよいのかを明らかにする。

雑音や残響の種類・特性、音源配置などをパラメータとする室内音響の特徴づけ

音声の音響特徴ならびに音質や音声明瞭度の劣化度具合を調査する

MTF / STIを通じて雑音・残響を同時に扱えるのか議論し、音環境ディバイドレス化の方略を提案

### 研究の成果

私たちの音声コミュニケーションの中で悪になる要因を排除する方法を探っている。すでに、アプローチはMTF / STIを通じて行う方向で固まり、そこから生じる諸問題を浮きあがらせ、その1つないし2つの解決に取り組んでいる。本研究により、健康者・聴覚障害者の区別なく、あるいは残音・残響などの影響を受けず、いつどこで誰とでも普通の会話ができる、安心して安全な「音のバリアフリー社会」を目指す。

鵜木研究室

情報科学研究科棟 1棟 9F

TEL 0761-51-1237 FAX 0761-51-1149

E-mail: unoki@jaist.ac.jp http://www.jaist.ac.jp/~unoki/



国立大学法人 北陸先端科学技術大学院大学

〒923-1292  
石川県能美市旭台1丁目1番地  
http://www.jaist.ac.jp

お問合せ

広報係

TEL 0761-51-1031 FAX 0761-51-1025 E-mail:kouhou@jaist.ac.jp



鵜木祐史 准教授

音処理情報の中でも特に聴覚の優れた能力に着目し、研究をしています。その研究過程で、電子音響透かしの技術や、新しい残響時間推定法を開発しています。

使用装置

心理物理実験装置 (TDT System IIIシリーズ・各ユニット 2セット) / 音環境測定装置 (B&K人工耳、騒音計、Head Torso等) / 高速計算機サーバー (IBM System x3500 2台、Huron HURAR) など

キーワード

聴覚特性 / 聴覚モデル / カクテルパーティー効果 / 選択的聴取 / CASA / 音の情報表現 / 分離・融合 / 変調伝達関数と音声明瞭度 / 残響・雑音除去 / 電子音響透かし

鵜木研究室

TEL 0761-51-1237

FAX 0761-51-1149

E-mail: unoki@jaist.ac.jp  
http://www.jaist.ac.jp/~unoki/

著作権保護情報を埋め込む電子音響透かし、あるいは人ごみの中でも測定可能な残響時間推定法の技術を、さまざまな形で有効活用してもらえればと考えています。

応用の可能性

デジタル音響機器・インターネット  
音響計測分野・建築音響分野 ほか

- 電子音響透かし…オーディオコーデックを利用しているデジタル情報機器・音響機器の市場すべて
- 残響時間推定法…良好な音響設計が要求される施設(コンサートホール、シアター、AVルームなど)、業務用音響設備、音環境を検討する必要がある施設

研究室の概要

— 聴覚の優れた特性に着目し、  
聴覚的な音信号処理の実現を目指す

私たち人間は、雑音や残響がある環境においても、狙った音をいとも簡単に聞き取ることができます。しかし、同じことを計算機で実現することは非常に難しい問題です。もしも計算機上に、聴覚と機能的に等価な信号処理システムを構築することができれば、音声認識や補聴システムといったさまざまな音信号処理に応用することができるはずです。鵜木研究室では、**聴覚の優れた能力に着目し、聴覚的な音信号処理の実現**を目指しています。具体的には、以下のようなプロジェクトを進行させています。

聴覚フィルタバンクの構築	聴覚情景解析の計算モデル化	聴覚特性に基づいた信号処理
聴覚の優れた周波数選択性の機能を解明し、その実験結果に基づいて、聴覚による信号分析と機能的に等価な聴覚フィルタバンクの構築を目指します。	音韻修復現象や選択的音抽出などの研究により、聴覚情景解析の計算をモデル化します。	選択的な信号分離法や雑音除去法、変調伝達関数に基づいた残響音声回復法、骨導音声の明瞭度回復の研究を行っています。これからご紹介させていただく2つのテーマは、聴覚特性に基づいた信号処理に関連して開発した技術です。

研究テーマ

Theme 1. デジタル音響信号へ知覚不可能な情報を埋め込む「電子音響透かし」

デジタル音楽コンテンツの世界では、違法コピー、違法配信が横行しています。これらの著作権を保護する目的で開発したのが「**電子音響透かし**」の技術です。

蝸牛の基底膜が音を周波数分解する際、低い周波数成分のほうが高い周波数成分よりも少しだけ遅れて伝播される「**蝸牛遅延**」という特性を埋め込みの基礎技術として利用します。そして、音声ファイルを聞いているリスナーにはまったく聞こえないけれども、管理者側ではある条件下でのみ検出できる情報を埋め込みます。

電子音響透かしに求められる要件は、「**知覚不可能性**」(透かし情報が埋め込まれたことに気づかないこと)、「**秘匿性**」(許可された者のみ透かし情報を検出できること)、「**頑健性**」(信号変形処理を施した後でも検出できること)の3つであると思います。スペクトル拡散法、エコーハイディング法、ビット置換(LSB)法、周期的位相変調法などの従来技術が存在していますが、これら3つをすべて満たすものではなく、特に知覚不可能性に改善の余地があります。私たちが新しく開発した「**蝸牛遅延に基づいた電子音響透かし**」の技術は、秘匿性、頑健性に優れています。特に知覚不可能性においては、従来技術と比べて非常に優れたものとなっています。

想定される用途として考えられるのは、1つ目はデジタル音響データの著作権保護（違法コピーへの対応、違法配信への対応）です。2つ目はデジタル音響データへの付加価値情報の追加（ステガノグラフィへの発展）であり、音情報へ文字情報等を埋め込むことにより、同一チャンネル内の複数情報の提供や、携帯情報端末での2次情報伝達（宣伝など）が可能になると思います。

特許はすでに申請済みです。デジタル携帯端末・デジタルオーディオプレーヤー等の開発企業の皆さん、デジタルデータに対する暗号化技術を持つ企業の皆さん、あるいはオーディオコーデック技術を持つ企業の皆さんとの共同研究を希望します。提案技術の試作のサポートも行いますので、ご相談ください。

## Theme 2. 室の伝達特性の測定を必要としない残響時間推定法の開発

スピーカー、マイク、アンプなどを必要とする従来技術の室内残響インパルス応答測定システムには、大きく3つの問題点がありました。1つめは、暗騒音によるダイナミックレンジの現象。2つめは、室内の形状変化や物体の移動などに伴う特性変化。3つめは、人を排除する必要があること。すなわち、実測するための条件整備が必要なため、駅など多くの人間が行き来する場所での実測が思うように行えませんでした。

この問題点を解消し、条件を選ばずに実測可能となる技術が、**ブラインド残響時間測定装置による測定法**です。変調伝達関数(MTF)や音声伝達指数(STI)を通じて、室のインパルス応答を測定せずに観測した残響音声から残響時間を推定します。

身の回りの音(残響音声)を利用して残響時間を推定するため、従来技術と比較して以下のようなメリットが挙げられます。1つめは、装置は集音マイクとレコーダーのみでOKなため、**コンパクトで持ち運び自由**であること。2つめは、簡単なデジタルフィルタ技術を利用するため、**実時間測定が可能**であること。3つめは、人を排除する必要が無いので、**実際に人がいる環境でも残響時間の推定に利用可能**なこと。これらのメリットから想定される用途としては、コンサートホールやホームシアターなどの室内音響学に基づく建築・音響設計の分野や、駅構内やホールなどの人を排除できない音環境での特性測定です。その他にも、装置が簡便化したことでさまざまな用途が考えられると思います。

この技術もすでに特許申請済みです。音響計測分野や建築音響分野の方々と、有意義なコラボレーションができればと考えています。

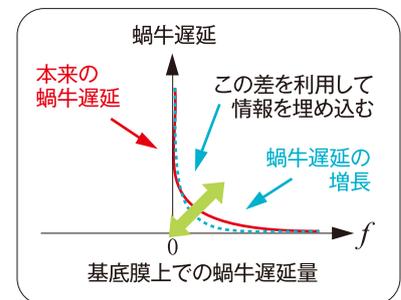
### 従来法の評価と新技術の狙い

手法	知覚不可能性	秘匿性	頑健性
スペクトル拡散法 (Cvejic & Seppän, 2001)	×	○	○
エコーハイディング法 (Gruhl, 1996)	○	×	○
ビット置換法 (Boneyら, 1996)	○	○	×
周期的位相変調法 (西村&鈴木, 2001)	△	○	△
<b>蝸牛遅延に基づいた電子音響透かし</b>	◎	○	○

### 着眼点

信号の蝸牛遅延量を増長してもわからない

{0.1}の情報埋め込みに使える



ベンチャーマインドあふれる企業の皆さんと二人三脚で、電子音響透かしや新しい残響時間測定法の普及に取り組みたいと思っています。

鵜木祐史准教授の「エクセレントコア形成支援プロジェクト研究内容」

残音、残響などの音環境ディバイドがどれだけ音声コミュニケーションに悪影響を与えるのか、また音環境ディバイドをどのように取り除けばよいのかを明らかにし、いつでもどこでも普通の会話ができる、安心して安全な「音のバリアフリー社会」を目指します。



国立大学法人 北陸先端科学技術大学院大学

〒923-1292  
石川県能美市旭台1丁目1番地  
<http://www.jaist.ac.jp>

お問合せ

広報係

TEL 0761-51-1031

FAX 0761-51-1025

E-mail: [kouhou@jaist.ac.jp](mailto:kouhou@jaist.ac.jp)

# 研究紹介 一口笛と骨伝導

森 幹男（福井大）

2012. 3. 11

日本音響学会北陸支部第4回音響・超音波ジョイント研究会

# 研究紹介 —口笛と骨伝導—

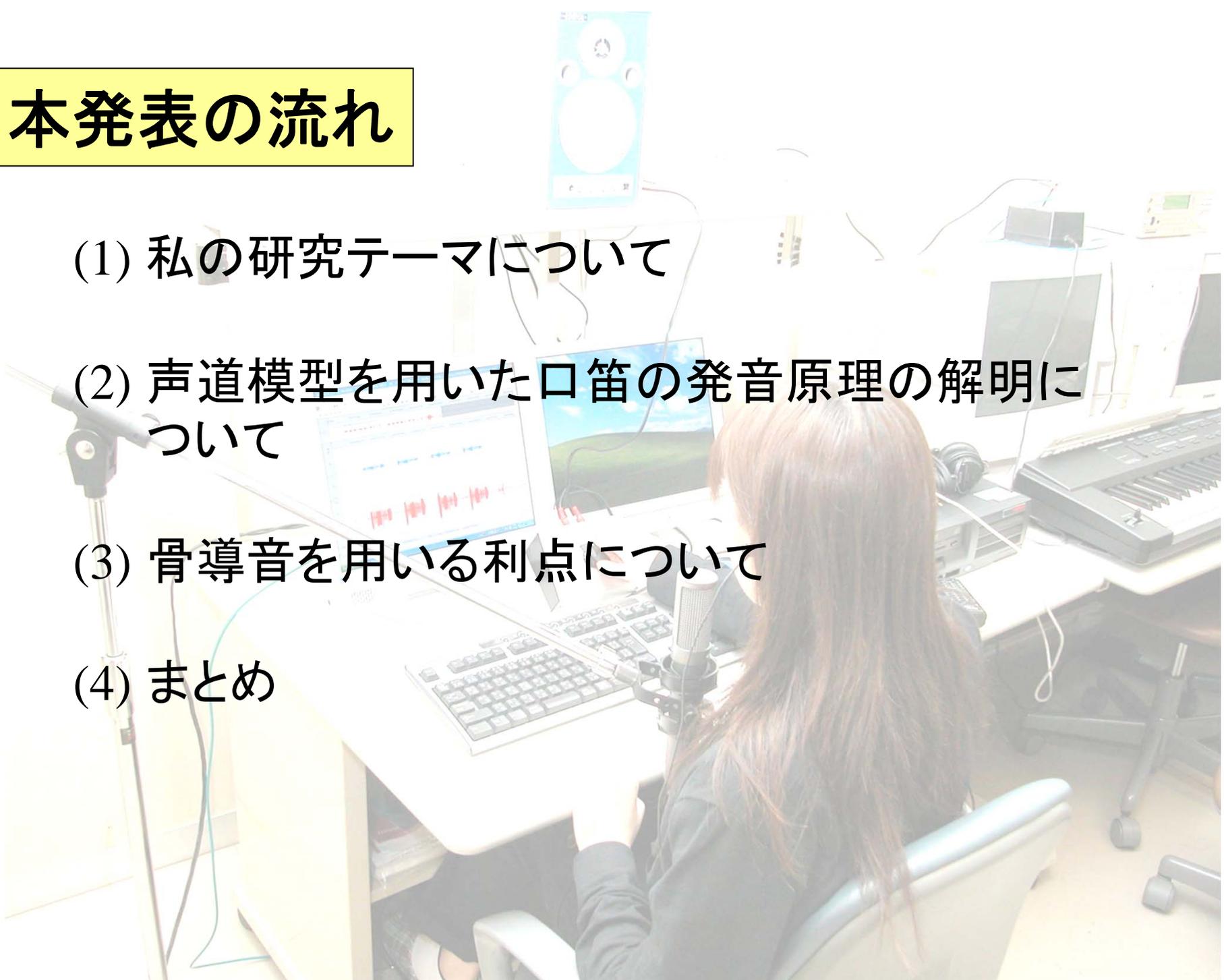
福井大学創立五十周年記念館

福井大学アカデミーホール

Fukui-University Academy Hall

福井大学大学院 工学研究科  
森 幹男

# 本発表の流れ

- (1) 私の研究テーマについて
  - (2) 声道模型を用いた口笛の発音原理の解明について
  - (3) 骨導音を用いる利点について
  - (4) まとめ
- 
- A person with long brown hair is sitting at a desk in a laboratory or office. They are facing a computer monitor displaying a waveform. A microphone is positioned in front of them. To the right, there is a keyboard and a mouse. In the background, there are various pieces of equipment, including a large blue speaker and a small electronic device on the desk. The scene is brightly lit.

# 研究テーマ

## ユビキタス社会に向けた骨導音の応用

- 骨導ヘッドホンの高音質化
- 気導音と骨導音の時間分解能の比較
- 歯を介した骨導音に関する研究
- 骨導口笛音を応用した製品の開発
- 声道模型を用いた口笛の発音原理の解明



骨伝導イヤホンマイク  
+HMD+無線通信

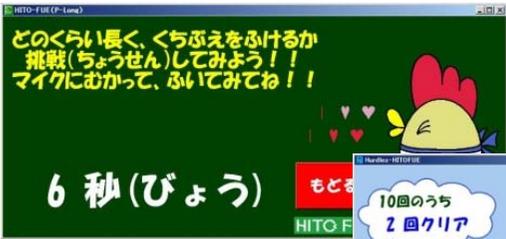
当初考えていたユビキタスHIのイメージ

# 製品像



＜共同研究・受託研究＞  
骨導ヘッドホン，補聴器  
(ハードウェア，信号処理)

＜2008年10月大学発ベンチャー起業＞



## 口笛トレーニング ゲーム



## 口笛検定 (ソフトウェア)



## 口笛練習用装置 (ハードウェア)

# 口笛音楽コンクール

＜日本オープン(日本口笛音楽協会主催)＞

2005年10月10日	第1回予選	2006年 3月12日	本選
2007年11月10日	第2回予選	2007年 11月10日	本選
2008年11月 8日	第3回予選	2009年 3月28日	本選
2009年11月15日	第4回予選	2010年 3月21日	本選
2011年 8月28日	第5回予選	2011年 11月26日	本選

＜国際口笛大会(米国1974年～)＞

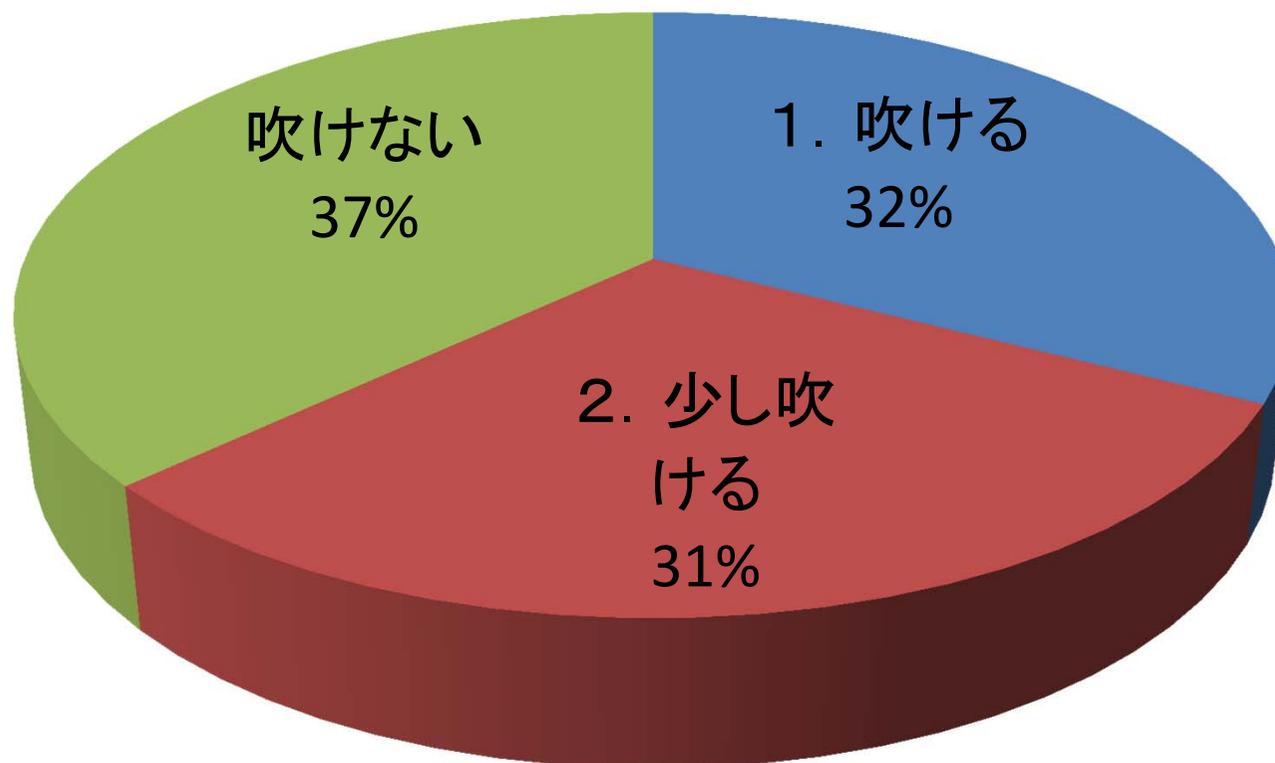
2007年 4月19日～21日	第34回 国際口笛大会(米国)
2008年 7月19日～21日	第35回 国際口笛大会(日本 牛久)
2009年 4月22日～25日	第36回 国際口笛大会(米国)
2010年 5月26日～30日	第37回 国際口笛大会(中国 青島)
2011年 4月 6日～10日	第38回 国際口笛大会(米国)
2012年 4月18日～22日	第39回 国際口笛大会(米国)

\* 口笛の経済効果を考えると・・・

日本の人口1.2億人 × 33%(口笛を吹ける) × 10% × 単価3千円 = 約120億円  
世界の人口70億人 × 1% × 単価3千円 = 約2000億円

# 口笛に関するアンケート結果

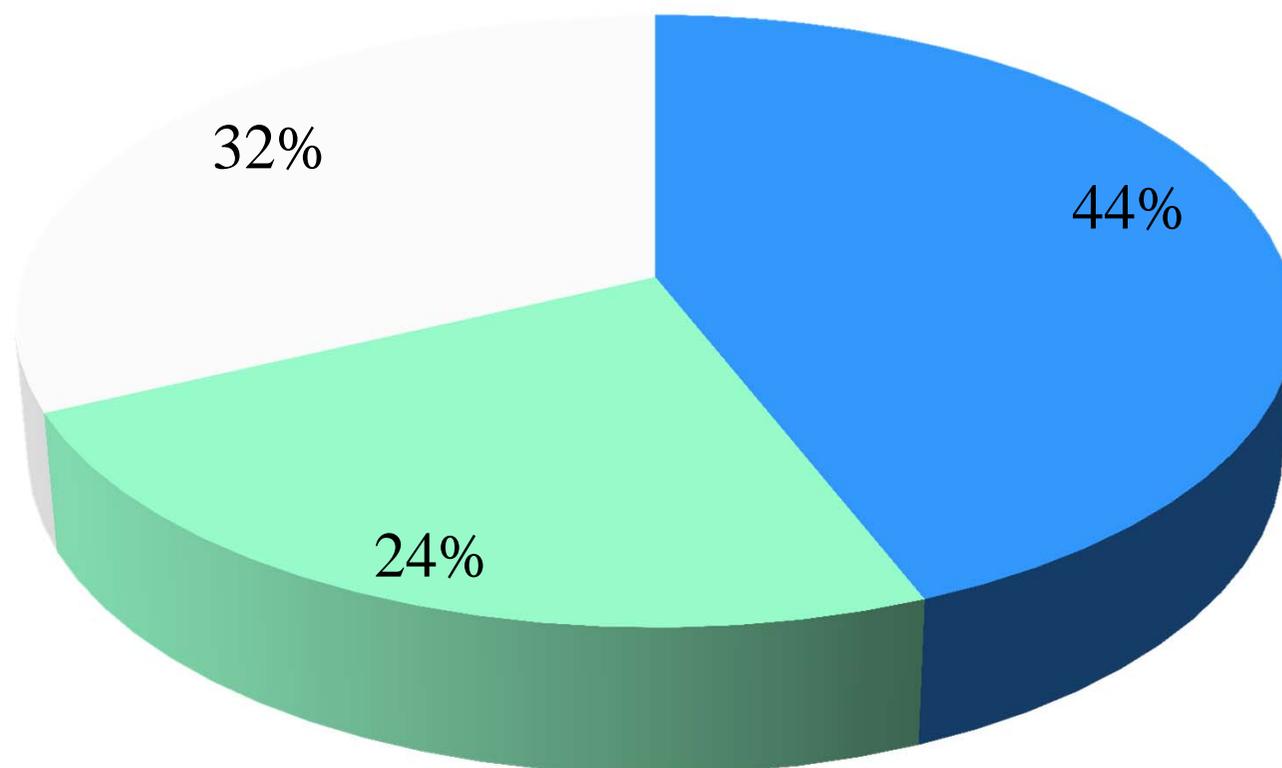
「あなたは口笛で曲を吹くことができますか？」



大学生206名に対するアンケート結果

# 音域測定結果の分布

■ 1オクターブ以上   ■ 1オクターブ未満   3. 吹けない

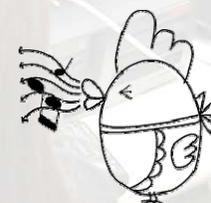


初級では1オクターブ音域を長音階で吹けることを目標とする。

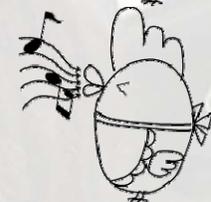
# 口笛検定試験システム

## 口笛音楽検定試験システムのグレード(級)

- 5級 単一の音を正確なピッチで100回発音できる。  
(合格ライン:80%)
- 4級 1オクターブの音域を上昇・下降(長音階)で  
それぞれ発音できる。(合格ライン:87.5%)
- 3級 5級の内容を吹音・吸音で行うことができる。  
(合格ライン:90%)
- 2級 4級の内容を吹音・吸音で行うことができる。  
(合格ライン:93.75%)
- 1級 課題曲を正確に演奏することができる。  
<現在, 研究・開発中。>

