

日本音響学会北陸支部
第3回音響・超音波ジョイント研究会

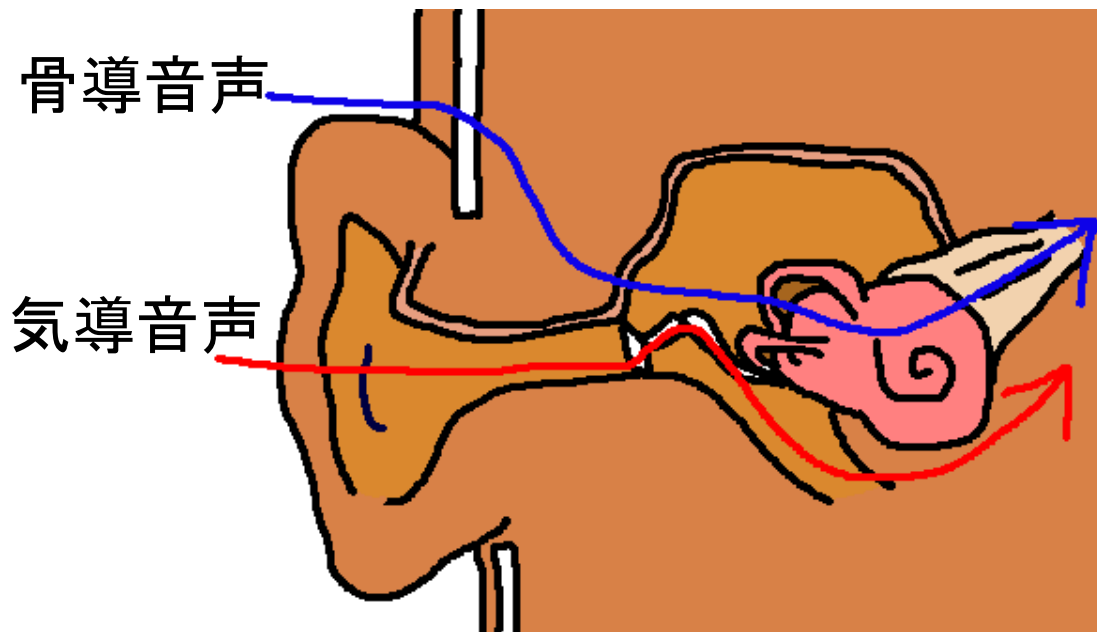
福井大学情報・メディア工学専攻音声研究グループの研究紹介

自己聴取音に占める気導音声・ 骨導音声の割合の推定に関する研究

福井大学

荻原 慎洋・森 幹男・谷口 秀次

骨導音声



気導音声：空気を介して
伝わる音声

骨導音声：発声者の骨を
介して伝わる音声

発声者は自身の声を気導音声＋骨導音声として無意識に認識している

発声者は機器類で録音された自身の声に聴感上の違和感を感じる

研究目的

- ▶ 発声時、発声者本人に聞こえている声の気導音声と骨導音声の割合を聴取実験によって明らかにする



* 「正しい発声」をPC上で自動判定し、教師音を発声するボイストレーニングシステムへの応用

* 骨導音声を併用した補聴器の研究・開発へ応用



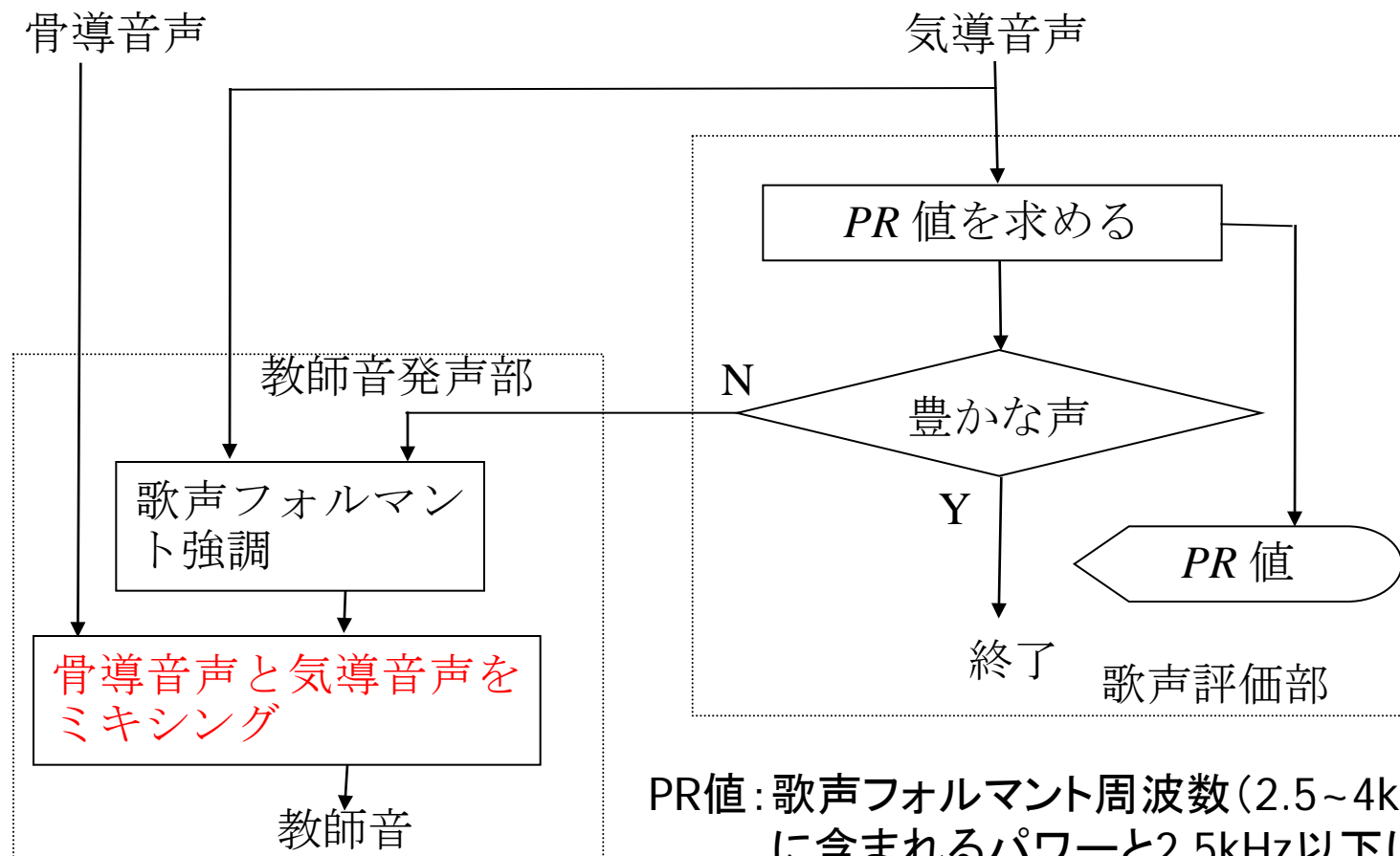
GOOD!



BAD!