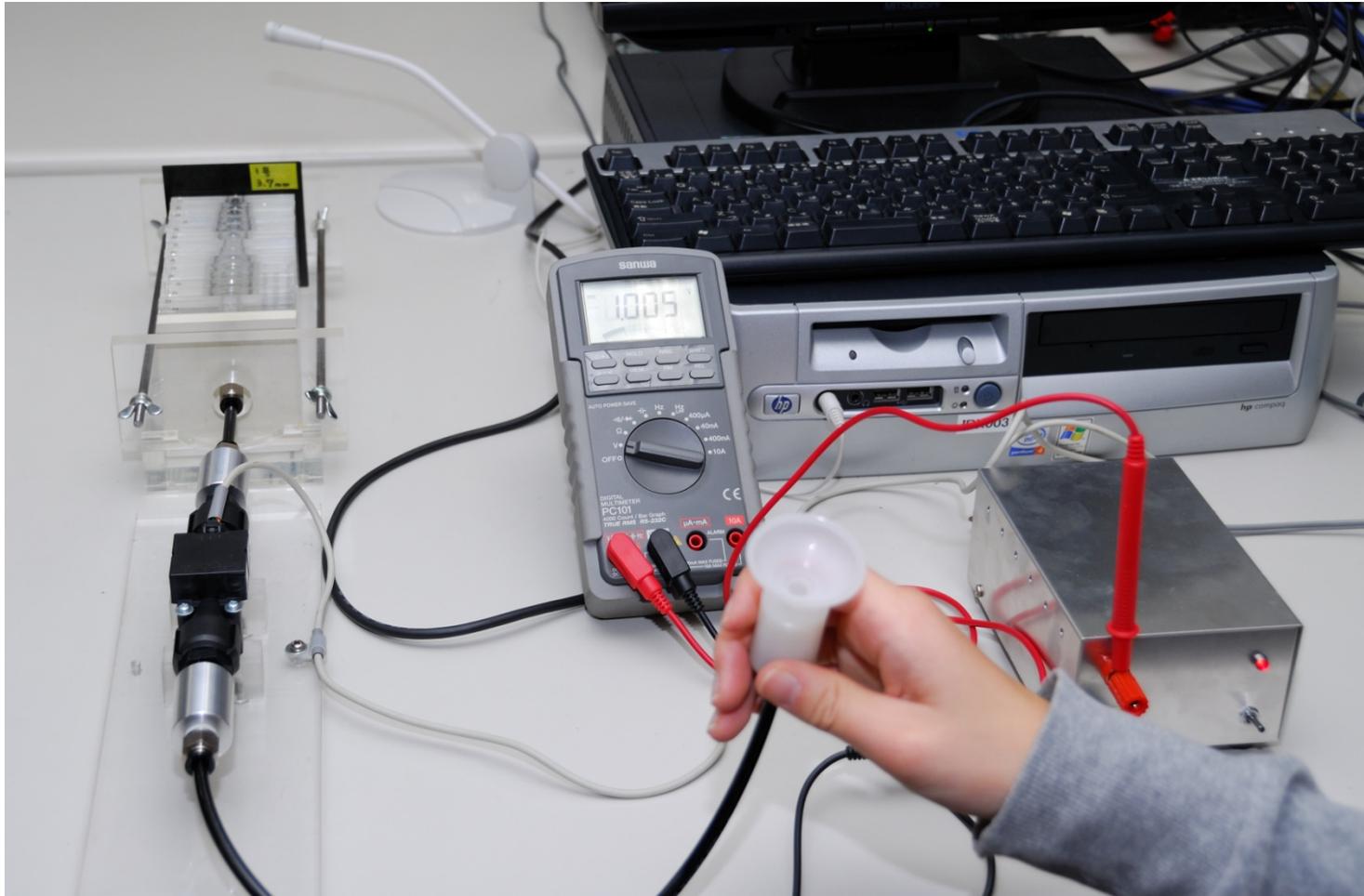


「吹音」  
吹いて発音



「吸音」  
吸って発音

# 口笛演奏時の声道模型



口笛演奏時の声道模型と流量計測・周波数測定実験系

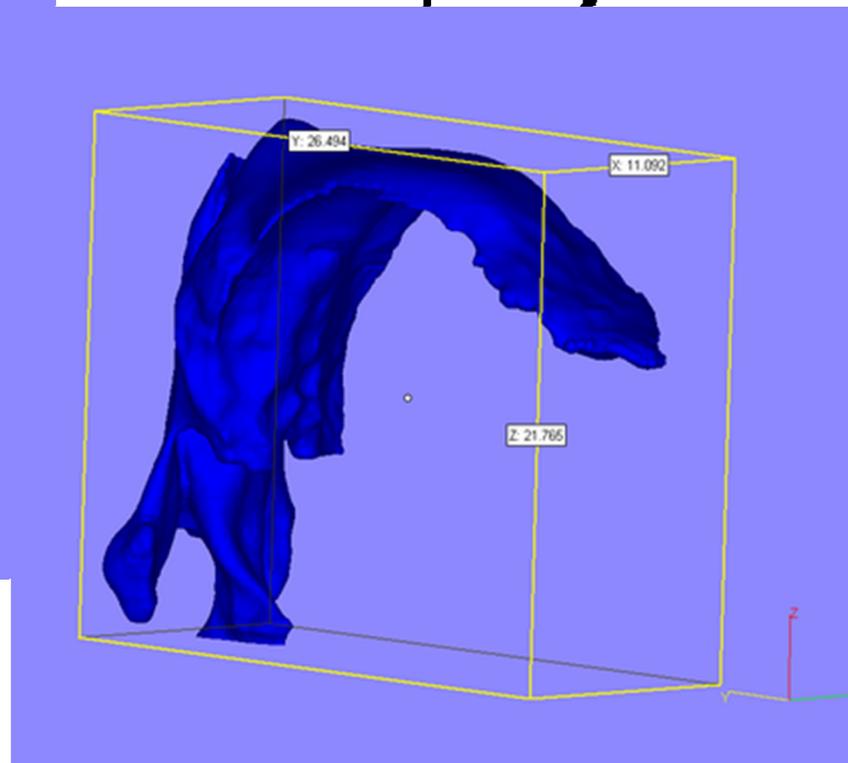
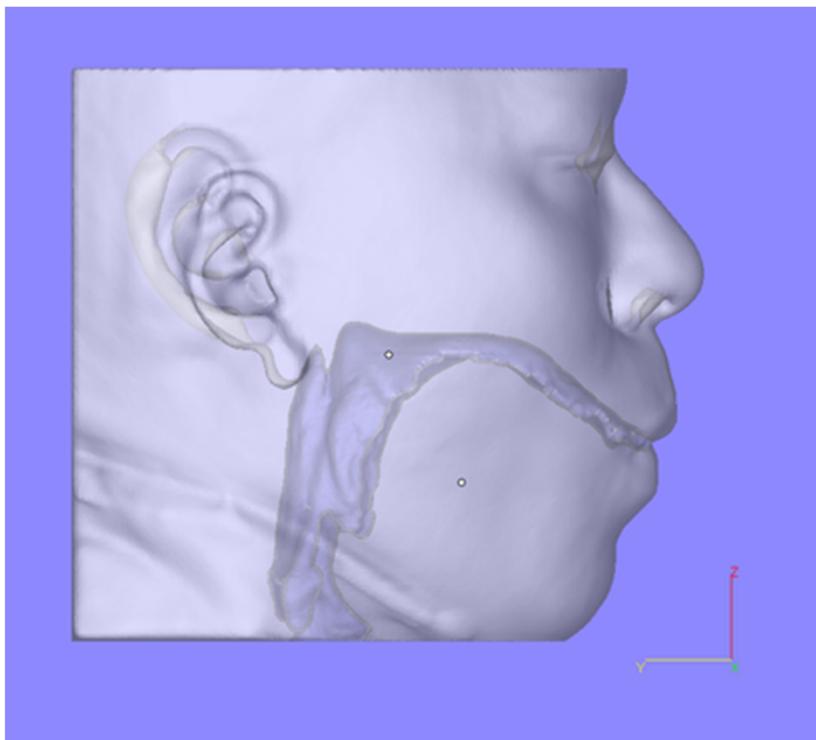
# 声道模型

- 声道模型教材
  - 人間の発声器官を機械的に模擬
  - プレートの組み合わせで任意の声道形状を模擬できる
  - 日本語5母音について音声合成を行う



声道模型教材 VTM-10

# 口笛演奏時の声道



本当にヘルムホルツ共鳴？

# 実験

## ➡ 1. 声道模型を用いたピッチ周波数測定

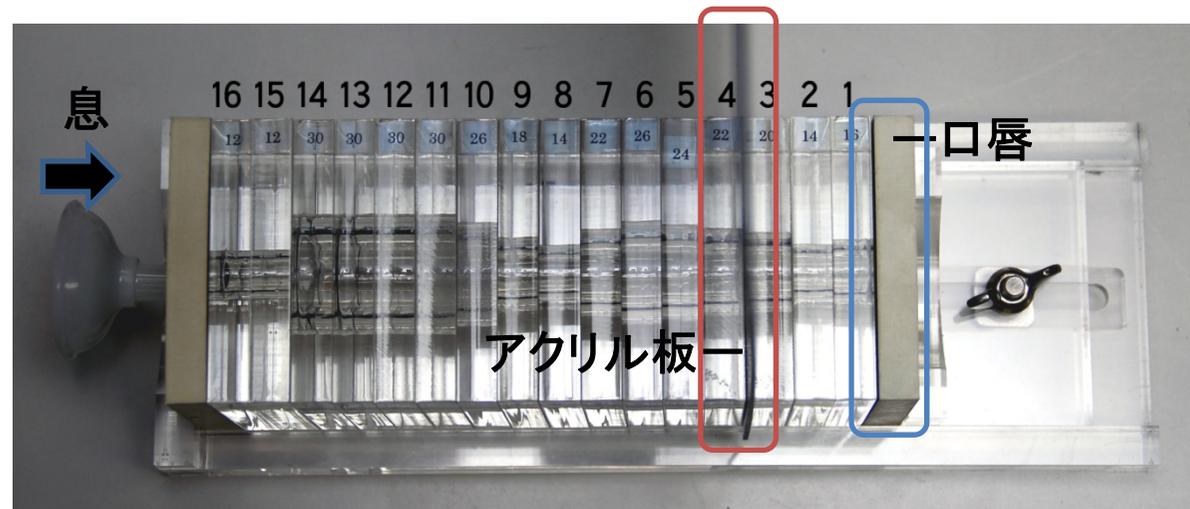
声道模型を利用して口笛演奏時の声道形状を模擬し、ピッチ周波数の測定を行う

## 2. 呼気の流速とピッチ周波数の関係

## 3. 単純化モデルの場合のピッチ周波数測定

# 口笛演奏時の声道模型

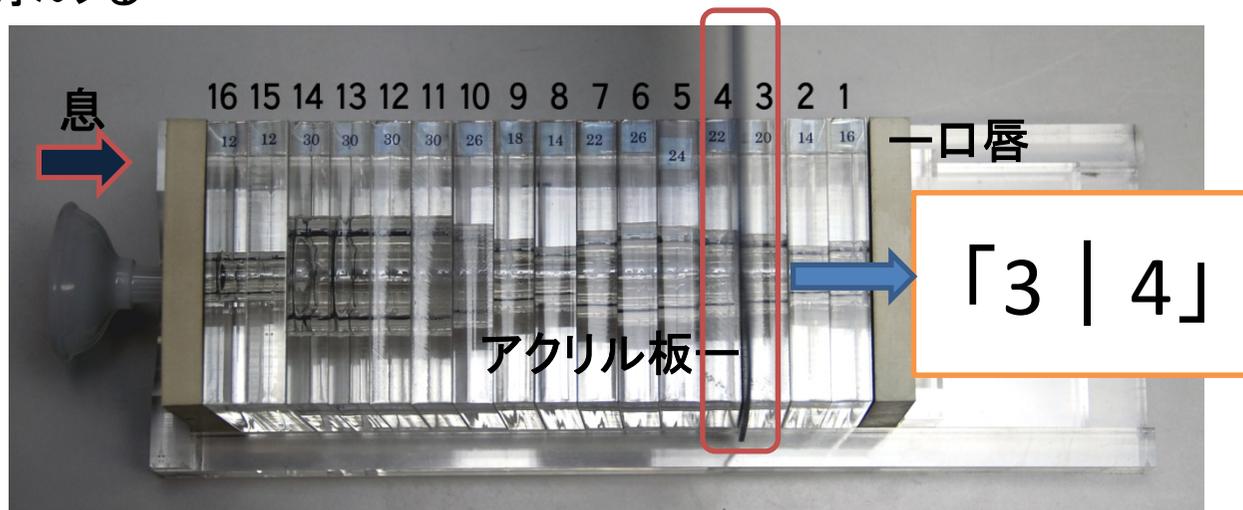
- /u/の形状にセット
- 声帯にあたるパーツを取り除く
- 穴の開いたアクリル板→舌と口蓋による狭め
  - ・厚さ1.5mm
  - ・穴の直径3.7mm
- 穴の開いたゴム板→すぼめた口唇
  - ・厚さ10mm
  - 穴の直径8mmと10mmの2種類



口笛演奏時の声道模型

# 実験方法

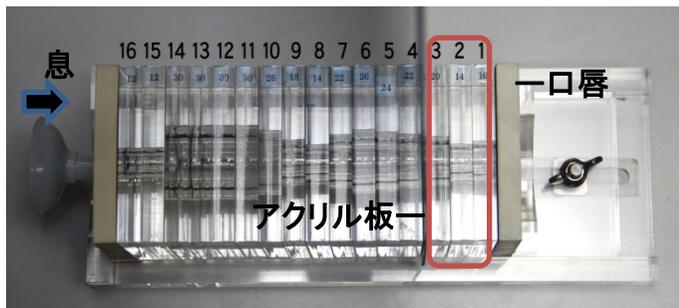
- 図の左側から息を吹き入れる
- 右側の口唇側から出力される音を、20cm離して置いたマイクロホンで録音する
- サンプル周波数44.1kHz    • 量子化ビット数16ビット
- 定常部の中央4096ポイントに対してハニング窓をかけ、FFTを行いピッチ周波数を求める



- 口唇側からゴム板は含めずに1番目と2番目のプレートの間にアクリル板を挿入した場合、その場所を「1 | 2」と表す。

# 実験結果

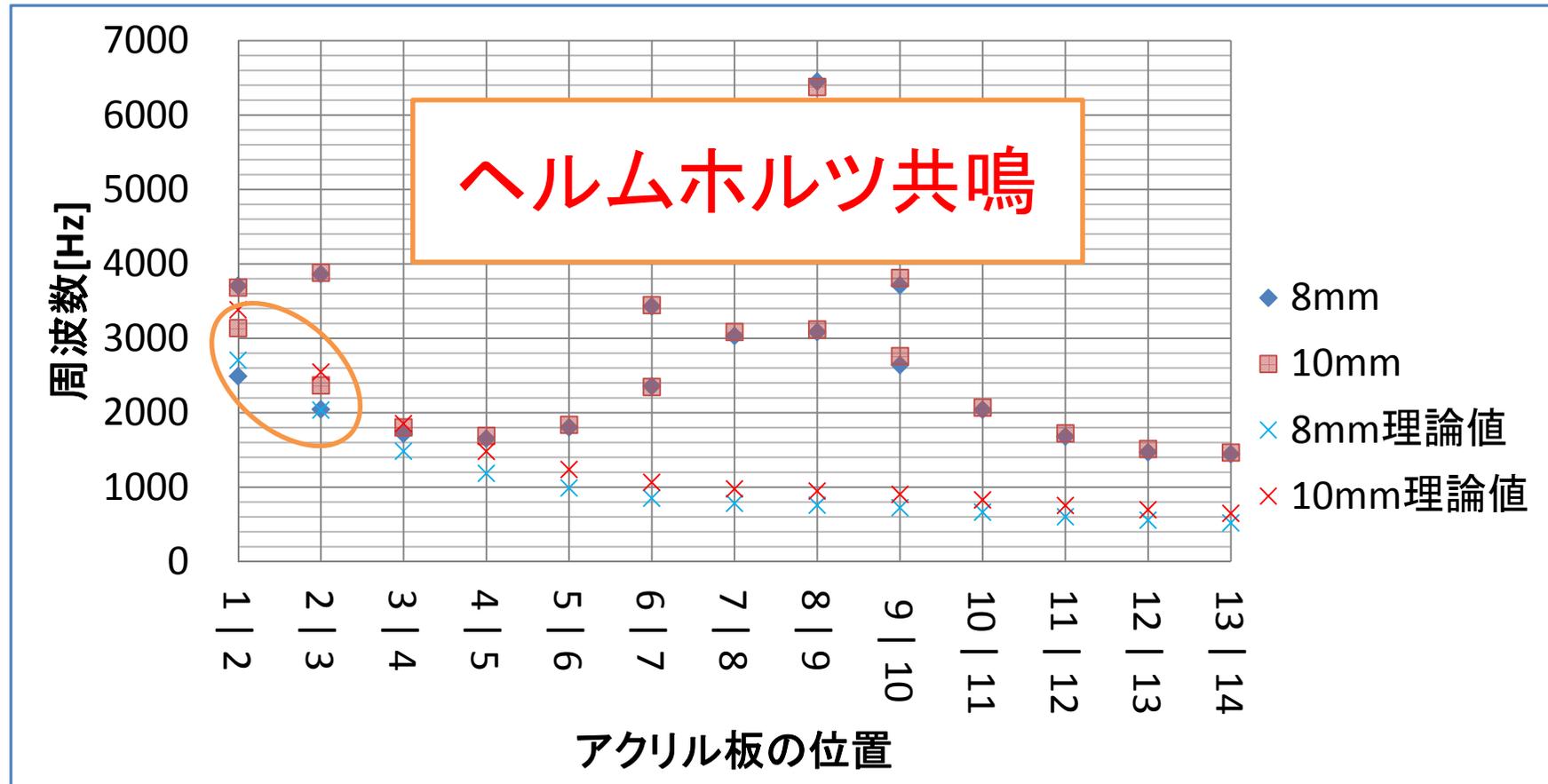
- アクリル板で仕切られた右側(口唇側)を共鳴腔と考え, ゴム板の穴の直径をそれぞれ8mmと10mmにしたときの  
**ヘルムホルツ共鳴周波数の理論値**



声道模型を用いた口笛音の共振周波数

アクリル板の位置	測定値[Hz]		ヘルムホルツ共鳴周波数の理論値[Hz]	
	8mm	10mm	8mm	10mm
1   2	2489.3	3137.4	2705.6	3382.0
	3705.8	3682.2		
2   3	2045.7	2366.5	2036.2	2545.2
	3860.9	3882.4		
3   4	1716.2	1802.3	1483.1	1853.9
4   5	1651.6	1690.4	1184.4	1480.5
5   6	1802.3	1836.8	990.0	1237.5
6   7	2355.8	2347.1	851.0	1063.7
	3428.1	3445.3		
7   8	3031.8	3085.8	781.0	976.3
	5165.8	5157.2		
8   9	3085.7	3120.0	757.3	946.6
	4573.7	6376.0		
9   10	2640.0	2760.5	722.3	902.9
	3710.2	3781.2		
10   11	2039.2	2069.4	662.6	828.3
11   12	1679.6	1722.7	602.2	752.7
12   13	1468.9	1513.8	555.7	694.7
13   14	1444.9	1464.3	518.6	648.3

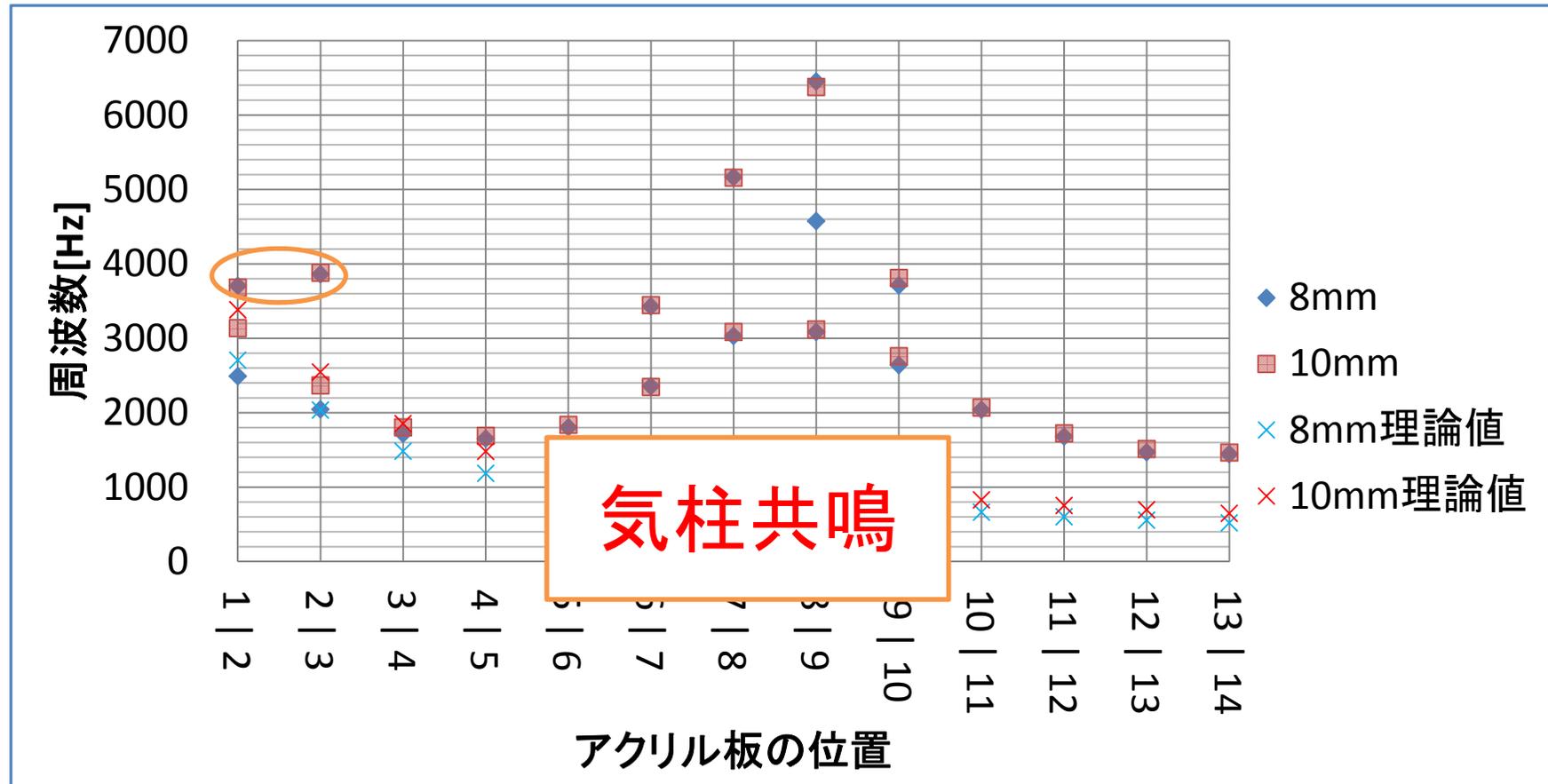
# 実験1の結果



「1 | 2」「2 | 3」において

- ・測定値と理論値の誤差が10%以内である
- ・ゴム板の穴の直径が8mmと10mmのときで共振周波数が大きく異なる

# 実験1の結果



オーバーブローイングで高い音が出る場合

- ・理論値とは大きく異なっている
- ・ゴム板の穴の直径が8mmと10mmでピッチがほとんど変わらない

## 実験2

1. 声道模型を用いたピッチ周波数測定
- ➡ 2. 呼気の流速とピッチ周波数の関係  
呼気の流速を測定する
3. 単純化モデルの場合のピッチ周波数測定

# 実験方法

流速計を取り付ける(ゴム板の穴の直径8mm)



ゆっくりと息を吹き始め、徐々に呼気の上速度を上げる

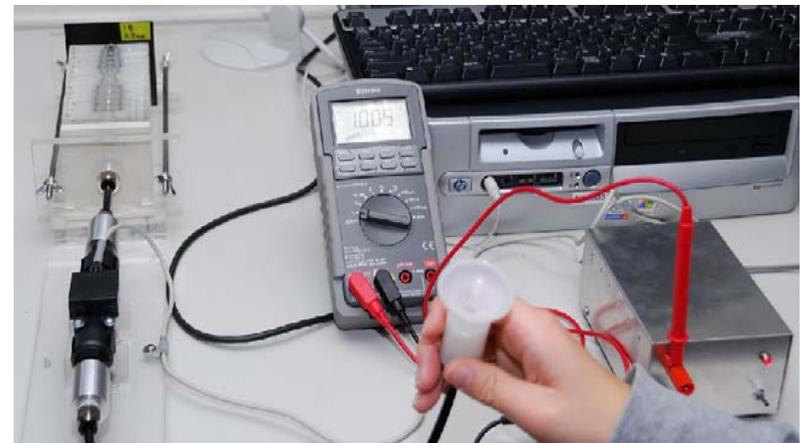


二か所の流速とピッチ周波数を測定する

- ・音が鳴り始める箇所
- ・音が鳴らなくなる直前

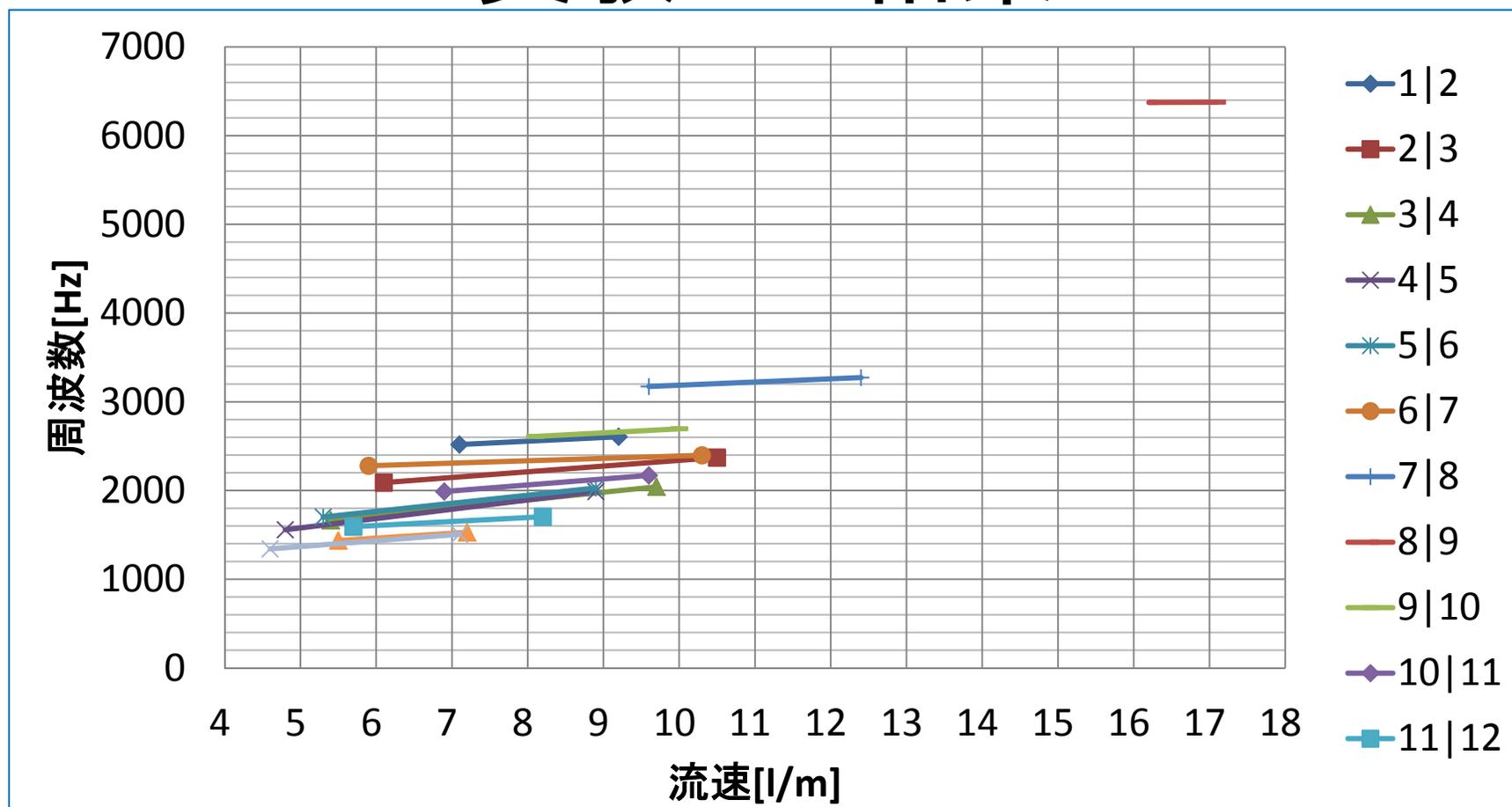


流速に比例した電圧が出力される  
デジタルマルチメータで測定する



実験2の様子

# 実験2の結果



- 呼気の流速が上がると、ピッチ周波数が高くなる
- 流速によるピッチ周波数の変動幅は最大で422.6cent

# 実験3

1. 声道模型を用いたピッチ周波数測定

2. 呼気の流速とピッチ周波数の関係

➡ 3. 単純化モデルの場合のピッチ周波数測定

気柱共鳴における共振周波数について詳細な  
検討を行う

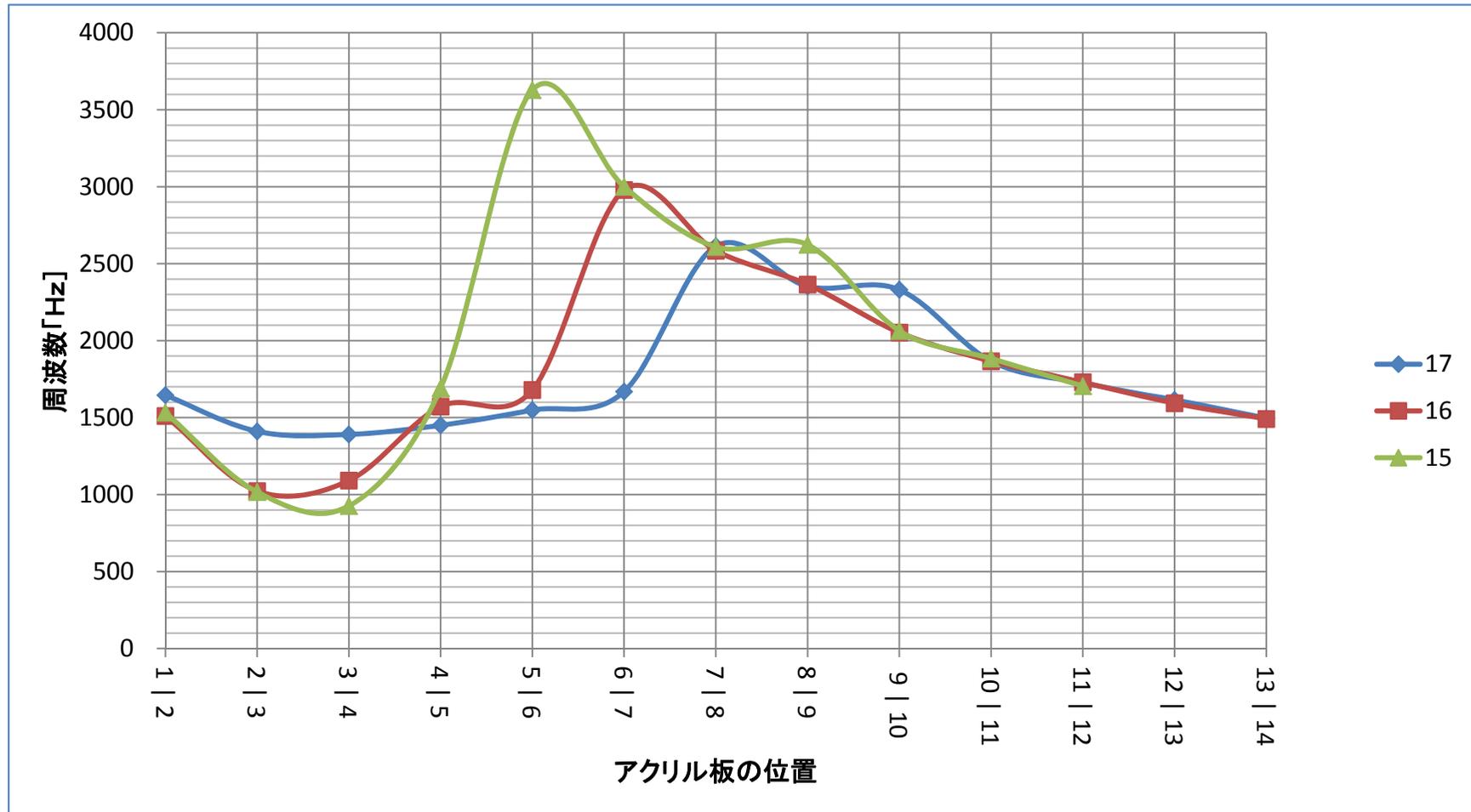
# 実験3の方法

- 声道模型の/u/の形状のプレートを穴の直径30mmに差し替える
- ゴム板の穴の直径8mm
- プレート枚数17枚、16枚、15枚の3種類

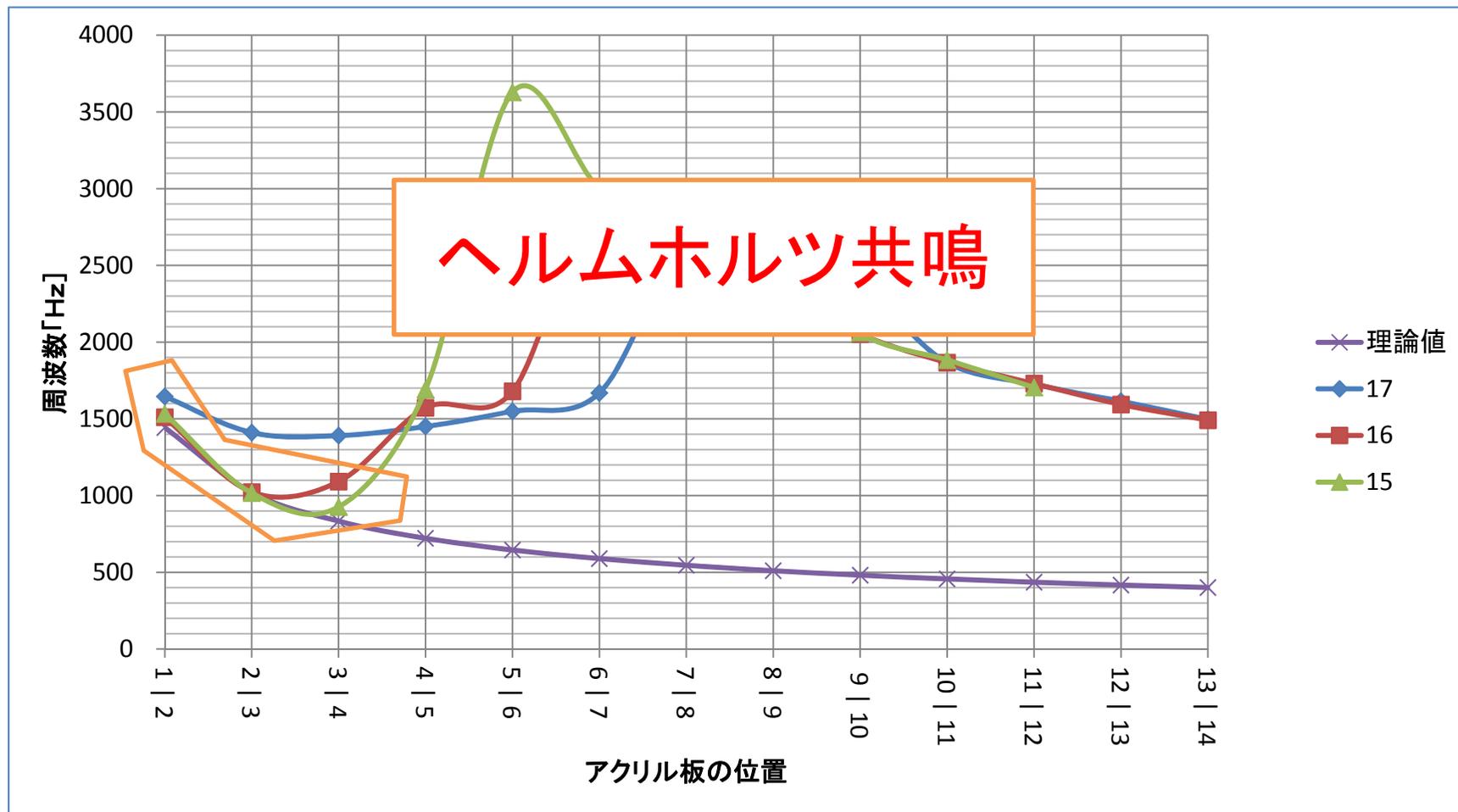


単純化モデル

# 実験3の結果

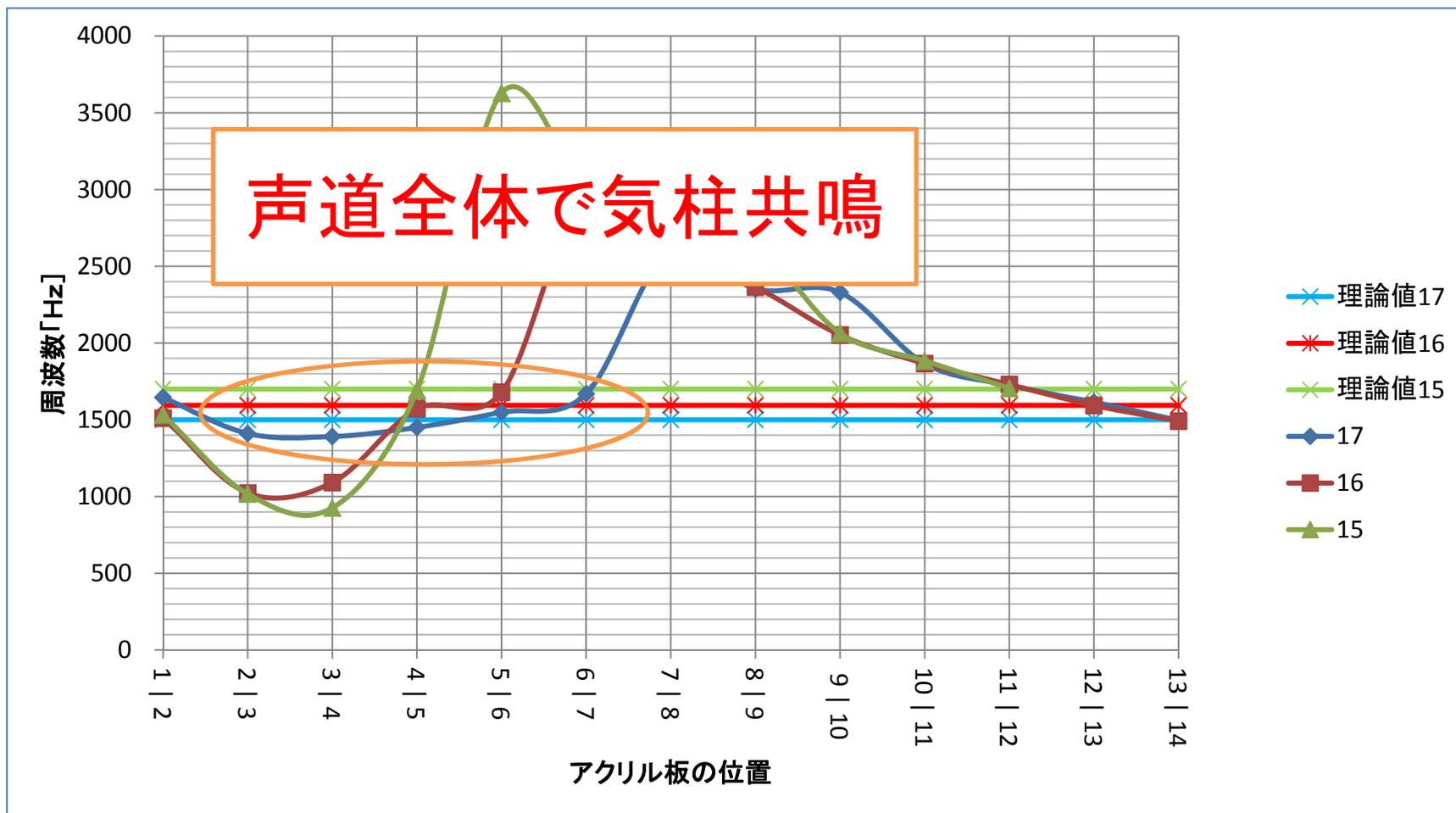


# 考察1



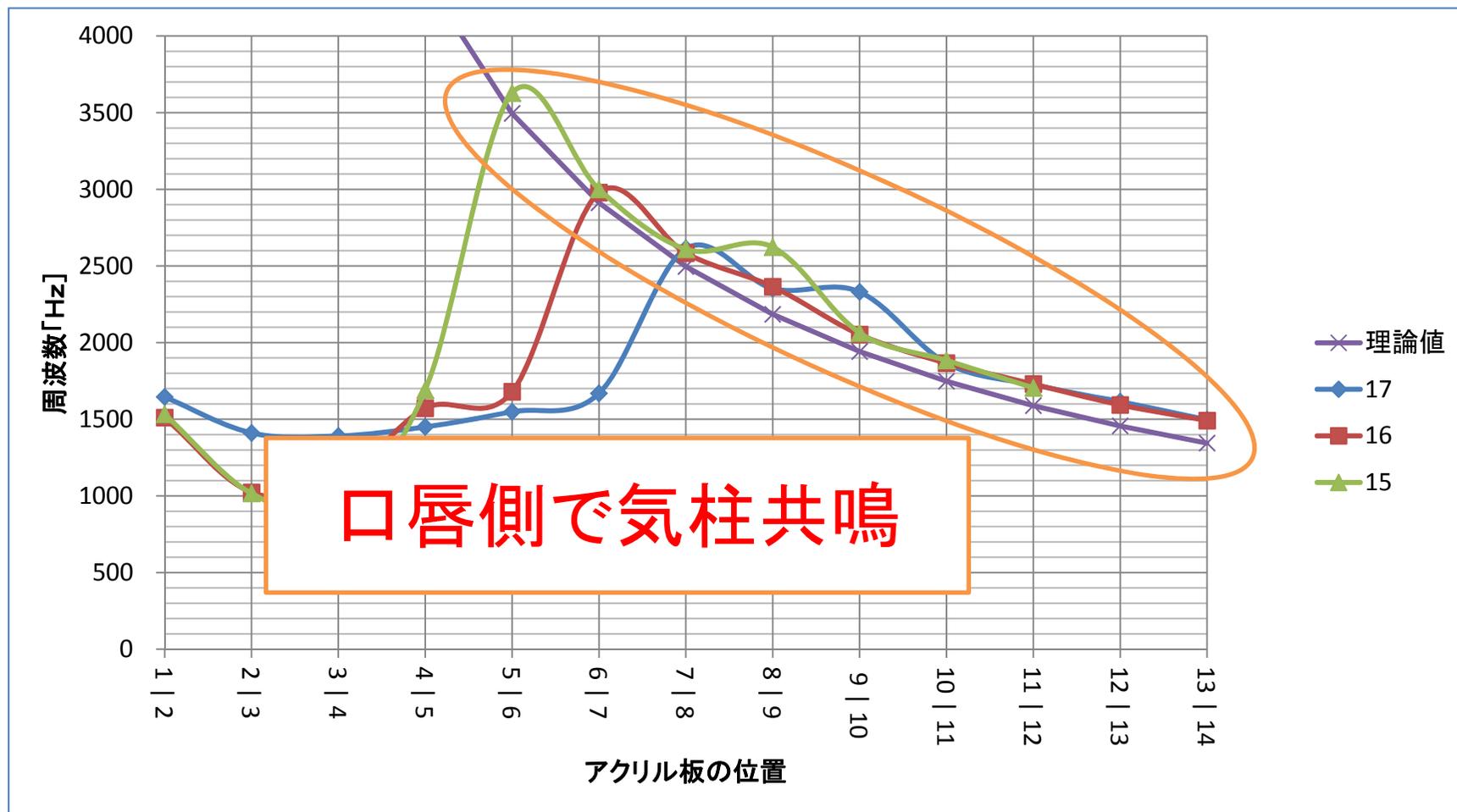
ヘルムホルツ共鳴周波数の理論値との比較

# 考察2

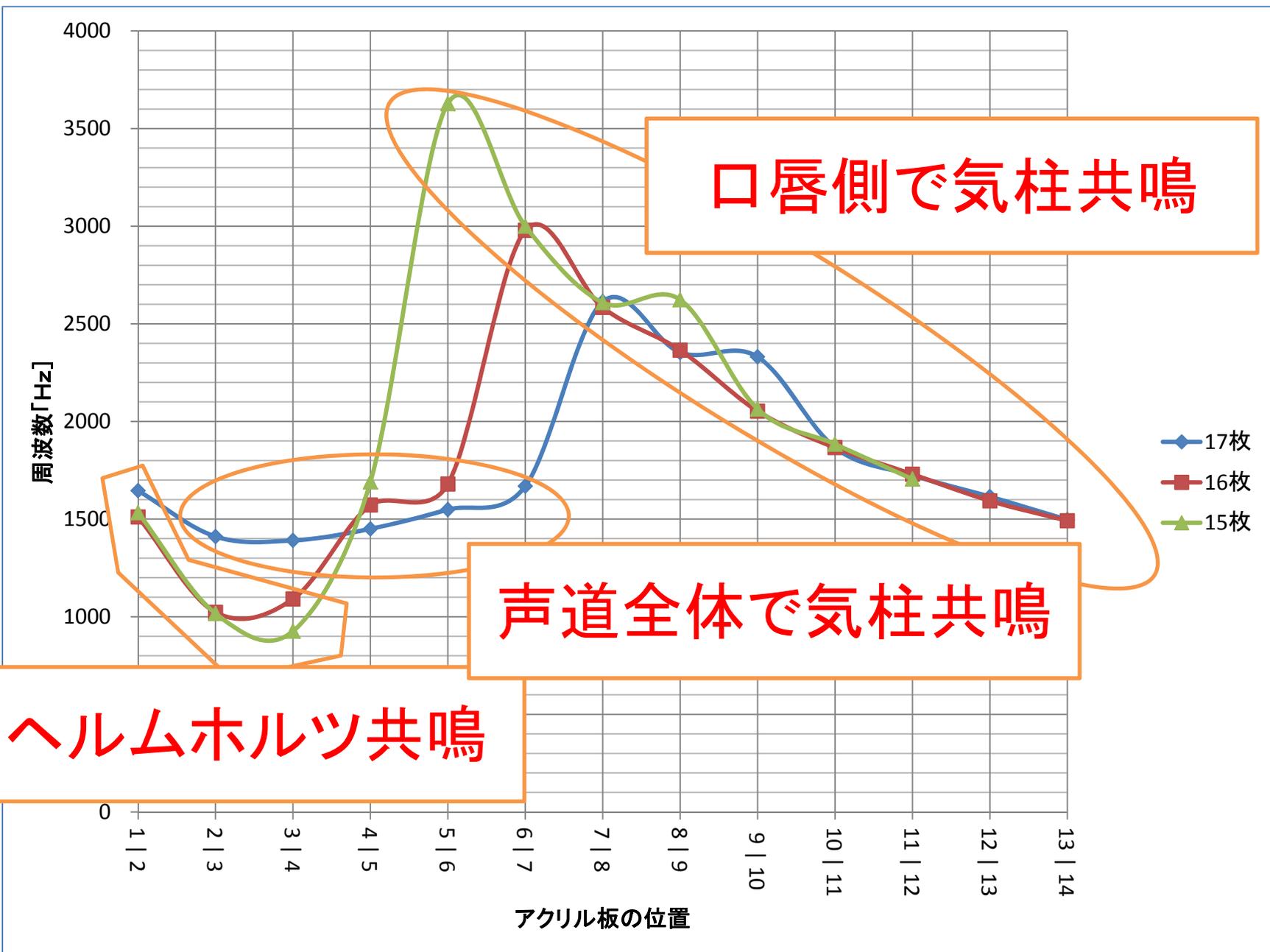


声道全体気柱共鳴の理論値との比較  
(片側開管の3倍振動)

# 考察3



口唇側気柱共鳴の理論値との比較(管内温度30°C,  $c: 349.5 \text{ m/s}$ )  
(両側開管の基本振動)



口唇側で気柱共鳴

声道全体で気柱共鳴

ヘルムホルツ共鳴

## まとめ

- 実験で用いたモデル, 実際の口笛はともに従来報告されているヘルムホルツ共鳴だけでなく**気柱共鳴の領域が存在する**と考えられる。
- 気柱共鳴の領域において、**声道全体**で気柱共鳴(片側開管)している場合と、**狭めの位置から口唇までの部分**で気柱共鳴(両側開管)している場合の2種類があると考えられる。

# 超音波を使う新しい教育研究領域の創成の試み

得永 嘉昭(金沢工大)

# 超音波を使う新しい教育研究領域の 創成の試み

金沢工業大学電気電子工学科  
得永 嘉昭



# 社会的背景

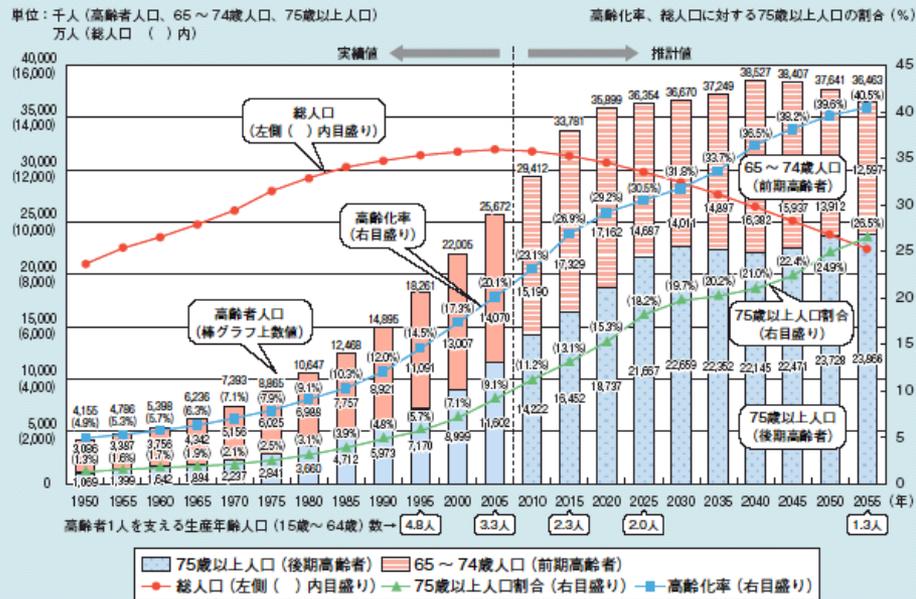
- 日本を取り巻く環境が急速に大きく変わった。
- 具体的に
- 日本が情報化と国際化で複雑系の開いた系になった。
- エネルギーと物質(人材)の導入でカオス状態になった。
- 急速な老齡化と少子化がそれを加速する。

新しい自己組織化が始まる。

教育はどうすべきか？

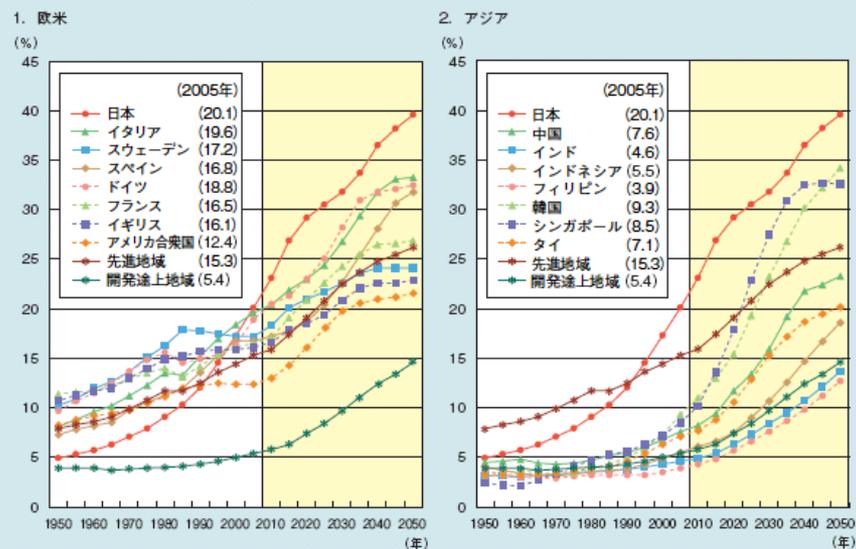


図1-1-2 高齢化の推移と将来推計



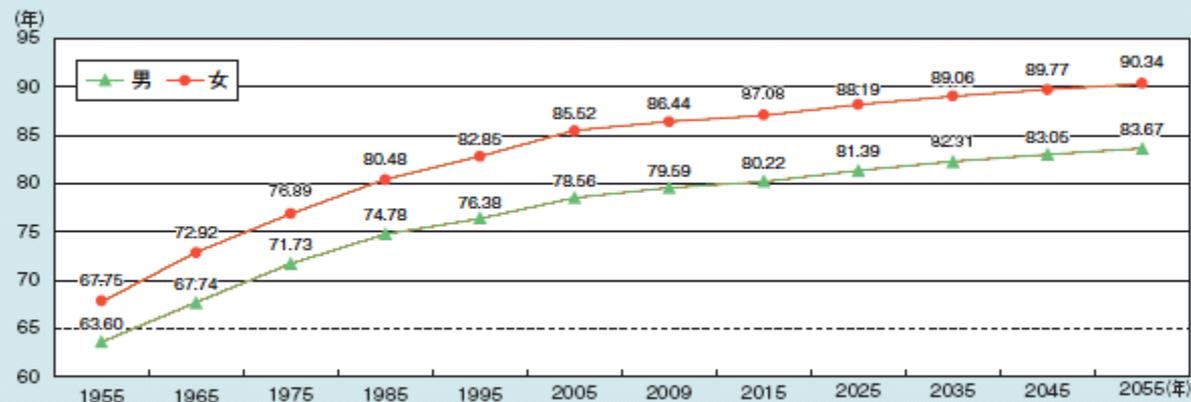
資料：2005年までは総務省「国勢調査」、2010年以降は国立社会保障・人口問題研究所「日本の将来推計人口(平成18年12月推計)」の出生中位・死亡中位仮定による推計結果

図1-1-6 世界の高齢化率の推移



資料：UN, World Population Prospects: The 2008 Revision  
ただし日本は、2005年までは総務省「国勢調査」、2010年以降は国立社会保障・人口問題研究所「日本の将来推計人口(平成18年12月推計)」の出生中位・死亡中位仮定による推計結果による。  
(注) 先進地域とは、北部アメリカ、日本、ヨーロッパ、オーストラリア及びニュージーランドからなる地域をいう。  
開発途上地域とは、アフリカ、アジア(日本を除く)、中南米、メラネシア、ミクロネシア及びポリネシアからなる地域をいう。

図1-1-4 平均寿命の推移と将来推計



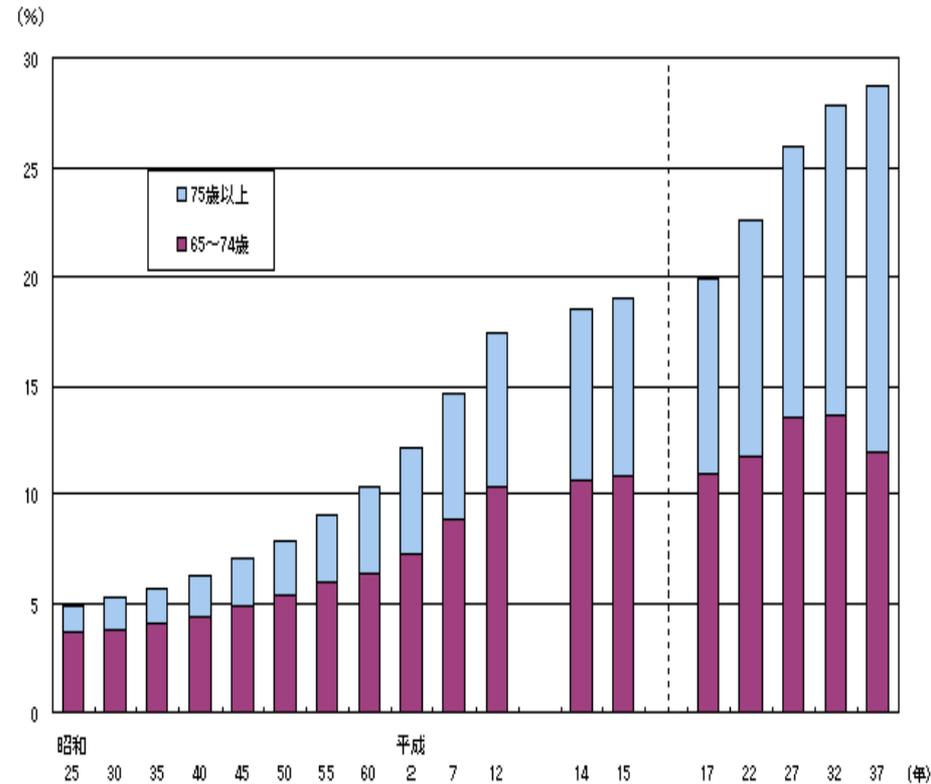
資料：2005年までは、厚生労働省「完全生命表」、2009年は厚生労働省「簡易生命表」  
2015年以降は、国立社会保障・人口問題研究所「日本の将来推計人口(平成18年12月推計)」の出生中位・死亡中位仮定による推計結果

平成23年版高齢社会白書  
「平成22年度 高齢化の状況及び高齢社会対策の実施状況」(内閣府)より抜粋

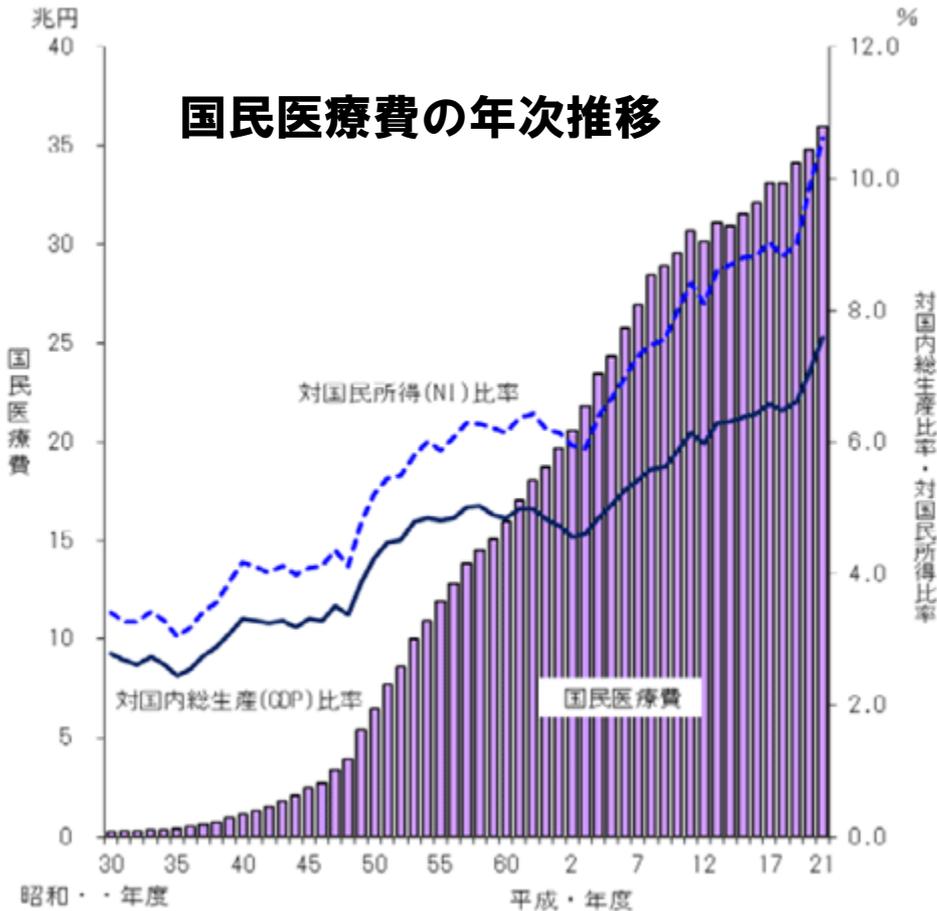
# 「高齢者人口の現状と将来」 (総務省統計局)より抜粋

資料:平成12年度までは「国勢調査」、平成13年および14年は「推計人口」、平成17年以降は国立社会保障・人口問題研究所「日本の将来推計人口ー平成14年1月推計」(中位推計)

図1 高齢者人口の割合の推移



## 国民医療費の年次推移



## 「平成21年度 国民医療費の概況」 (厚生労働省)より抜粋



# 太陽は東から昇り西へ沈む

- 自然界の大法則
- これまでの科学技術もしかり
- 新しい科学技術はアメリカで創成され、
- それが日本へ導入改良されてきた。
- 現在はタイやベトナムへと移行していく
- (情報とエネルギーの流れ)
- “太陽は東から上がり西に沈むという
- 事実と合致する“
- 



# 製造業の空洞化、高齢化問題

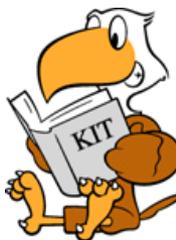
自然の摂理に逆らうこと辞めるべき  
逆説的にそれをチャンスととらえるべきで  
ある。

例：高齢化問題(事実は小説より奇なり)

他の国に先駆けて多くのノウハウを蓄積できる  
そのための電子部品は高信頼性あり、

システムの構築が苦手

実践的な問題点の発掘と解決法の構築



# これからのキーワード

膨大な金額が必要

- 宇宙
- 海洋
- 環境

- 食糧
- 人間

金のかかるものもあるが、  
身近にできるものもある



# 新しい**教育研究領域**の**創成**の問題点

- **パラダイムの壁**

1. **アプローチの仕方**

分子生物学、バイオ化学

2. **理論的バックグラウンドの有無**

バイオは確率論的、超音波は決定論的

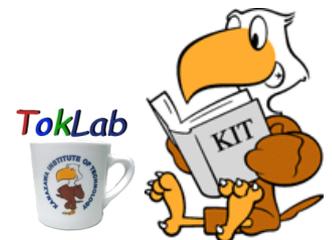
3. **マクロの世界とミクロの世界**

バイオは細胞の世界、超音波は波動の世界



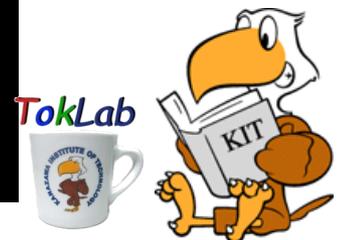
# 1.パラダイムの違いの例

- 導入
- 超音波の領域      injection    introduction
- バイオの領域      delivery    transfection
  
- ベクター
- 超音波の領域      方向と大きさをもったもの
- バイオの領域      ウイルスを細胞内へ入れる



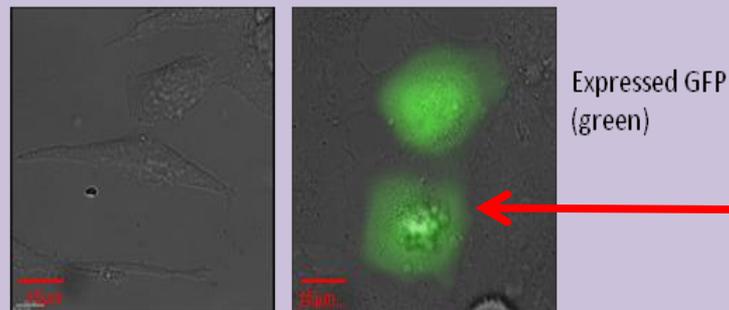
## 2.理論的バックグラウンド

- 超音波の世界は**決定論的**
- 理論的な予想ができる
  
- バイオの世界は**確率論的**
- 対象が生き物であり**多様性**と**冗長性**を有する



# 3. ミクロの世界とマクロの世界

- **バイオ**はミクロの世界



細胞核、  
ミトコンド  
リア、ゴル  
ジ体、小  
胞体その  
他の器官

- **超音波**はマクロの世界
- 協同作用としての波
- 

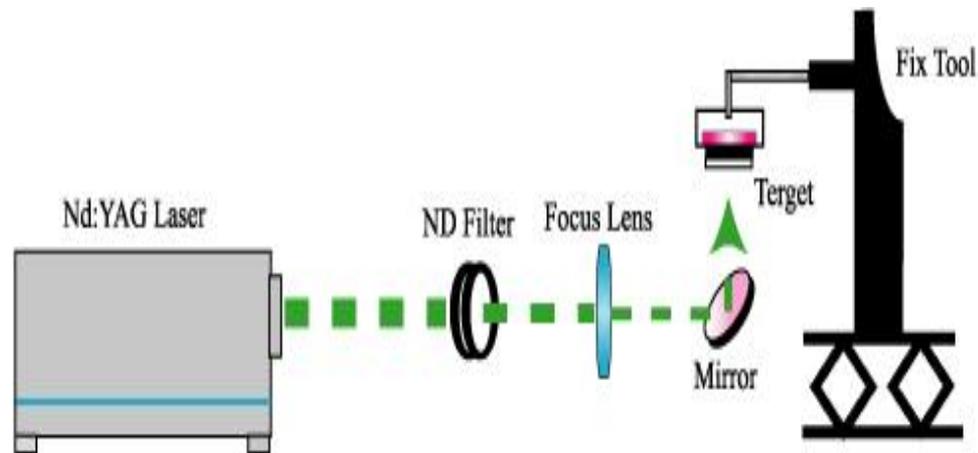


# 挑戦的萌芽教育研究の試み

- レーザ、超音波関係
- 電気系教授 得永 嘉昭
- 細胞と遺伝子導入関係
- バイオ化学教授 小 木美恵子
- デバイス関係
- 電気系教授 會澤 康治
  
- その概略を簡単に述べる:

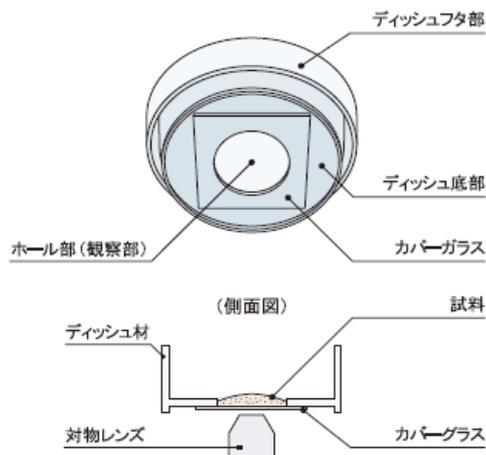


# レーザーとターゲットの相互作用

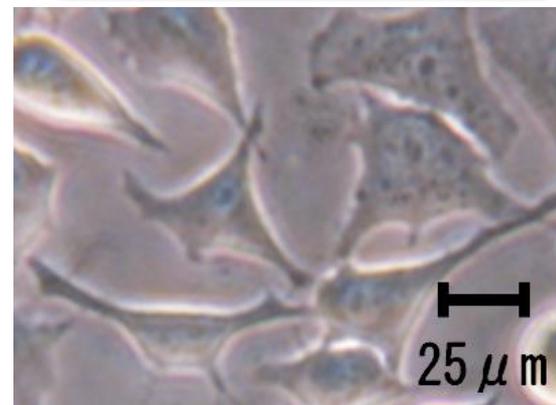


## ターゲットの種類、構造

### 細胞と培養液を入れるための道具

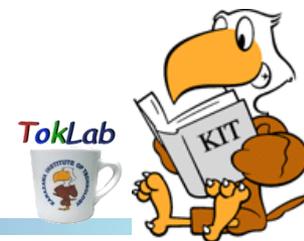


### 導入される細胞



HeLa細胞  
(ヒト子宮頸癌由来の細胞)

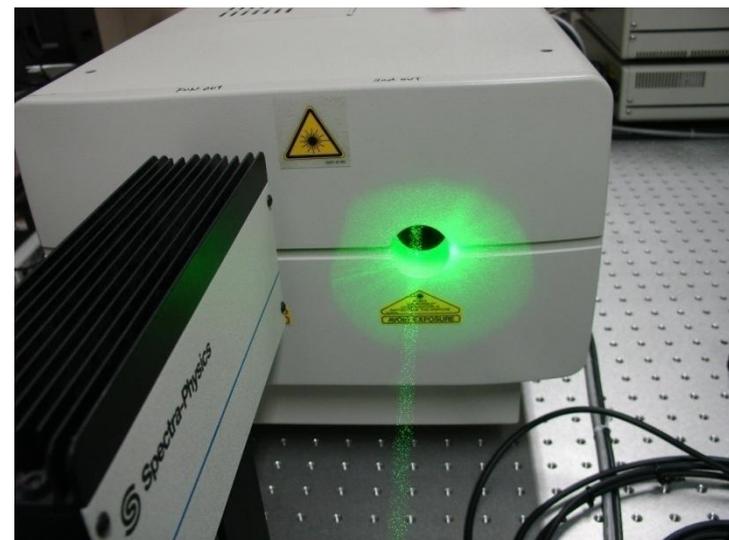
# ラボタイプNd:YAGレーザー



- Nd:YAG(Spectra Physics Quanta-Ray LAB-130)
- 第二高調波(532nm)
- 単発発振が可能



繰り返し周波数10[Hz]  
パルス幅 7-10[ns]  
レーザー出力200[mJ/pulse]

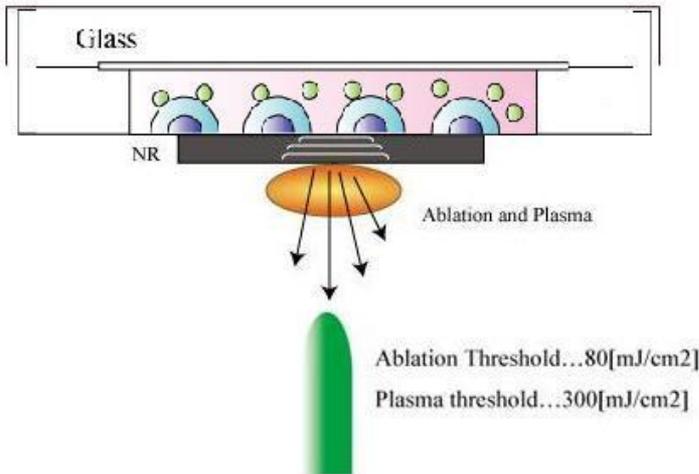


# Experimental conditions and figure of the principle



## Experimental conditions

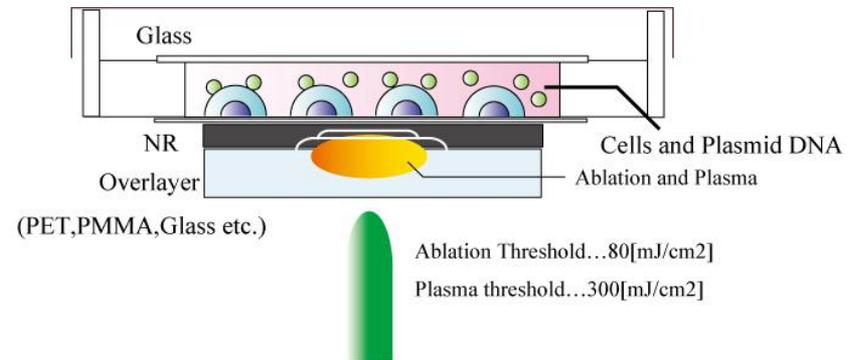
The thickness of the NR [mm]	0.5
Laser Fluence [J/cm <sup>2</sup> ]	2.4
cell	HeLa cell
The concentration of Plasmid DNA [μg/μl]	1.0



Irradiated Pulse Laser  
**Direct type**

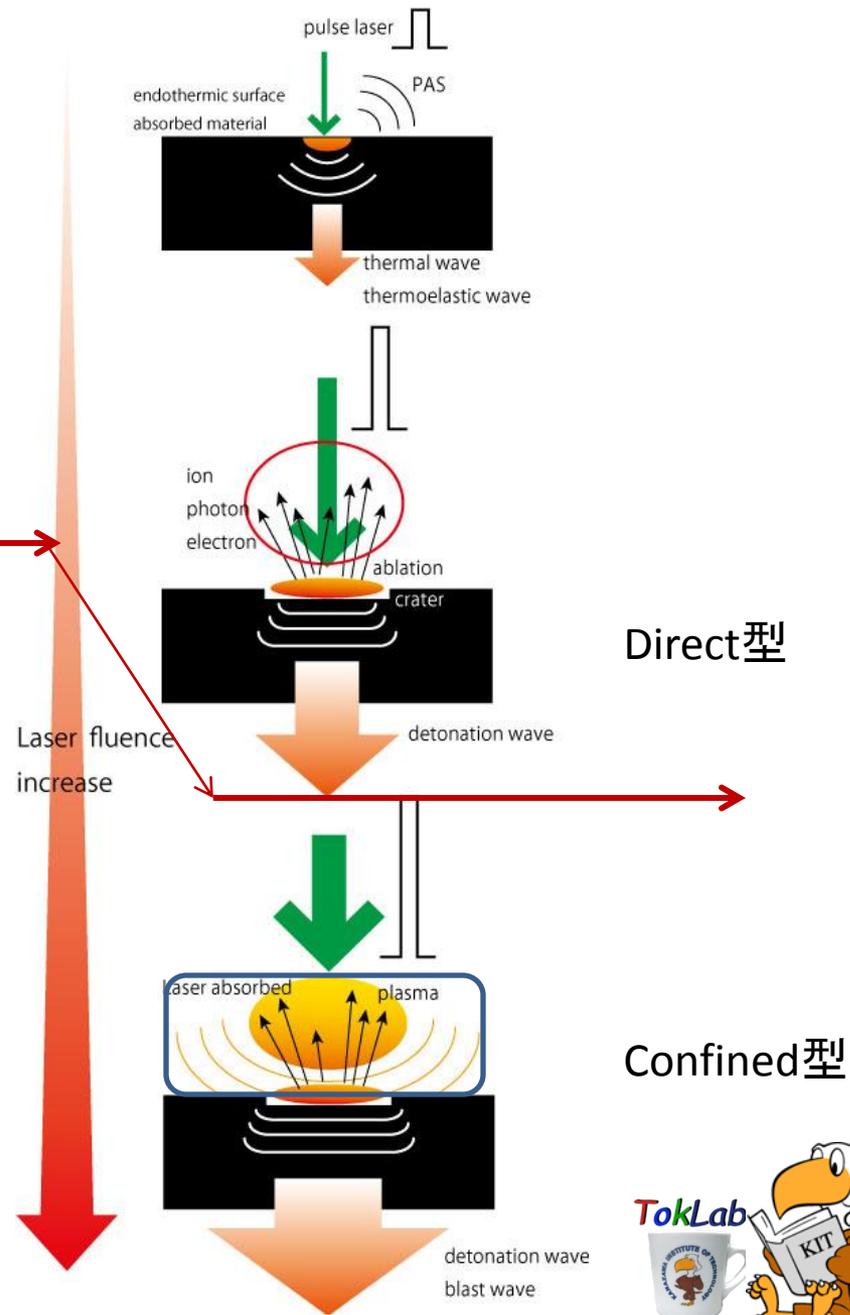
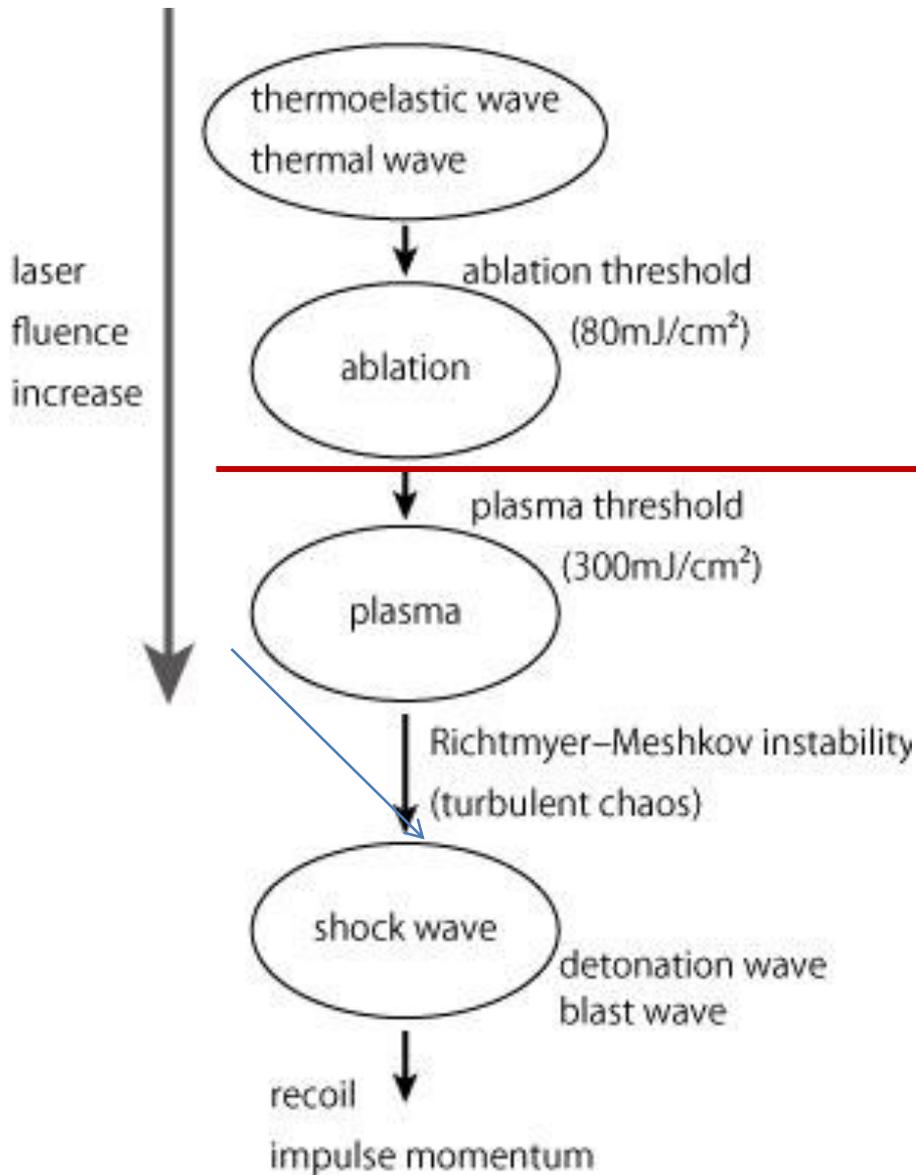
## Experimental conditions

The thickness of the NR [mm]	0.5
The thickness of the PET [mm]	1.0
The thickness of the Adhesive [μm]	20
Laser Fluence [J/cm <sup>2</sup> ]	2.4
cell	HeLa cell
The concentration of Plasmid DNA [μg/μl]	1.0



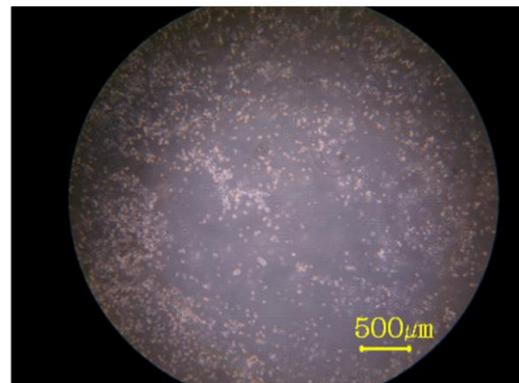
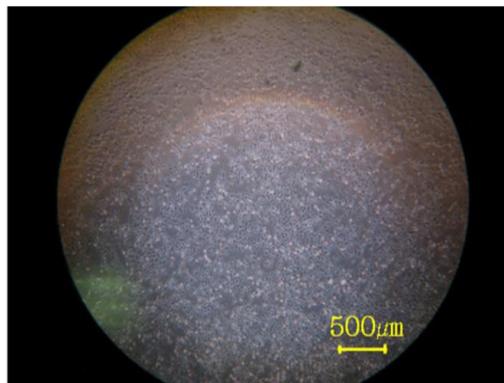
Irradiated Pulse Laser  
**Confined type**



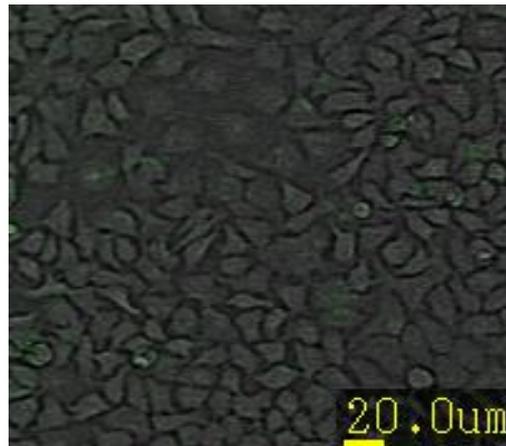


# Confined型素子の遺伝子導入結果

細胞接着状態



pEGFP発現細胞



LISW処理前



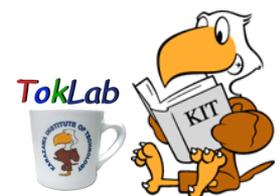
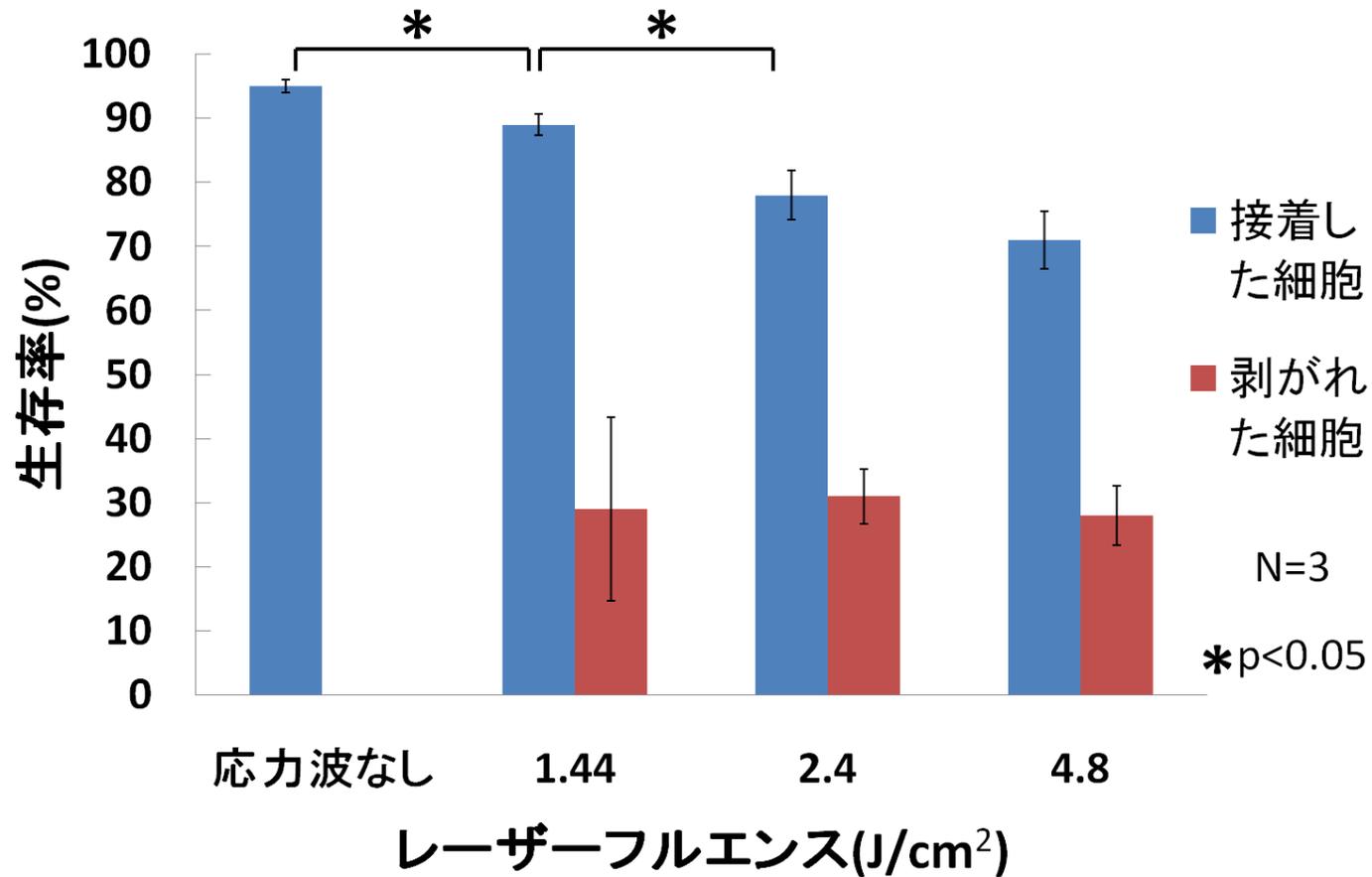
LISW処理後

導入に成功したHeLa細胞

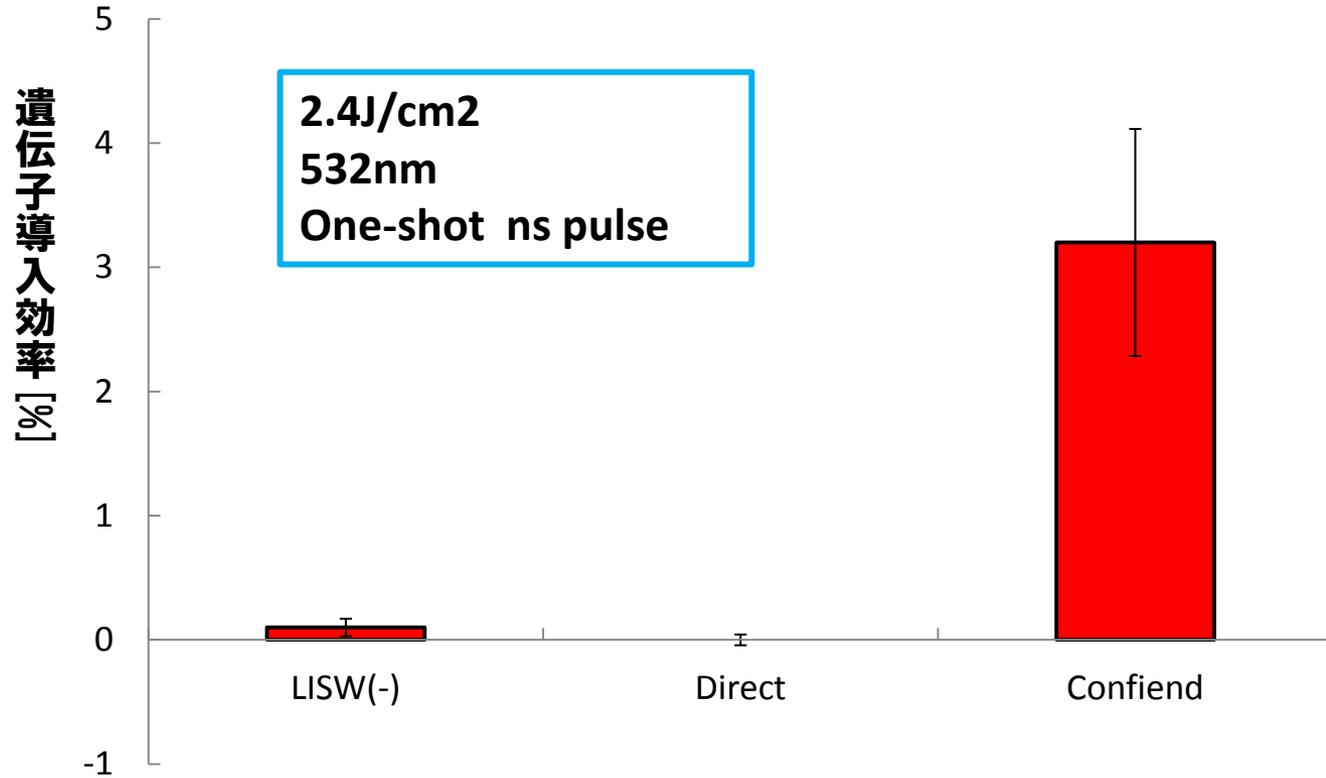
Direct型では遺伝子導入はされなかったが、  
Confined型素子では遺伝子導入を行うことができた。



# HeLa細胞の生存率



# 遺伝子導入効率



# 今後の進め方

- HL60細胞(白血病の癌細胞)への導入
- 皮膚癌への薬品導入法
- 超音波領域として
  - 効率的なターゲットの研究
  - 物理的、生理的な学問的視点の検討

# 謝辞

- 本教育研究を進めるにあたり、ご協力とご指導をいただいている小木美恵子教授、會澤康治教授に心より謝意を表します。

実験に協力をしてきている西脇 基晃君(得永研院生)、北村祐士君、長崎沙織君、牧田 裕起君(小木研学部生)にお礼を申し上げます。

本研究の一部は文部科学省の挑戦的萌芽研究助成の基に行われている。



キャラクターボイスの印象について：

萌える声、萌えない声

山田真司（金沢工大）

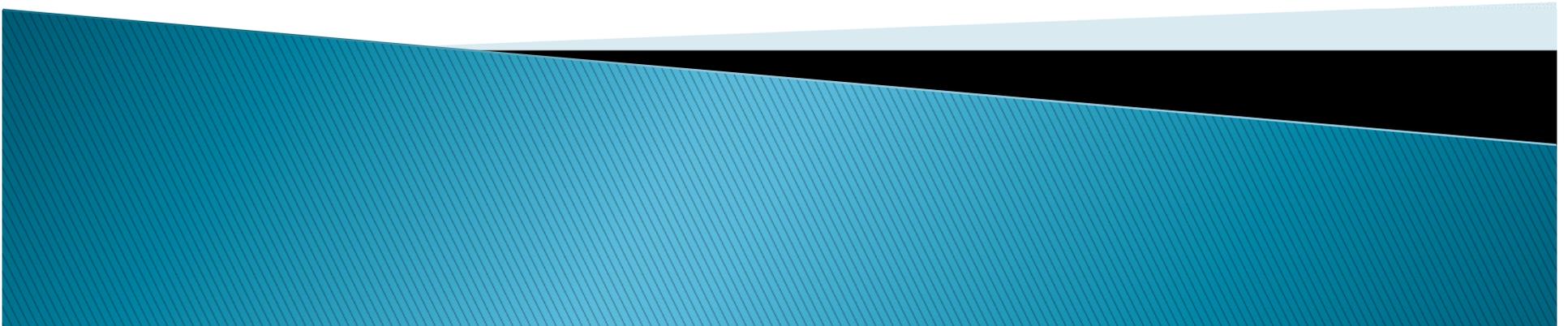
塚本将成

加藤典

小柳拓巳

# キャラクターボイスの印象について: 萌える声、萌えない声

金沢工業大学  
山田真司



# 第1回音響・超音波JMにて

- ▶ 山田研におけるSD(意味微分)法(Osgood, 1952)と因子分析を用いた、音楽の聴取印象やゲームの印象に関する研究について紹介
- ▶ SD法＋因子分析の手法は様々な対象に応用可能
- ▶ 音響学の分野では、もともと音色の研究で盛んに用いられてきた



# 音の3要素

- ▶ 音の高さ  
低い音—高い音 : 1次元
- ▶ 音の大きさ  
小さな音—大きな音 : 1次元
- ▶ 音色  
明るい音, 迫力のある音, 荘厳な音,  
さわやかな音,... : 多次元 ←何次元?



	非常に	かなり	やや	どちら でもない	やや	かなり	非常に	
きれい	○							きたない
滑らかな	○							粗い
力強い					○			弱々しい
迫力ある					○			迫力のない
明るい						○		暗い
鋭い						○		鈍い
うるさい						○		静かな
華やかな					○			みすぼらしい
潤った	○							かさかさした
荘厳な				○				みすぼらしい
派手な					○			地味な
重い					○			軽い

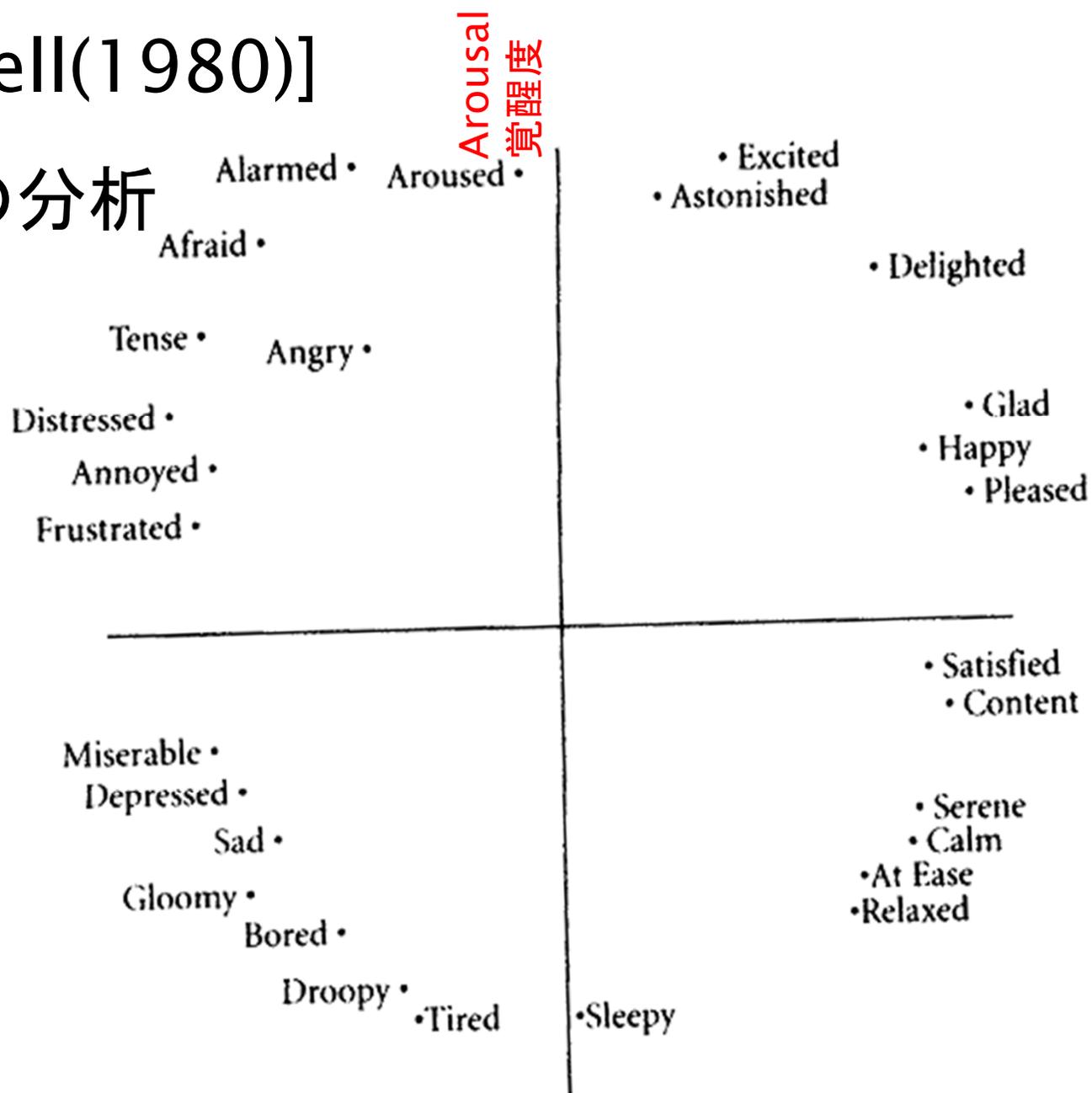
# このような結果を因子分析

- ▶ 同様な動きをする形容詞は同様な意味
- ▶ 独立な形容詞の組は？ → 次元数を決定
- ▶ 数学的には、行列の固有値問題を解くことに相当
- ▶ その結果、音色は、「美的因子」、「迫力因子」、「明暗（金属性）因子」の3次元で構成されていることが分かっている



# [Russell(1980)]

## 感情の分析



Valence  
快-不快

# 「萌え」の研究に 응용

- ▶ 「萌え」: Japan Cool(マンガ、アニメ、ゲーム)の代表的概念
  - 典型的には、秋葉原の「メイド喫茶」
  - 1866億8000万円の市場規模(浜銀総研,2007)
  - 「國文学」で2008に特集: 7篇の論文
  - 「招き寄せたい」という感情ではないか?(山口,2008)
  - 科学的な実験に基づく研究はされていない
  - 統制の取れた実験素材を得にくい



# キャラクターボイスの印象について: 萌える声、萌えない声 研究の背景

- ▶ 近年ゲーム中のセリフに声を付与
  - 大量の音声データを収録
  - 開発費が高騰
- ▶ 声優の起用
  - ディレクターの勘と経験に依存
  - イメージと合わない声優を起用してしまう例も
- ▶ 声優に対してのインストラクション
  - 抽象的なシチュエーション設定が使われている
  - 欲しい発話がなかなか得られない



# 目的

- ▶ 声質によって印象がどのように変化するのか
- ▶ 発話とシチュエーション設定の関係を調査
  - どのようなシチュエーション設定
  - どのような発話を得られるのか



# 実験素材

## ▶ 実験素材の用意

- 「おにいちゃんCD」
  - 声優12名
  - 100種類のシチュエーション設定
  - 「おにいちゃん」というセリフ
  - 合計1200トラック収録

🔊 声優A: チャンネル争いしているような

🔊 声優A: ちょっと甘えるように

🔊 声優B: ちょっと甘えるように

# 実験概要

- ▶ 実験素材
  - 「おにいちゃんCD」に収録された5名の声優
  - 35種類のシチュエーション設定
  - 合計175刺激
- ▶ 実験場所
  - 金沢工業大学八束穂キャンパス内 防音室
- ▶ 刺激呈示
  - ヘッドフォンを使用して50～86[dB(L<sub>A</sub>)] で呈示
- ▶ 実験参加者
  - 金沢工大生 10名



# 評定尺度

## ▶ 26対両極7段階評定尺度を使用

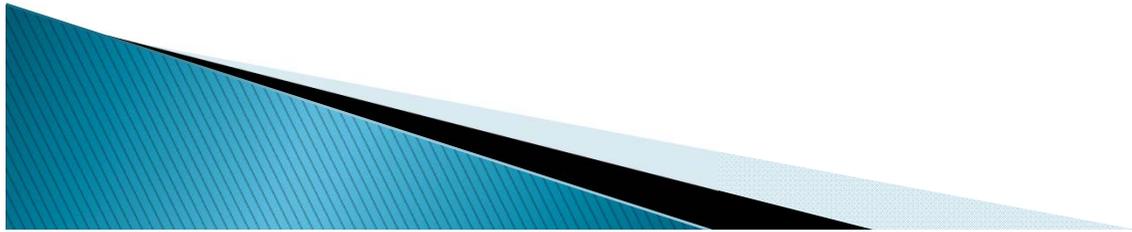
癒される - 疲れさせる  
快い - 不快な  
かわいらしい - にくらしい  
好き - 嫌い  
きれいな - きたない  
丁寧な - ぞんざいな  
温かい - 冷たい  
やさしい - 厳しい  
甘えた - そっけない  
元気のある - 元気のない  
力強い - 弱々しい  
活動的な - おとなしい  
強気な - 弱気な

意欲的な - 無気力な  
明るい - 暗い  
派手な - 地味な  
うきうきした - 沈んだ  
陽気な - 悲しい  
かよわい - 勇ましい  
静かな - うるさい  
落ち着きのある - 落ち着きのない  
重い - 軽い  
子供っぽい - 大人びた  
自然な - 不自然な  
なめらかな - しどろもどろな  
生意気な - 生意気でない

# 評定尺度

- ▶ 基本情動(喜怒哀楽及び恐怖)、「萌え」の度合い
  - 単極7段階評定尺度を使用

	1	2	3	4	5	6	7	
全く喜んでいない								非常に喜んでいる
全く恐れていない								非常に恐れている
全く怒っていない								非常に怒っている
全く哀しんでいない								非常に哀しんでいる
全く楽しんでいない								非常に楽しんでいる
全く萌えない								非常に萌える



# シチュエーション判断

- ▶ その発話がどのようなシチュエーション設定のもとに発話されたかを判断
  - 合計63種類のシチュエーション設定
    - 使用した35種類のシチュエーション設定
    - CD収録の28種類のシチュエーション設定
  - 複数回答可



# シチュエーション判断

## ▶ シチュエーションの判断の評価例



声優A: ちょっと喜んで

チェック	シチュエーション	チェック	シチュエーション
<input checked="" type="checkbox"/>	ちょっと喜んで	<input type="checkbox"/>	すごく喜んで
<input type="checkbox"/>	ちょっと怒って	<input type="checkbox"/>	すごく怒って
<input type="checkbox"/>	ちょっと悲しい	<input type="checkbox"/>	すごく悲しい
<input checked="" type="checkbox"/>	ちょっと楽しい	<input type="checkbox"/>	すごく楽しい
<input type="checkbox"/>	ちょっとバカにして	<input type="checkbox"/>	すごくバカにして
<input type="checkbox"/>	ちょっと泣いている	<input type="checkbox"/>	すごく泣いている
<input type="checkbox"/>	ちょっとくすぐったい	<input type="checkbox"/>	すごくくすぐったい
<input type="checkbox"/>	ちょっと暑い	<input type="checkbox"/>	すごく暑い
<input type="checkbox"/>	ちょっと寂しい	<input type="checkbox"/>	すごく寂しい
<input type="checkbox"/>	ちょっと寒い	<input type="checkbox"/>	すごく寒い
<input type="checkbox"/>	ちょっと遠くから呼んでみる	<input type="checkbox"/>	すごく遠くから呼んでみる
<input type="checkbox"/>	ちょっと恥ずかしい	<input type="checkbox"/>	すごく恥ずかしい



# 分析結果

- ▶ 聴取者にわたる平均評定値
- ▶ 因子分析(主因子法)
  - 第2因子までで累積寄与率が約79 %
  - 2因子解を採用



# 因子負荷量

- ▶ 第1因子
  - 快さ
- ▶ 第2因子
  - 興奮度

Russell (1980)  
の感情の分析結果  
と合致

評定尺度		因子負荷量	
		快さ	興奮度
癒される	- 疲れさせる	-0.98	0.08
快い	- 不快な	-0.96	0.01
かわいらしい	- にくらしい	-0.94	0.14
好き	- 嫌い	-0.93	0.03
きれいな	- きたない	-0.93	0.08
丁寧な	- ぞんざいな	-0.89	0.31
温かい	- 冷たい	-0.80	-0.26
やさしい	- 厳しい	-0.78	0.34
甘えた	- そっけない	-0.71	0.27
元気のある	- 元気のない	-0.06	-0.99
力強い	- 弱々しい	0.28	-0.93
活動的な	- おとなしい	0.28	-0.92
強気な	- 弱気な	0.29	-0.90
意欲的な	- 無気力な	-0.17	-0.89
明るい	- 暗い	-0.38	-0.88
派手な	- 地味な	0.04	-0.85
うきうきした	- 沈んだ	-0.38	-0.81
陽気な	- 悲しい	-0.33	-0.81
かよわい	- 勇ましい	-0.49	0.82
静かな	- うるさい	-0.51	0.75
落ち着きのある	- 落ち着きのない	-0.38	0.63
重い	- 軽い	0.68	0.49
子供っぽい	- 大人びた	-0.51	0.04
自然な	- 不自然な	-0.67	-0.02
なめらかな	- しどろもどろな	-0.64	-0.26
生意気な	- 生意気でない	0.63	-0.64
累積寄与率(%)		45.13	78.91

# 印象と基本情動との関係

## ▶ 重回帰分析

- 独立変数: 因子得点
- 従属変数: 喜怒哀楽及び恐怖、萌えの度合いの平均評定値

	決定係数	標準化偏回帰係数	
		快さ	興奮度
恐怖	0.56	-0.06	-0.75
喜び	0.57	0.55	0.52
怒り	0.48	-0.58	0.37
哀しい	0.64	-0.10	-0.79
楽しい	0.53	0.48	0.55
萌え	0.87	0.93	-0.01

- 特に「萌え」は「快さ」によってほぼ決定される



# 布置の違いについて

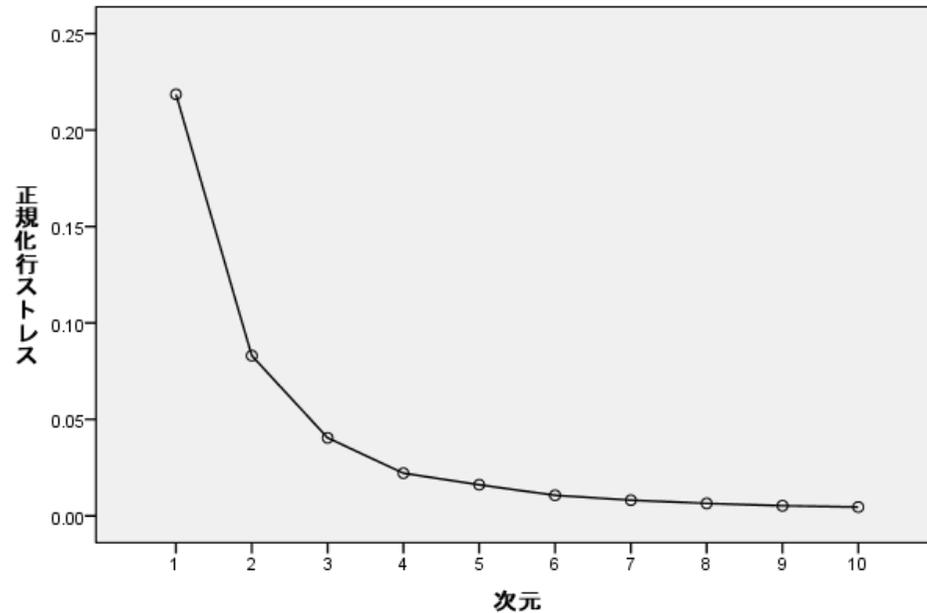
## ▶ 結果のデモ

- 声優毎に「快さ」が変化
  - 声優の「声質」が「快さ」に強く影響
  - 「萌える」声の持ち主はどんな口調でも「萌える」
  - 「萌えない」声の持ち主はどんな口調でも「萌えない」



# シチュエーション判断の分析

- ▶ 各項目の相関行列から距離行列
  - $1-r$
- ▶ MDS(多次元尺度構成法)を用いて分析
  - 2次元解を採用





# まとめ

- ▶ 音声の印象は「快さ」と「興奮度」で表現が可能
  - 声優の「声質」が「快さ」に強く影響
    - 「萌える」声の持ち主はどんな口調でも「萌える」
    - 「萌えない」声の持ち主はどんな口調でも「萌えない」
- ▶ 似た発話音を得られるシチュエーション設定のマップ
  - 声優への円滑なインストラクションが可能に



# 2次元デジタルフィルタを用いる 指向性制御方法の応用例

山川昭彦（金沢大）

伊藤栄太

齋藤毅

三好正人

# 2次元デジタルフィルタを用いる 指向性制御方法の応用例

金沢大学 オーディオ情報処理研究室  
伊藤 栄太, 山川 昭彦, 齋藤 毅, 三好 正人

# 目次

- ①マイクロホンアレーによる移動音声の高S/N收音について
- ②集束ビームを用いて生成した仮想音源の評価

# マイクロホンアレーによる移動音声の高S/N收音について

## 1.はじめに

- 講演者が移動しながら話をする時

空調雑音



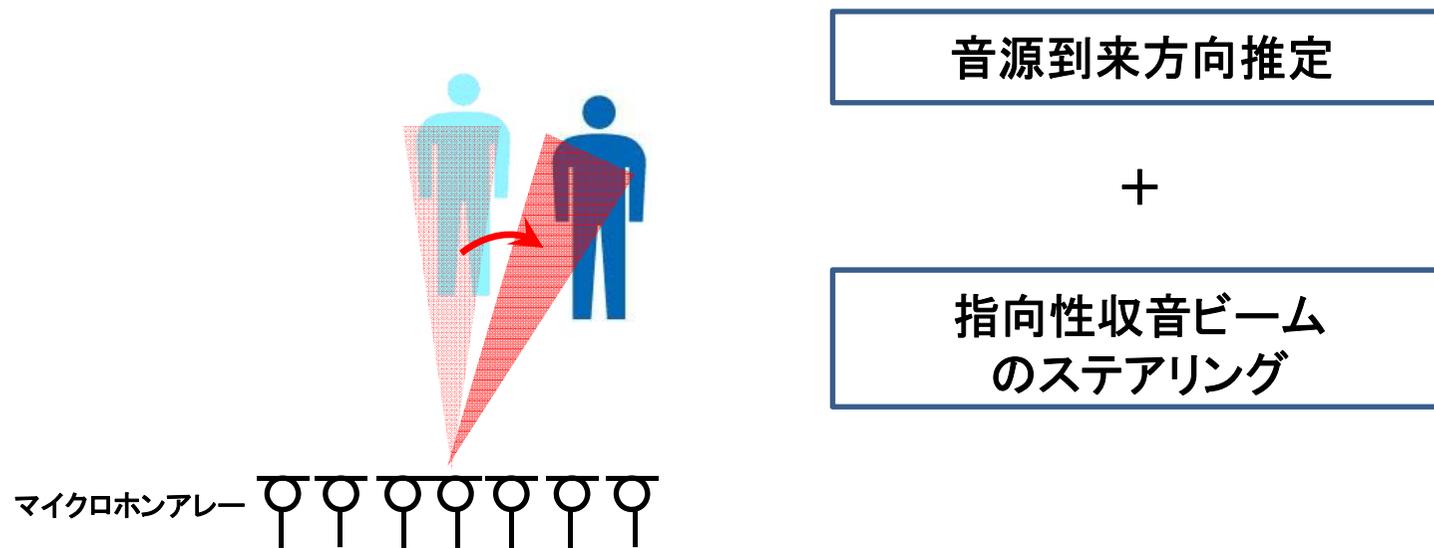
生徒のざわめき

会場内に設置した固定マイクロホンで  
講演者音声のみを收音したい

## 2.提案システム

### 2.1 システム原理

- 固定マイクロホンで、講演者音声のみを高S/N收音する



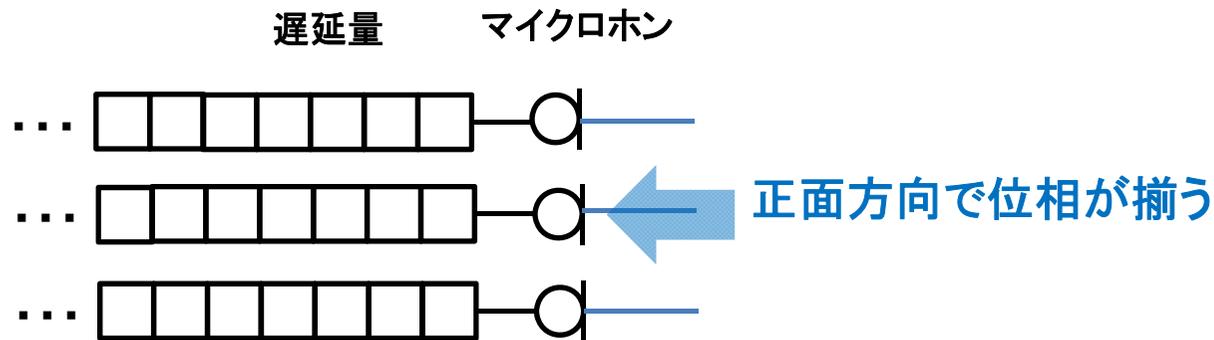
推定方向に指向性收音ビームをステアリングする

## 2.提案システム

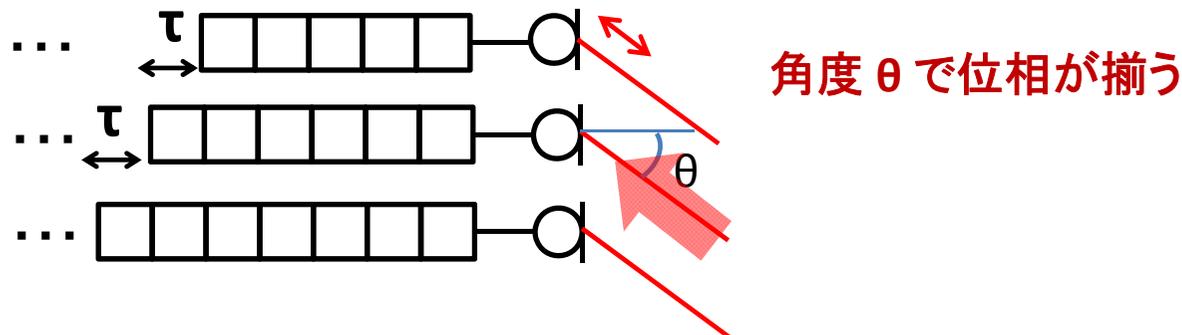
### 2.1 システム原理

- ビームステアリング

- ・全てのマイクにかける遅延量が同じ場合



- ・マイク毎の遅延量を変化させる場合

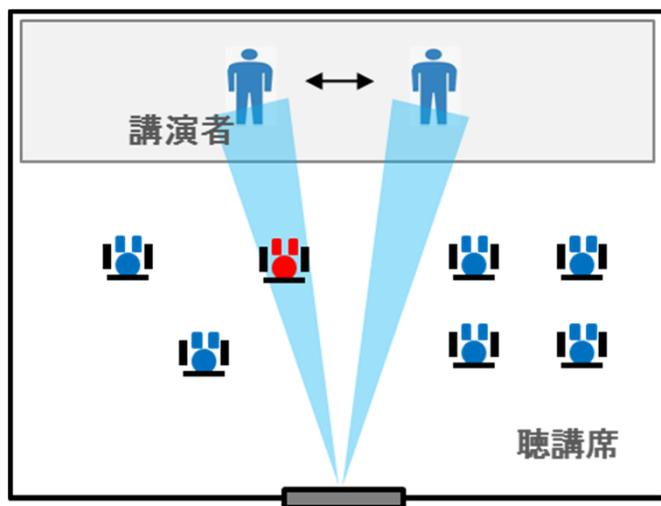


## 2.提案システム

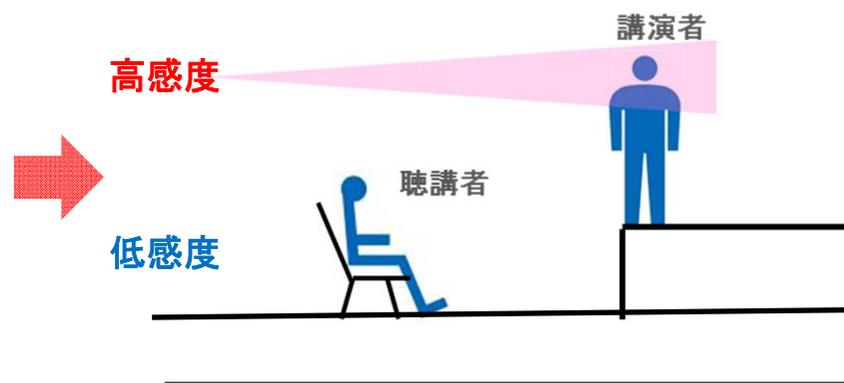
### 2.1 システム原理

- ビーム内に雑音を含んでしまう場合

左右方向が分かったとしても...



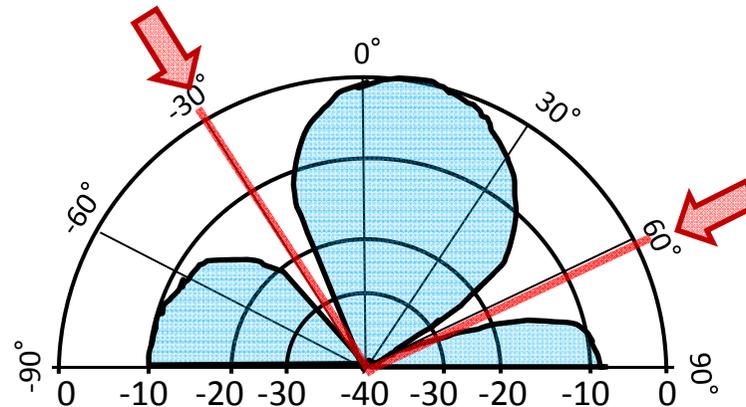
聴講者音声などが含まれる場合がある



音源の高さの違いを利用する

## 2.提案システム

### 2.2 ヌルビームによる方向推定

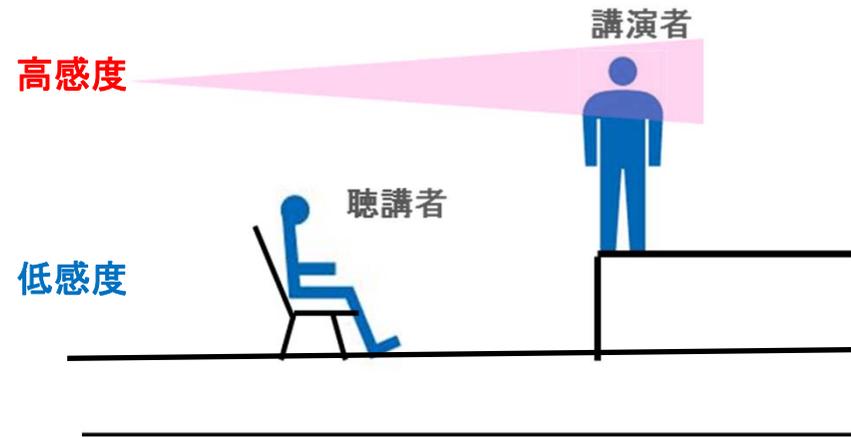


#### ヌルビームの問題点

- ①複数の方向のうち、どれが講演者の音声なのかを判断する基準が必要
- ②方向推定出来る音源数が「アレーの総素子数-1」個に限られる  
※以下、アレーの自由度と呼ぶ

## 2.提案システム

### 2.2 ヌルビームによる方向推定



#### 解決案

①複数の方向のうちどれが講演者の音声なのかを判断する基準が必要 ⇒パワーの最も大きい方向に定める

②方向推定出来る音源数が「アレーの総素子数-1」個に限られる

※自由度

⇒パワーの大きい到来方向を優先させる

## 2.提案システム

### 2.2 ヌルビームによる方向推定

<ヌルビームによる 到来方向推定法 >

超分解能

指導原理	素子数(K)と到来波数(L)	計算原理	特徴	備考
線形予測法	$K-1=L$	方向サーチ	予測誤差最小化	熱雑音の影響有
最小ノルム法	$K-1=L$	固有値, 方向サーチ	電力最小化	-
MUSIC法	$K-1 \geq L$	固有値, 方向サーチ	電力最小化	-
ESPRIT法	$K-1 \geq L$	固有値	サブアレー間位相差	アレイ形状は同形

音源数が自由度を越えてもパワーの大きい方向を優先させる, **最小ノルム法**を採用

## 2.提案システム

### 2.3 最小ノルム法の原理

- アレーの出力電力を最小化する方向を求める  
⇒それぞれのマイクの出力信号間で相関行列を作成し、その行列の固有値・固有ベクトルを利用する

#### 最小ノルム法の計算式

$$\min(P_{out} = \frac{1}{2} \mathbf{W}^H R_{xx} \mathbf{W})$$

subject to  $\mathbf{W}^H \mathbf{W} = 1$

$$P_{MN}(\theta) = \frac{1}{|\mathbf{W}_{MN}^H a(\theta)|^2}$$

( $\mathbf{W}_{MN}$ :  $R_{xx}$ の最小固有値の固有ベクトル)

$P_{MN}(\theta)$  が高い値を取る $\theta$ が到来方向となる

$P_{out}$  :アレーの出力電力

$R_{xx}$  :マイク間の相関行列

$\mathbf{W}$  :素子ウェイト

## 2.提案システム

### 2.3 最小ノルム法の原理

Lagrangeの未定係数法により置き換える

$$Q(\mathbf{W}) = \frac{1}{2} \mathbf{W}^H \mathbf{R}_{xx} \mathbf{W} + \frac{\lambda}{2} (1 - \mathbf{W}^H \mathbf{R}_{xx} \mathbf{W})$$

$\lambda$ : 未定係数



$$\nabla_{\mathbf{W}} Q(\mathbf{W}) = \mathbf{R}_{xx} \mathbf{W} - \lambda \mathbf{W} = 0$$



$$\therefore \underline{\mathbf{R}_{xx} \mathbf{W} = \lambda \mathbf{W}}$$

固有値問題

## 2.提案システム

### 2.3 最小ノルム法の原理

$R_{xx} \mathbf{W} = \lambda \mathbf{W}$  の両辺に  $\mathbf{W}^H$  を掛けると

$$\mathbf{W}^H R_{xx} \mathbf{W} = \lambda \mathbf{W}^H \mathbf{W} = \lambda$$

$(P_{out} = \frac{1}{2} \mathbf{W}^H R_{xx} \mathbf{W})$  より,  $\lambda$ は出力電力の2倍に相当することが分かる

最小の $\lambda$ に対応する固有ベクトル $\mathbf{W}$ を用いて,以下の式を得る

$$P_{MN}(\theta) = \frac{1}{|\mathbf{W}_{MN}^H a(\theta)|^2}$$

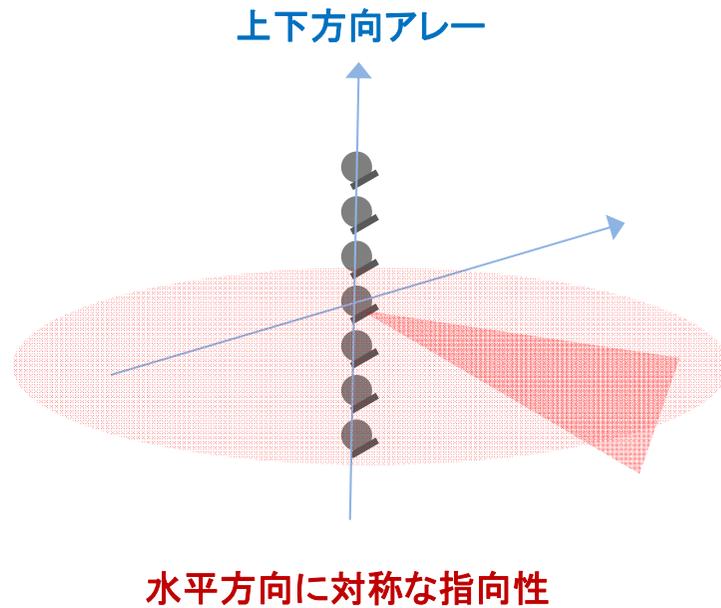
( $\mathbf{W}_{MN}$  :  $R_{xx}$ の最小固有値の固有ベクトル)

$P_{MN}(\theta)$  が高い値を取る $\theta$ が到来方向となる

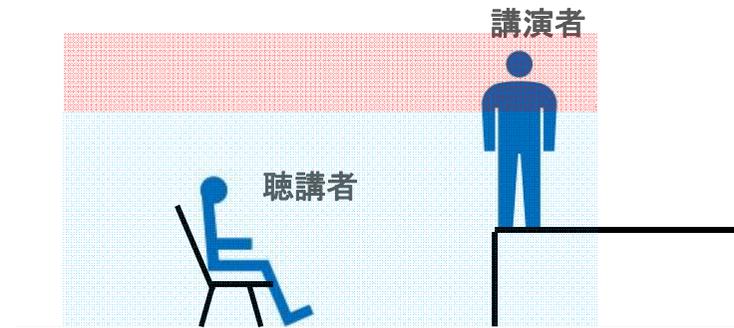
## 2.提案システム

### 2.3 指向性制御

- 講演者音声の感度が高い指向性の実現



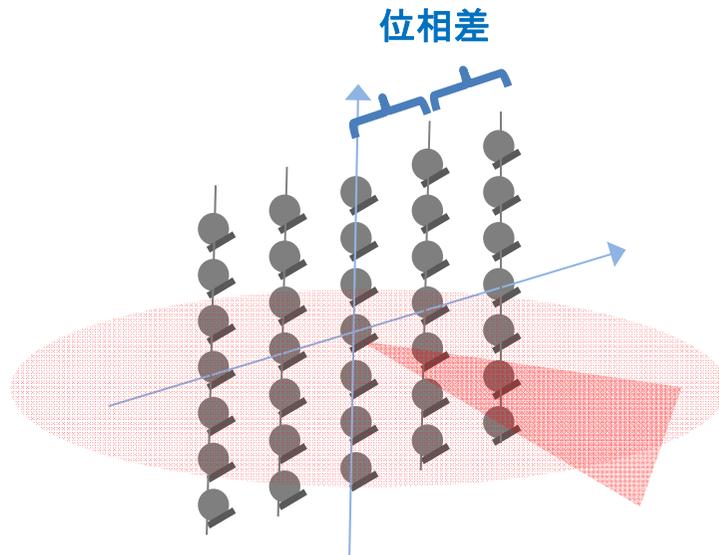
1つの高さ方向アレーを【1ユニット】とする



## 2.提案システム

### 2.3 指向性制御

- 話者の左右方向移動での方向推定



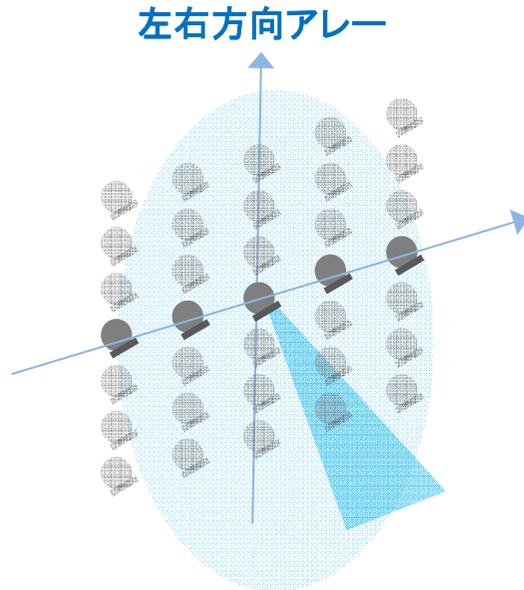
水平に並べた複数の高さユニット間で  
左右方向の方向推定

## 2.提案システム

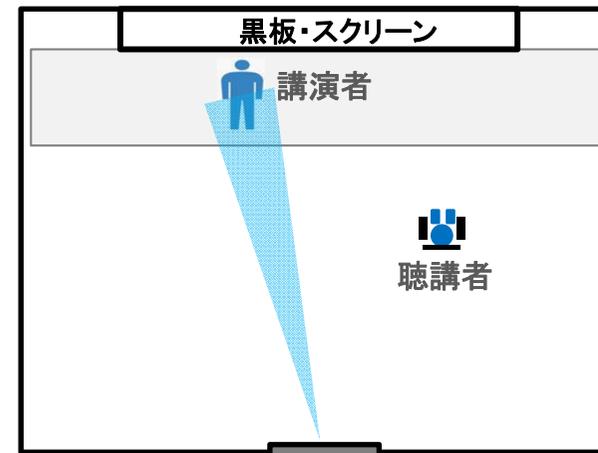
### 2.3 指向性制御

- 左右方向に指向性收音ビームを向ける

高さアレーの中心素子を代表させて...



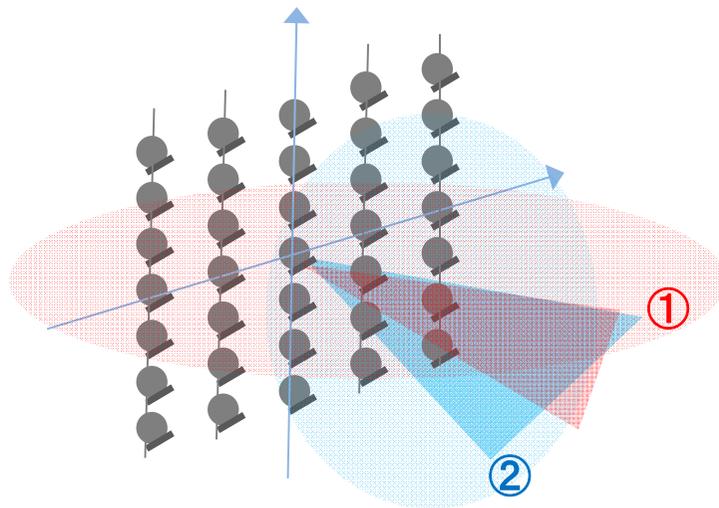
鉛直方向に対称な指向性



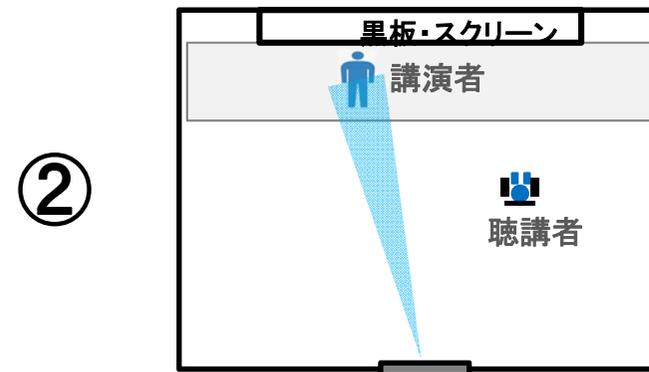
## 2.提案システム

### 2.3 指向性制御

- 推定された話者方向に指向性收音ビームを向ける



水平・鉛直指向性の積となり  
スポットライト型ビームで講演者音声が強調される

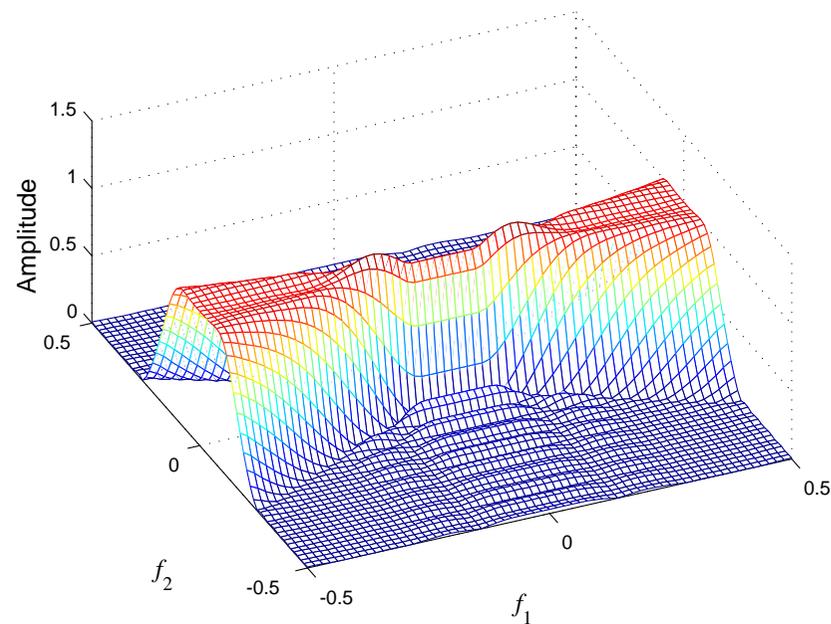


## 2.提案システム

### 2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$



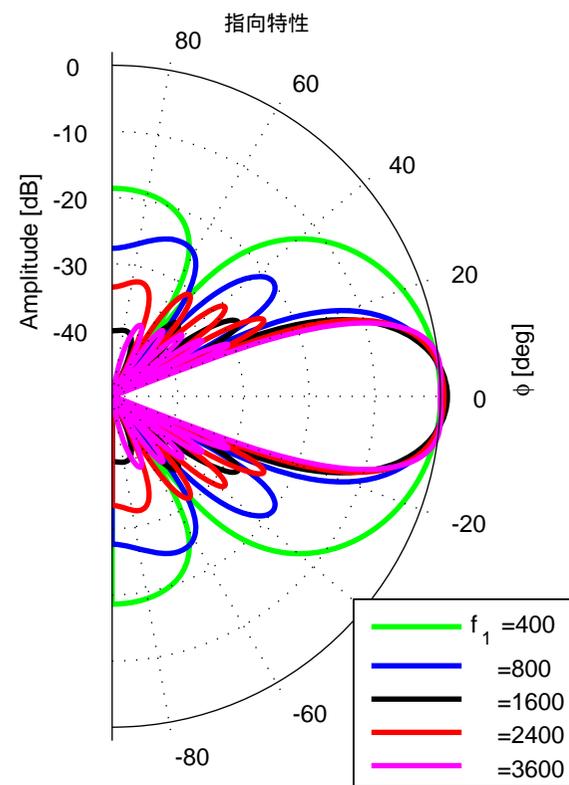
< 振幅特性 >

## 2.提案システム

### 2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$



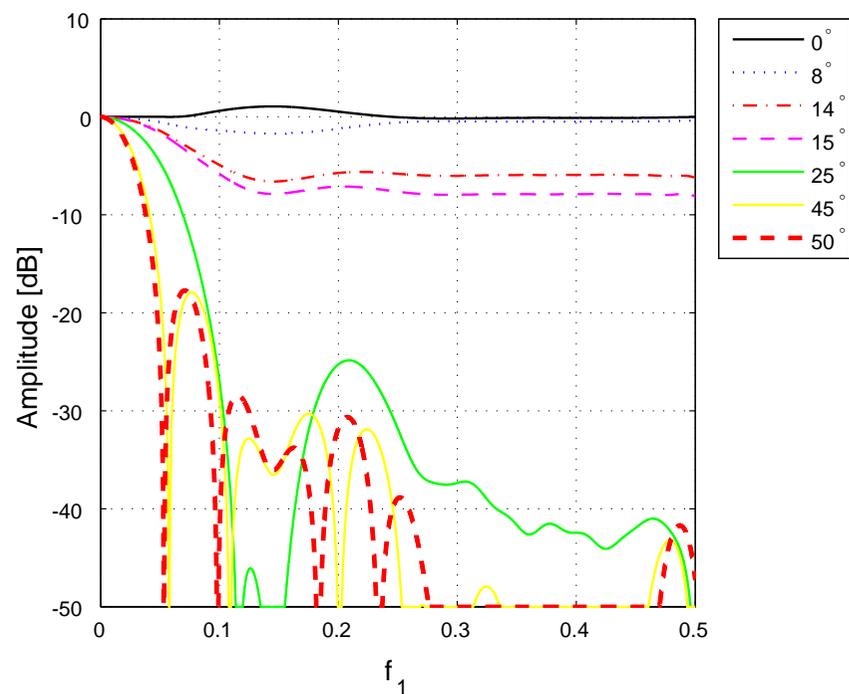
< 指向特性 >

## 2.提案システム

### 2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$

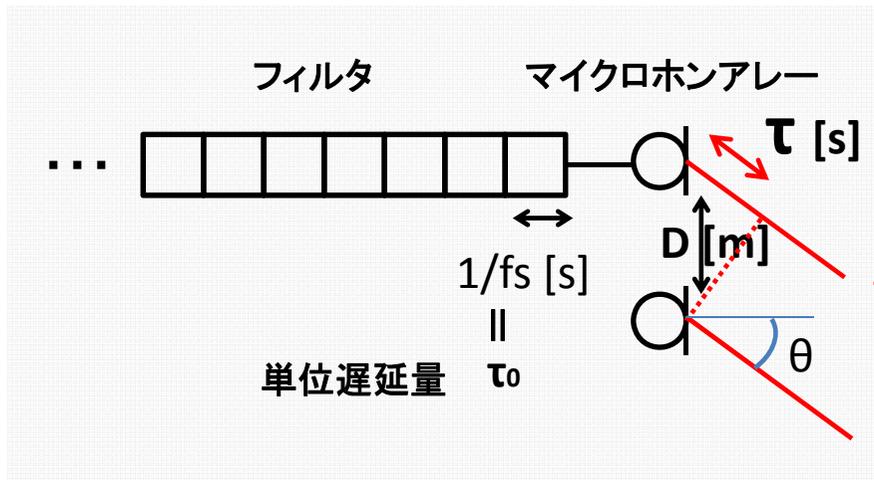


< 周波数特性 >

# 3. シミュレーション

## 3.1 条件定義

### ・受信時のサンプリング周波数の決定



隣り合うマイク間の時間差:

$$\tau = \frac{D \cdot \sin \theta}{c}$$

↓ 単位遅延量に合わせると

$$\text{測定できる角度} : \theta = \sin^{-1} \left( \frac{c \cdot (\tau_0 \cdot n)}{D} \right)$$

< 測定可能角度 >

例えば	サンプリング周波数	40[kHz]	のとき...	単位遅延数n	1	2	3	4
	間隔D	4.25[cm]		θ	11.5	23.6	36.9	53.2

# 3.シミュレーション

## 3.1 条件定義

### ・受信時のサンプリング周波数の決定

今回は10° 刻みの移動を追う

< 推定可能な角度分解能 >

サンプリング周波数	間隔D	単位遅延数	1	2	3	4	5	6	7
64[kHz]	4.25[cm]	対応角度	7	14	22	30	39	49	61

サンプリング周波数	64 [kHz]
間隔d	4.25 [cm]

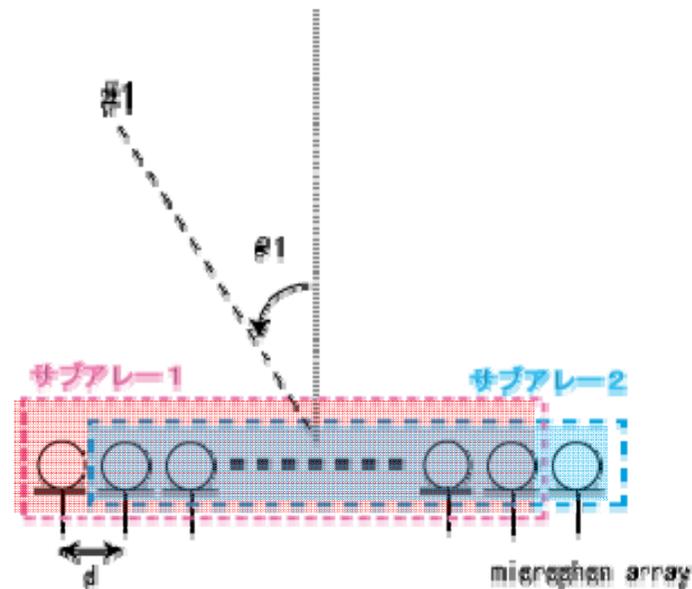
を採用する



# 3. シミュレーション

## 3.1 条件定義

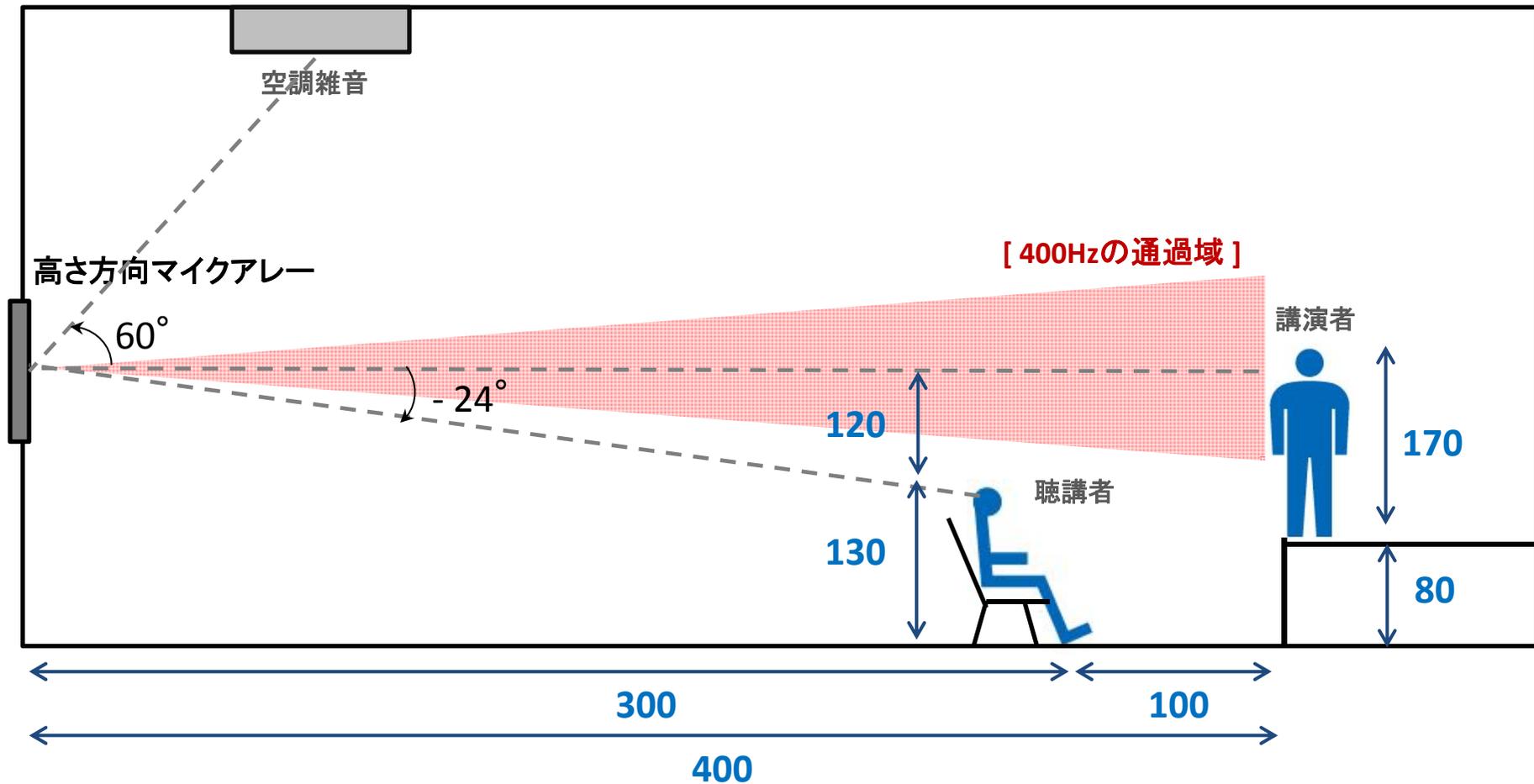
- ESPRIT法・・・2つのサブアレー間の位相差を固有値計算により求めて方向を得る



- 推定精度が非常に高い
- × 音源数が自由度を越えると推定できなくなる

# 3.シミュレーション

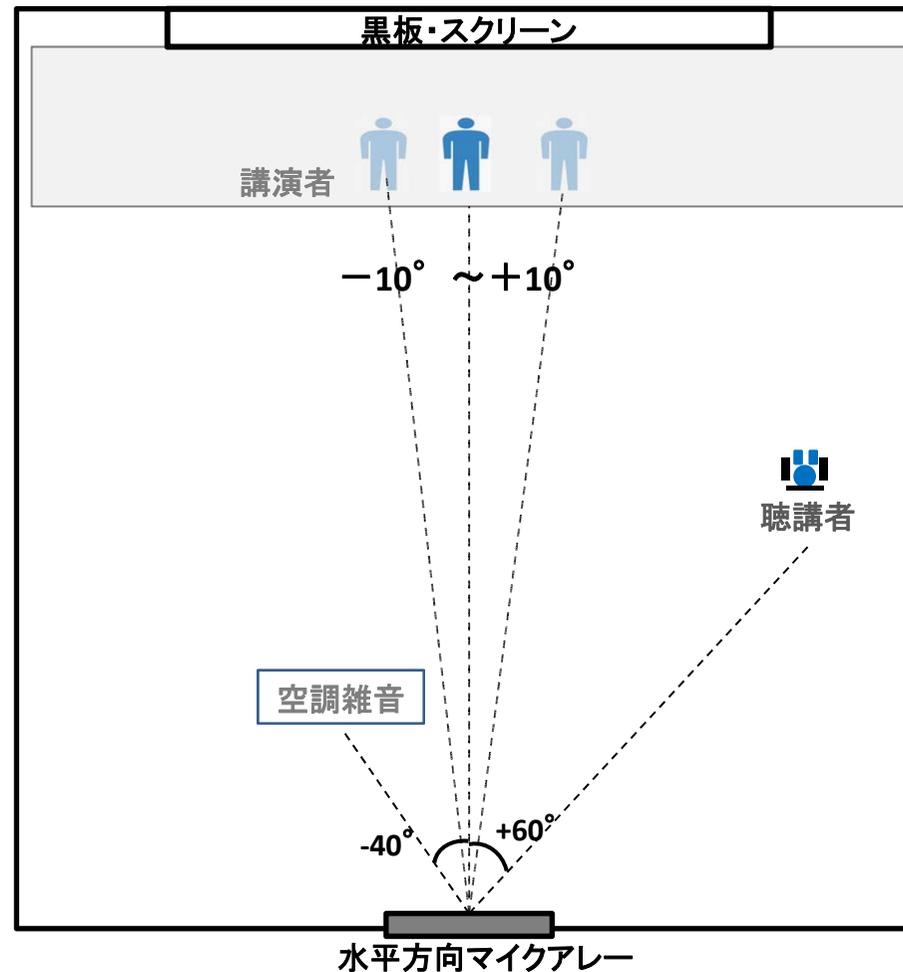
## 3.2 シミュレーション方法



< 高さ方向で見た部屋の空間配置 >

# 3. シミュレーション

## 3.2 シミュレーション方法

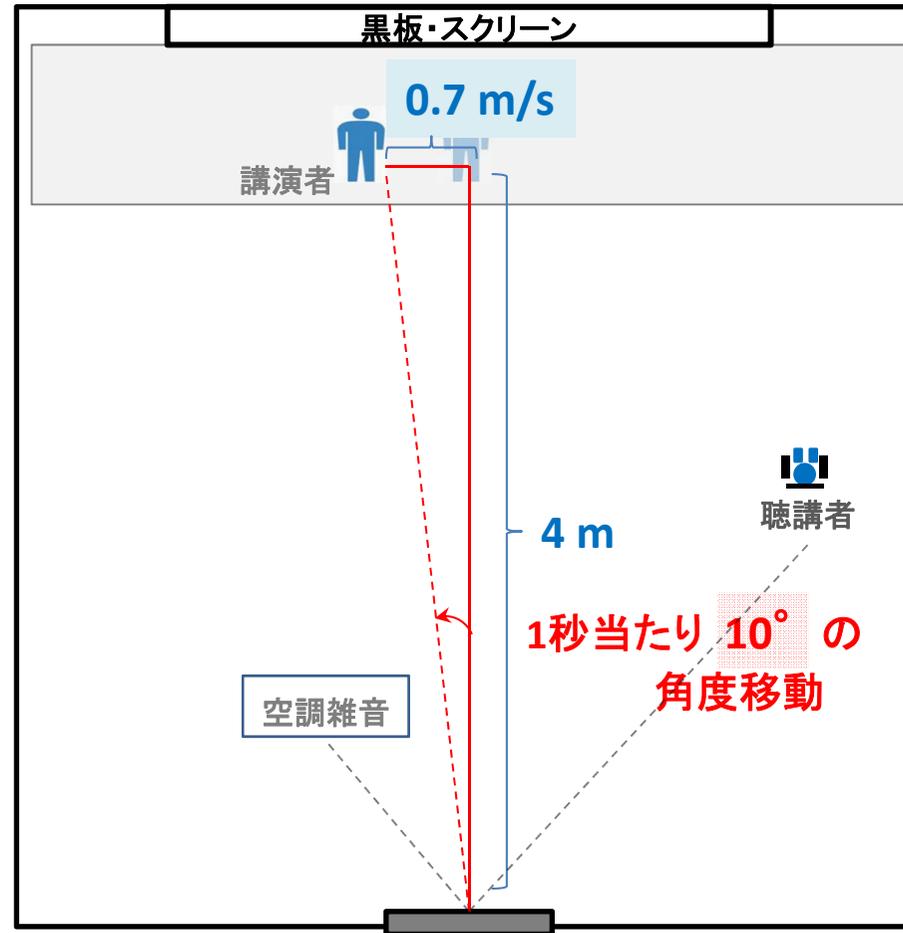


< 水平方向で見た部屋の空間配置 >

# 3. シミュレーション

## 3.2 シミュレーション方法

講演者の移動速度を  
0.7 m/s とすると...

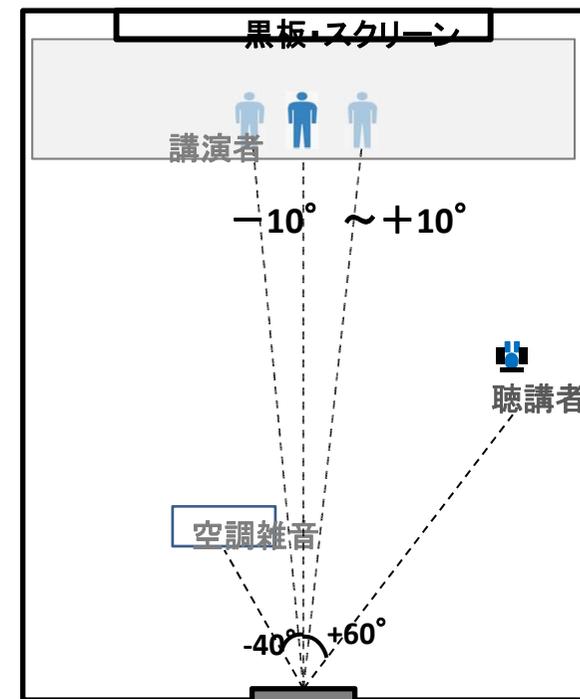
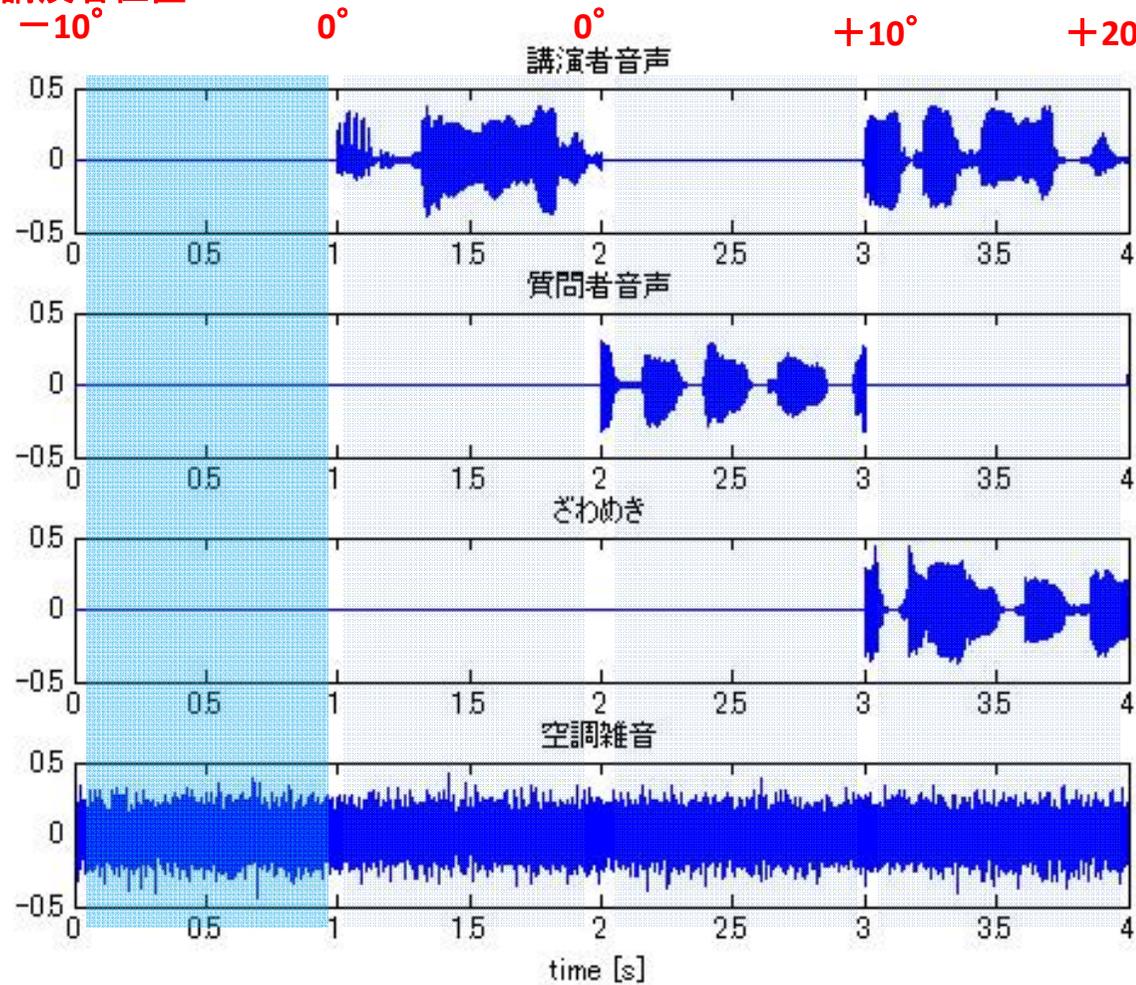


< 水平方向で見た部屋の空間配置 >

# 3.シミュレーション

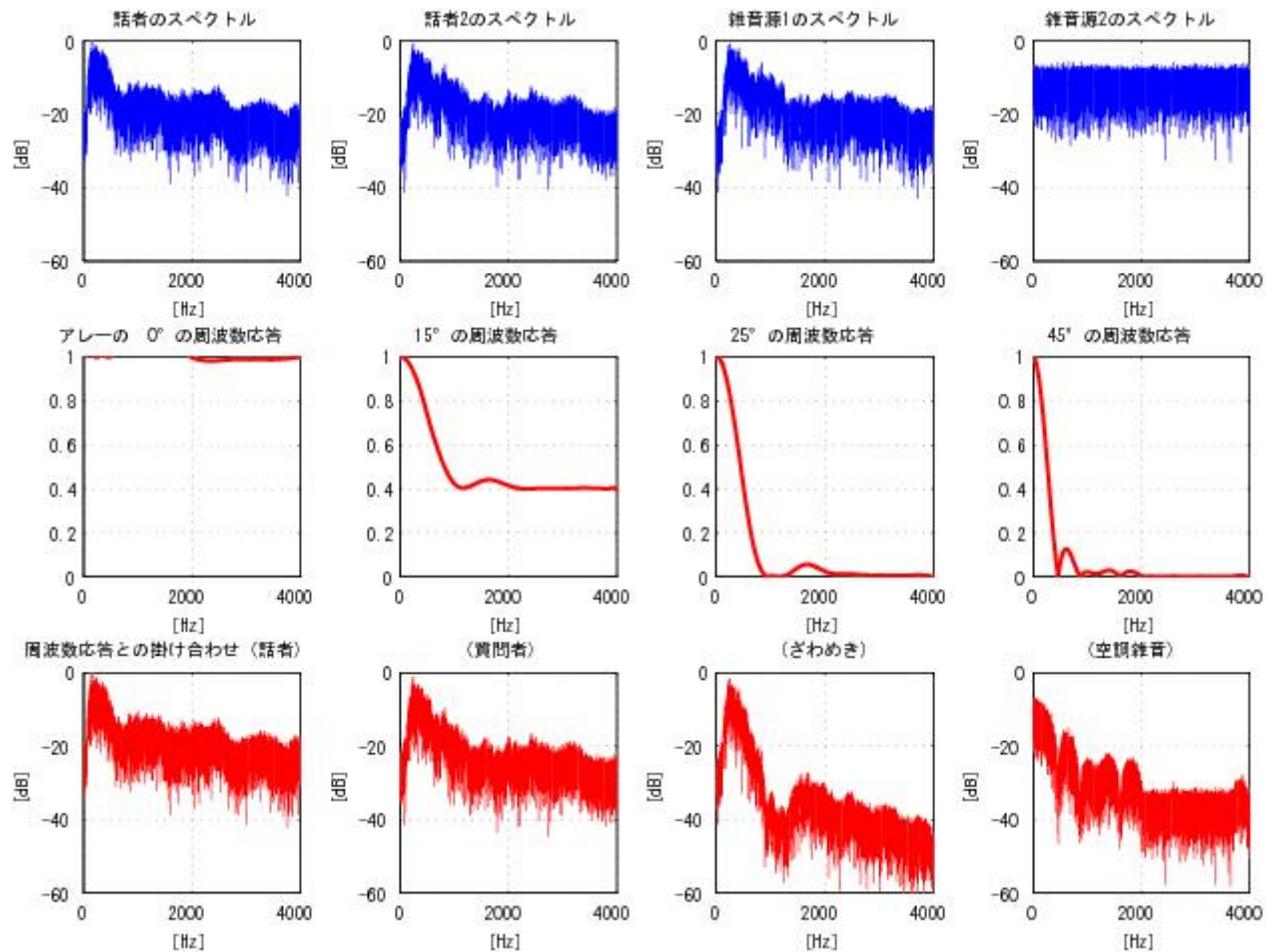
## 3.2 シミュレーション方法

講演者位置  
-10°



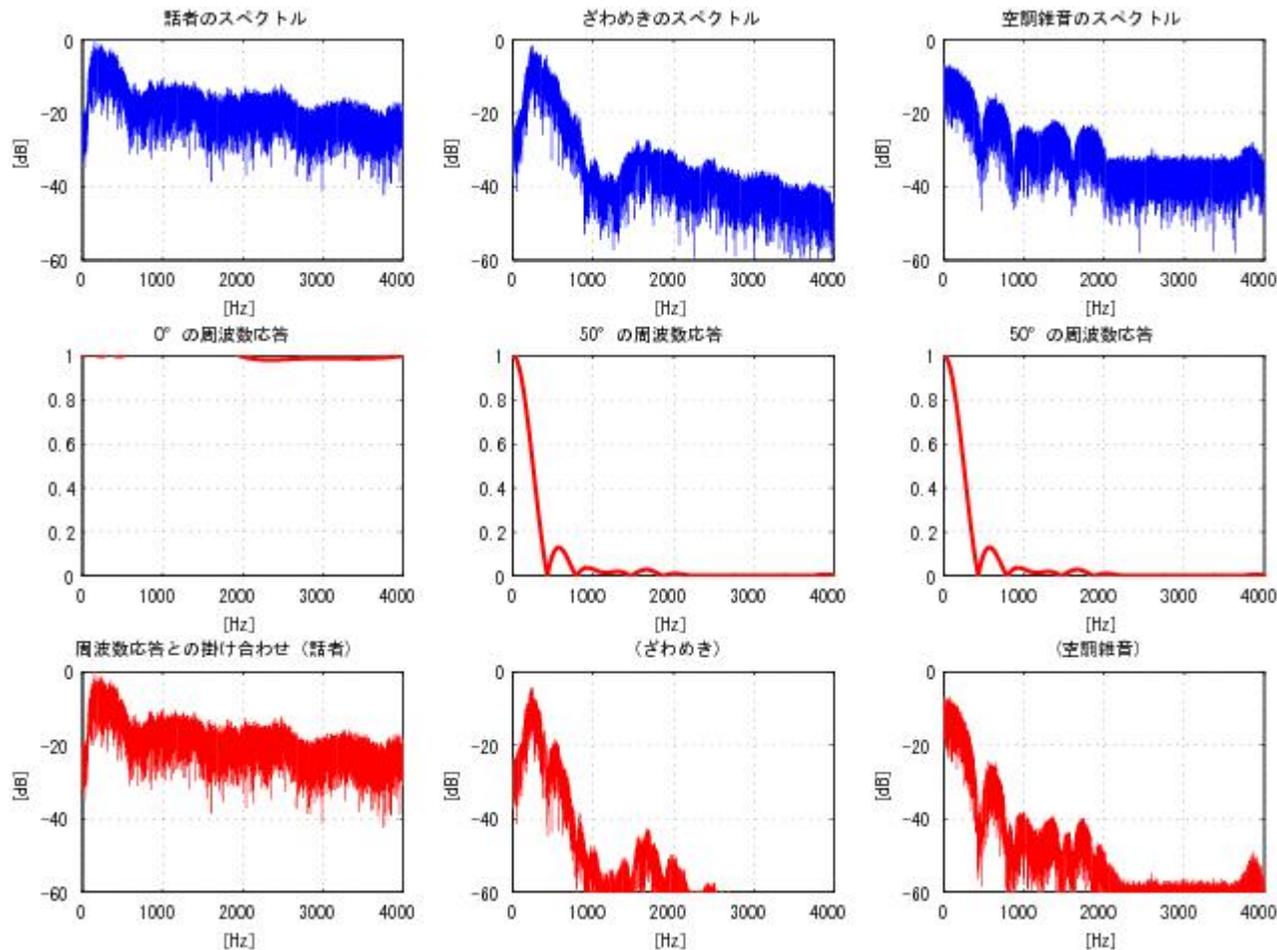
# 3. シミュレーション

## 3.3 シミュレーション結果



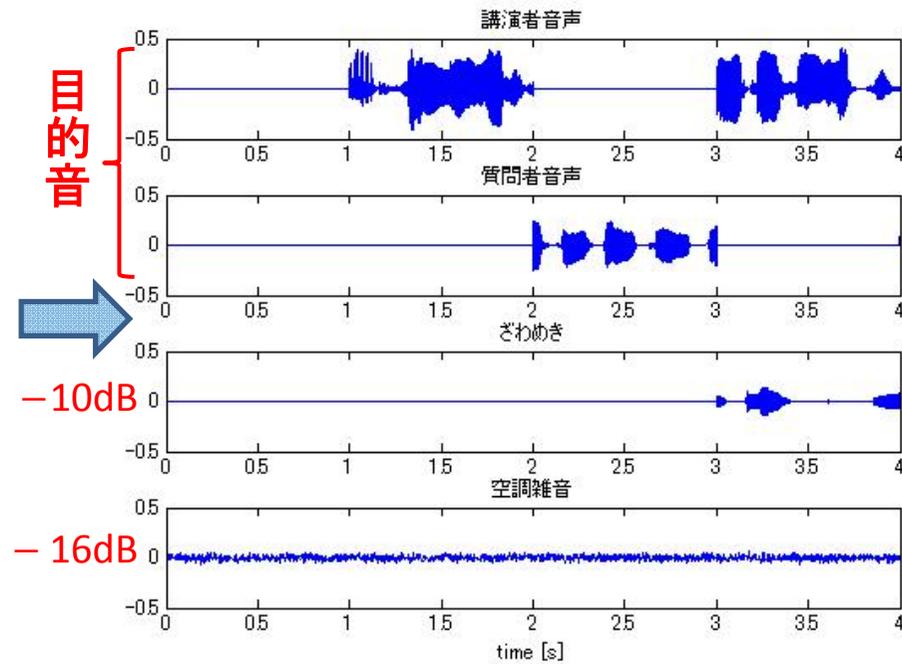
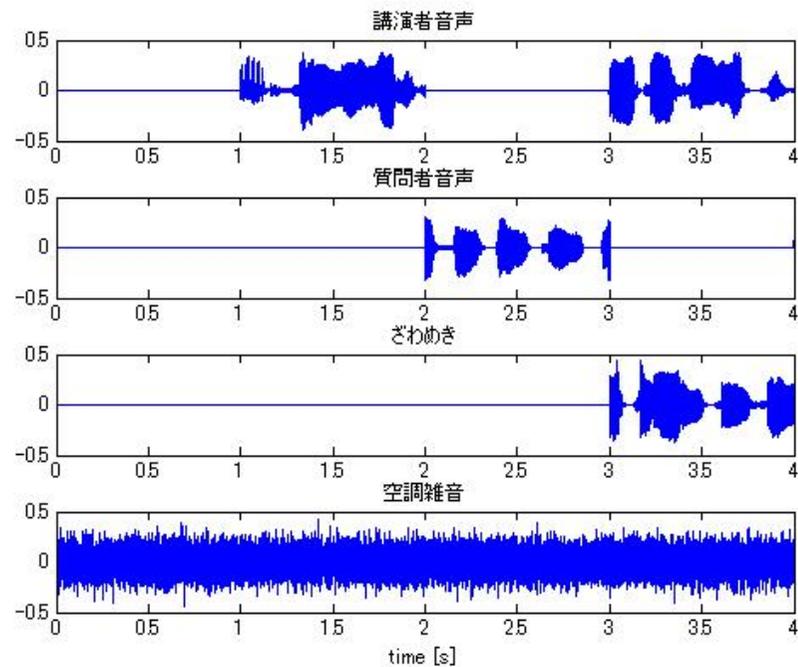
# 3. シミュレーション

## 3.3 シミュレーション結果



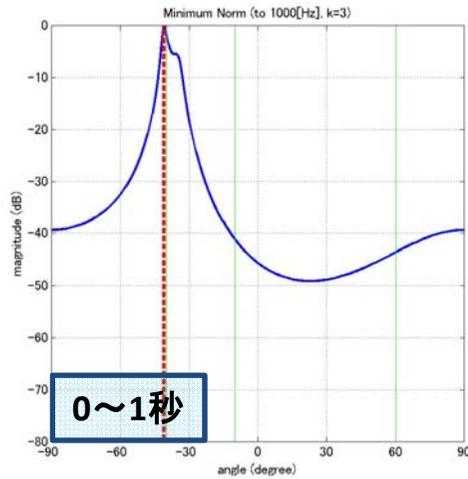
# 3. シミュレーション

## 3.3 シミュレーション結果



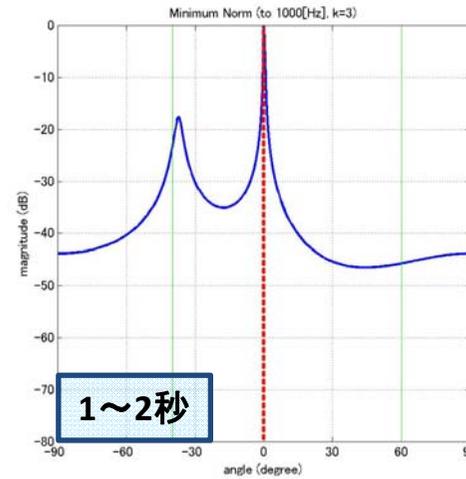
# 3. シミュレーション

## 3.3 シミュレーション結果



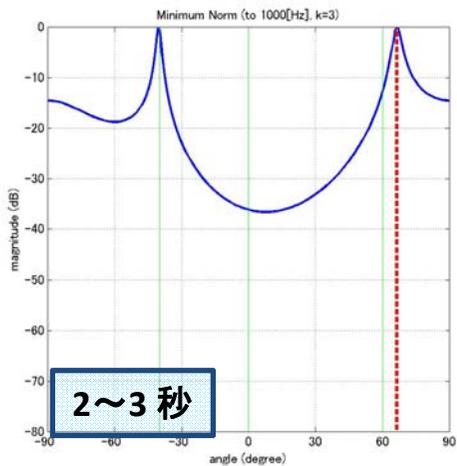
ESPRIT

$-41.7^\circ$   
0.9



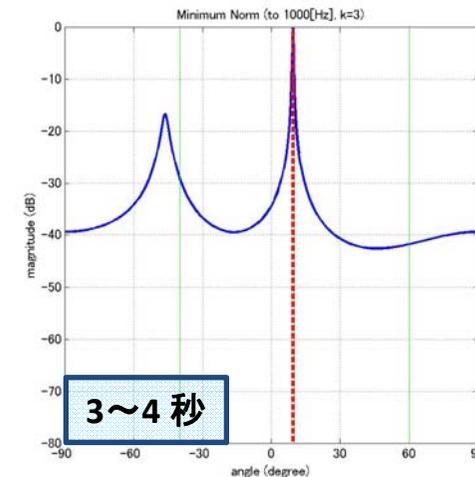
ESPRIT

$0.1^\circ$   
0.9  
 $-37.0^\circ$   
0.1



ESPRIT

$66.4^\circ$   
0.7  
 $-40.5^\circ$   
0.1



ESPRIT

$16.1^\circ$   
1.1  
 $35.1^\circ$   
0.2  
 $-38.7^\circ$   
0.1

# 3.シミュレーション

## 3.3 シミュレーション結果

<各周波数での方向推定平均結果>  
(500, 600, 800, 1000, 1200, 1600, 2000, 2500, 3000[Hz])

(各時刻で最もパワーの大きい音源)

空調:-40°    講演者:0°    質問者:60°    講演者:10°

時間[s]	0~1	1~2	2~3	3~4
最小ノルム	-40.6°	0.2°	64.5°	10.2°
ESPRIT	-40.6°	0.1°	64.3°	-0.8°

音源数がアレーの自由度を越えるときでも、  
最小ノルム法では講演者音声の方向を高精度に推定できる

## 4.むすび

- 講演者音声のみを高S/N收音する
  - 指向性收音ビーム + 音源到来方向推定法

{シミュレーション実験}

1. 水平面の指向性形成
  2. 講演者音声のパワーを用いる方向推定
  3. 指向性收音ビームを講演者方向へ
- 今後の課題
    - 上下方向アレーの小型化
    - より実際的なシミュレーション

# スピーカ系について

## - 集束ビームを用いて生成した仮想音源の評価 -

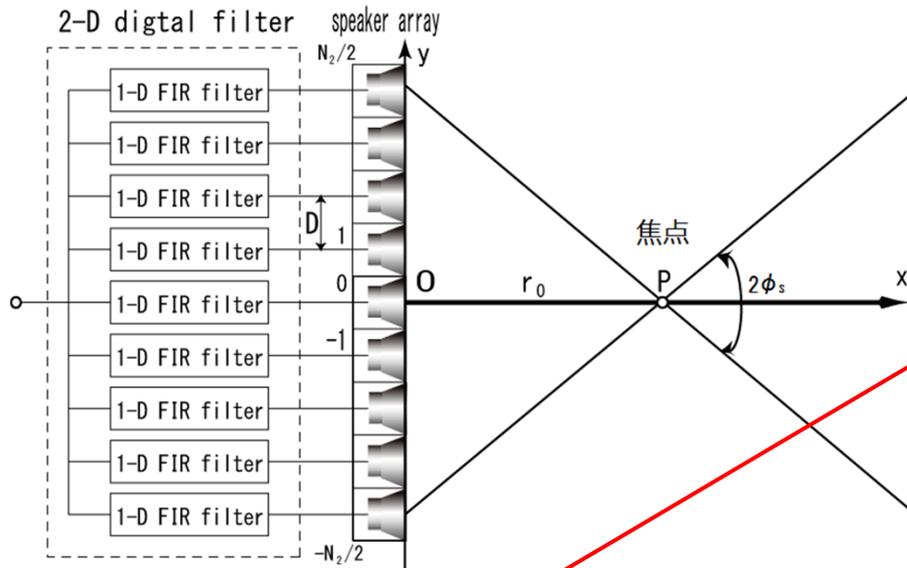
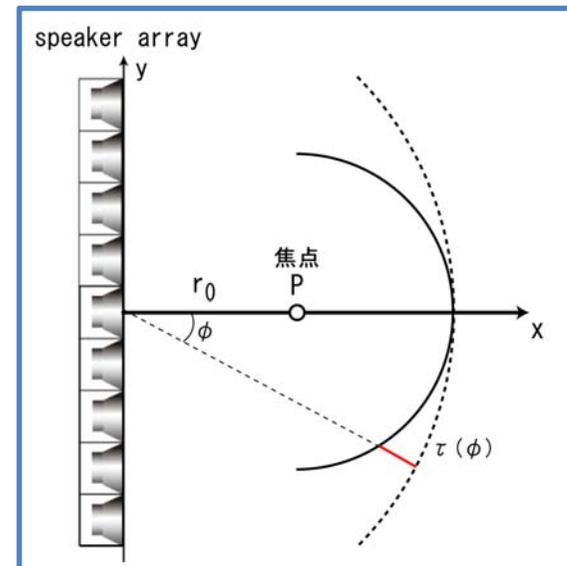
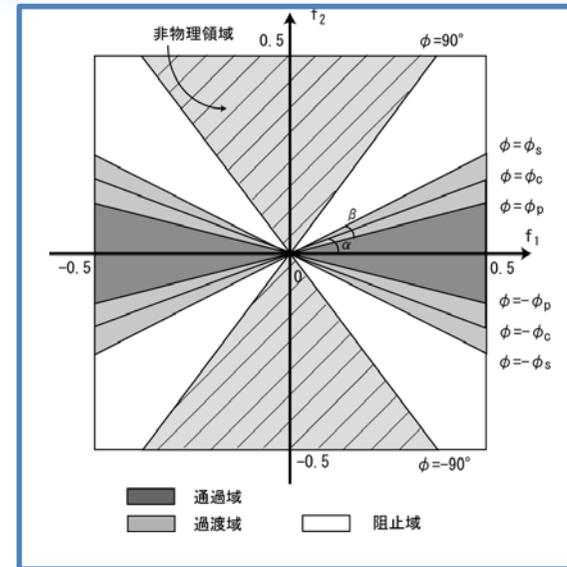


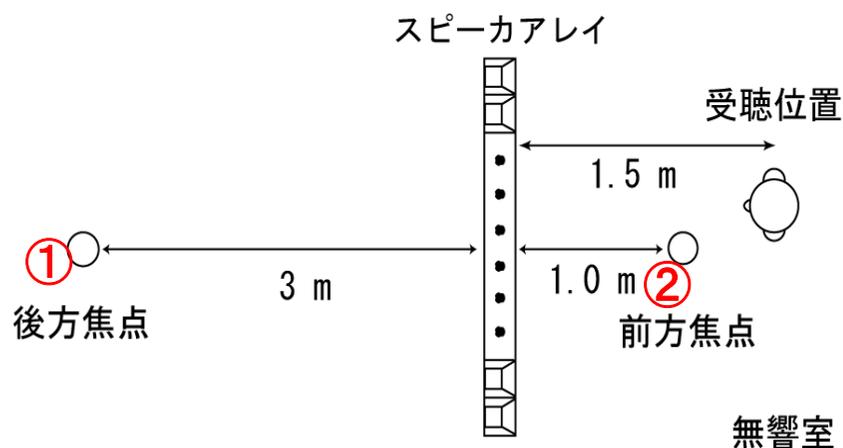
図 高音圧焦点の生成方法

振幅特性と位相特性の  
制御により高音圧焦点を生成



# 音像制御

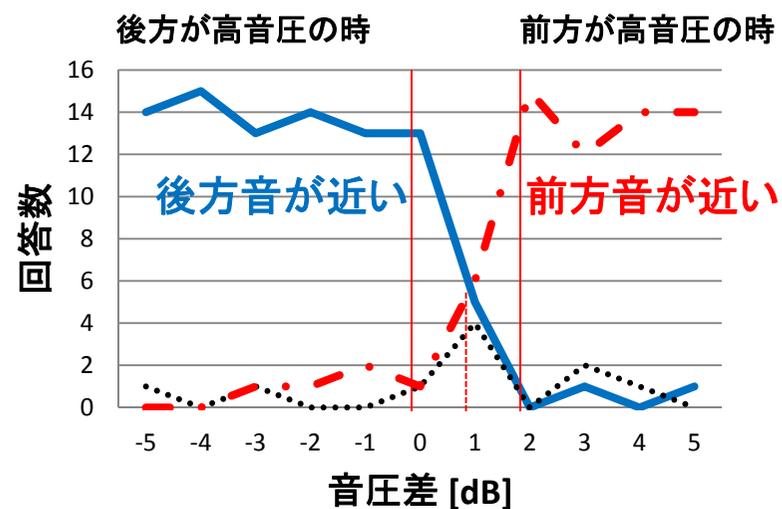
## • 音圧差と距離感



### 聴取実験

- ・後方焦点から音を鳴らした後、前方焦点から音を鳴らす
- ・後方と前方の音圧差を変えていき、聴取者に音の遠近感を判断させる

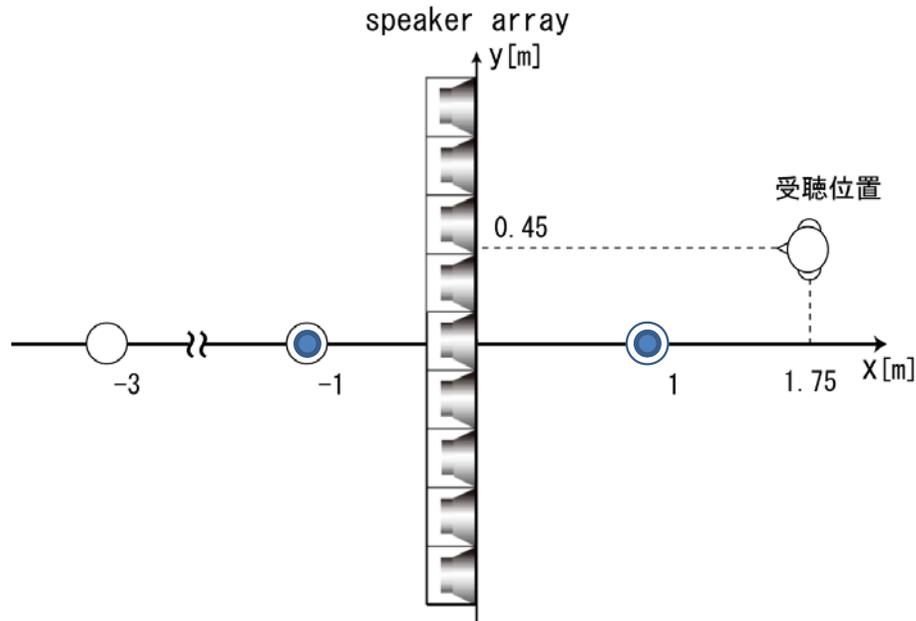
### <回答結果>



遠近感 は 音圧差 1~2dB に依存

# 音像制御

## ● 波面と方向感・距離感

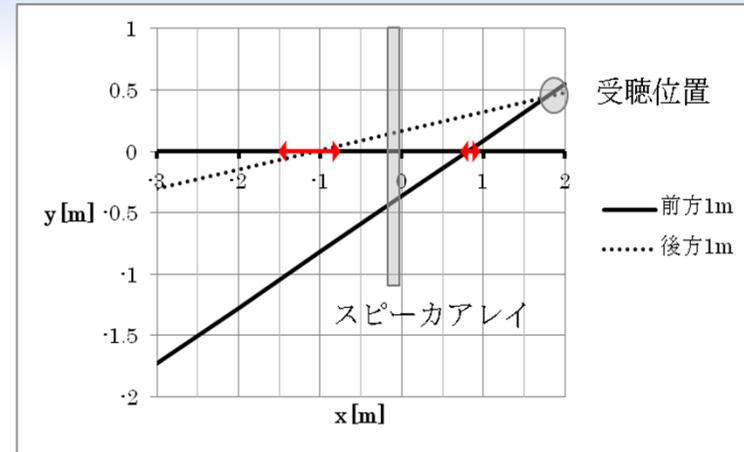


フィルタ次数 = (60, 14)

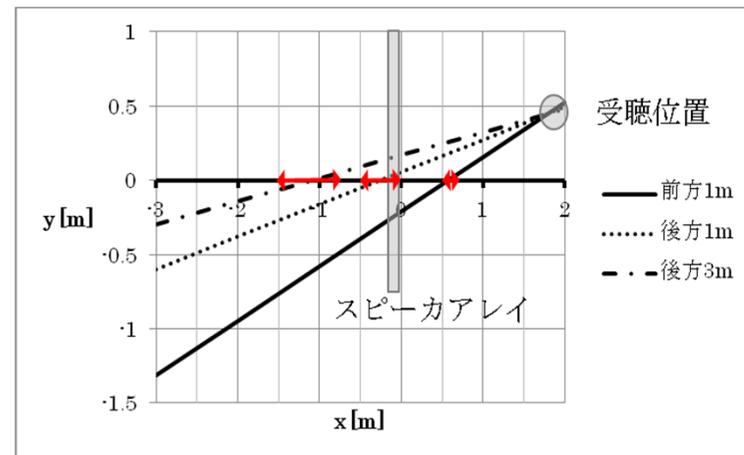
スピーカ間隔 = 14cm (スピーカ長 : 約2m)

信号帯域 = 2kHz

試験音 : 男声朗読音声「かえでいろづく」



実音源



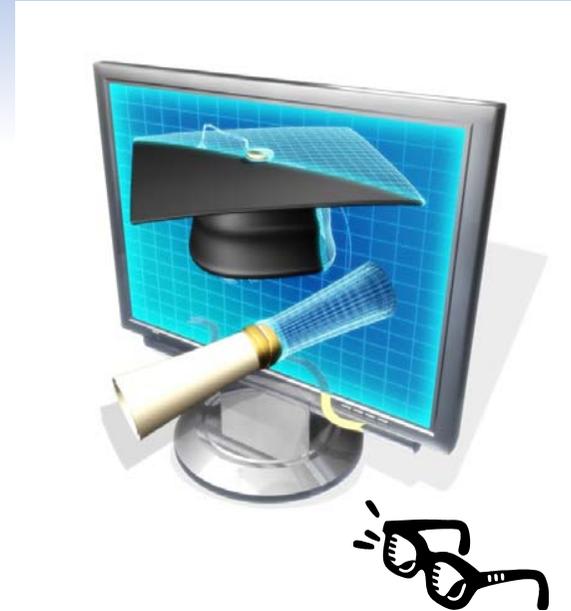
仮想音源

# まえがき

近年、3D映像技術が発達している



奥行きはきわめてリアルに表現されている



音響の遠近感に関する研究では、

多数のスピーカを用いる**波面合成**や**焦点形成**が検討されている

我々は、

比較的小規模のスピーカアレイを用いた  
焦点生成による音像制御方法の検討

# まえがき

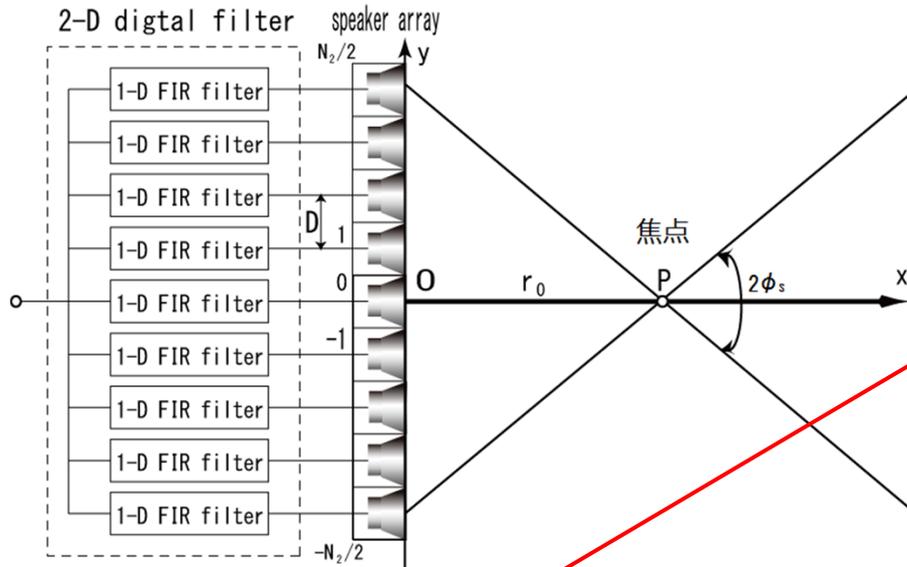
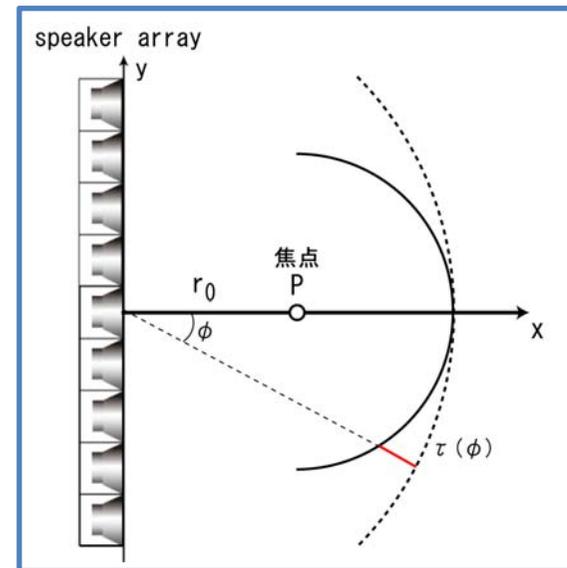
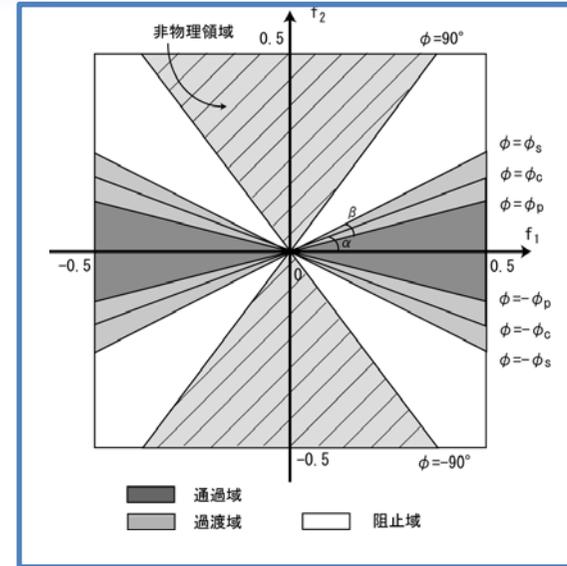


図 高音圧焦点の生成方法

振幅特性と位相特性の  
制御により高音圧焦点を生成



# まえがき

- 音源位置が異なった場合

- 音の大きさ



- 受聴位置での音圧差

- 波面の違い



- 受聴位置付近の位相差

- 壁による反射



- 直接音残響音比

スピーカアレイの前・後方に生成した仮想音源に関する聴取実験を行い、遠近感を評価する

# 遠近感について

- **音圧差**

遠近感に最も影響を与え、1~2 dB程度の差があれば十分影響を与える

- **波面**

実音源...低周波数の波面が遠近感に影響を与える

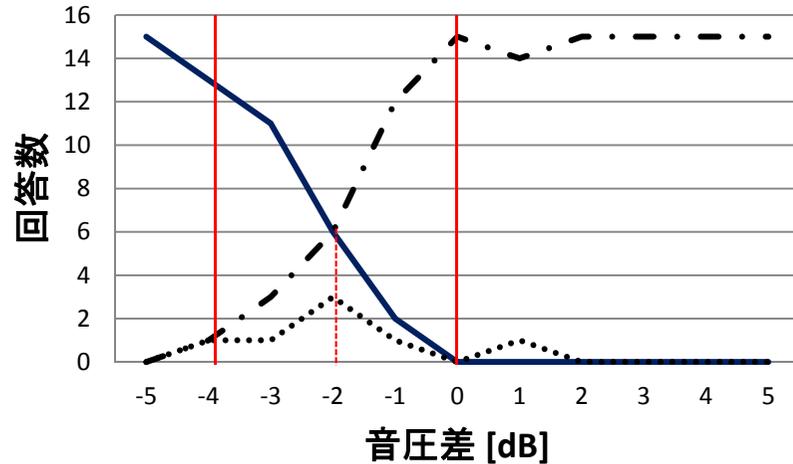
集束ビーム...低周波数の波面が形成出来ていないため検討は不十分

- **壁による反射**

遠近感に影響を与えるが、音圧差ほどの効果はない

# 音圧差について

前方焦点音→後方焦点音

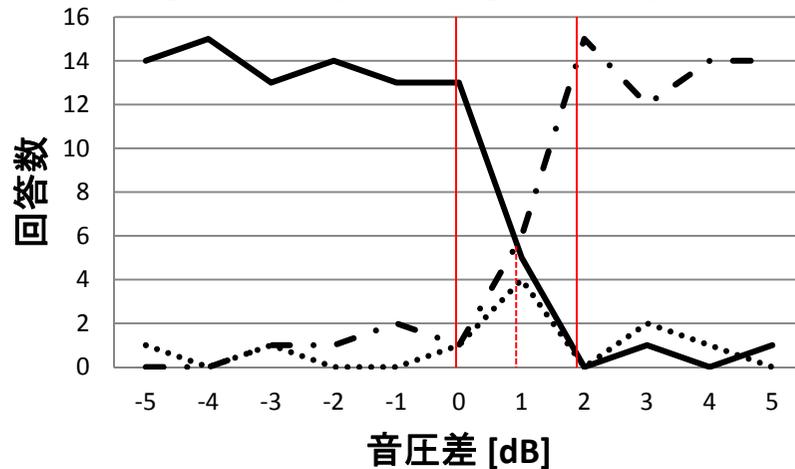


— 後続音の方が遠い  
..... 遠近判別出来ない  
- · - 後続音の方が近い

遠近判定の境目: -2 dB

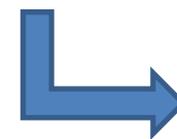
回答が偏るまでの音圧差の幅: 約2 dB

後方焦点音→前方焦点音



遠近判定の境目: 1 dB

回答が偏るまでの音圧差の幅: 1 dB



遠近判断は  
音圧差に依存  
(1~2 dB)

# 遠近感について

- 音圧差

遠近感に最も影響を与え、1~2 dB程度の差があれば十分影響を与える

- 波面

実音源...低周波数の波面が遠近感に影響を与える

集束ビーム...低周波数の波面が形成出来ていないため検討は不十分

- 壁による反射

遠近感に影響を与えるが、音圧差ほどの効果はない

# 波面について

聴取実験を行った結果

実音源 ⇒ ○: **正確な遠近感**が得られる  
集束ビーム ⇒ ×: 遠近感は得られなかった

先行研究より... **低周波数の波面**が聴感に影響を与えやすい



低周波数の波面を調べると...

実音源 ⇒ ○: 波面が**形成されている**  
集束ビーム ⇒ ×: 狙いどおりの波面が**形成されていない**

本研究についても

**低周波数の波面が遠近感に影響を与えることが示された**

➡ 集束ビームについて波面を形成する必要がある

※ **アレイ長・スピーカ間隔**を検討

# まとめ・今後の展望

- 集束ビームにより生成した仮想音源の遠近感に関する評価を行った
  - 遠近感は受聴位置での音圧差が最も影響を与える(1~2 dB)

不足点 : 低周波数における焦点の形成精度

今後は,

- 特に低周波数において高精度に波面を形成し, 波面の効果について再検討
  - アレイ長・スピーカ間隔を慎重に検討



**ご清聴ありがとうございました**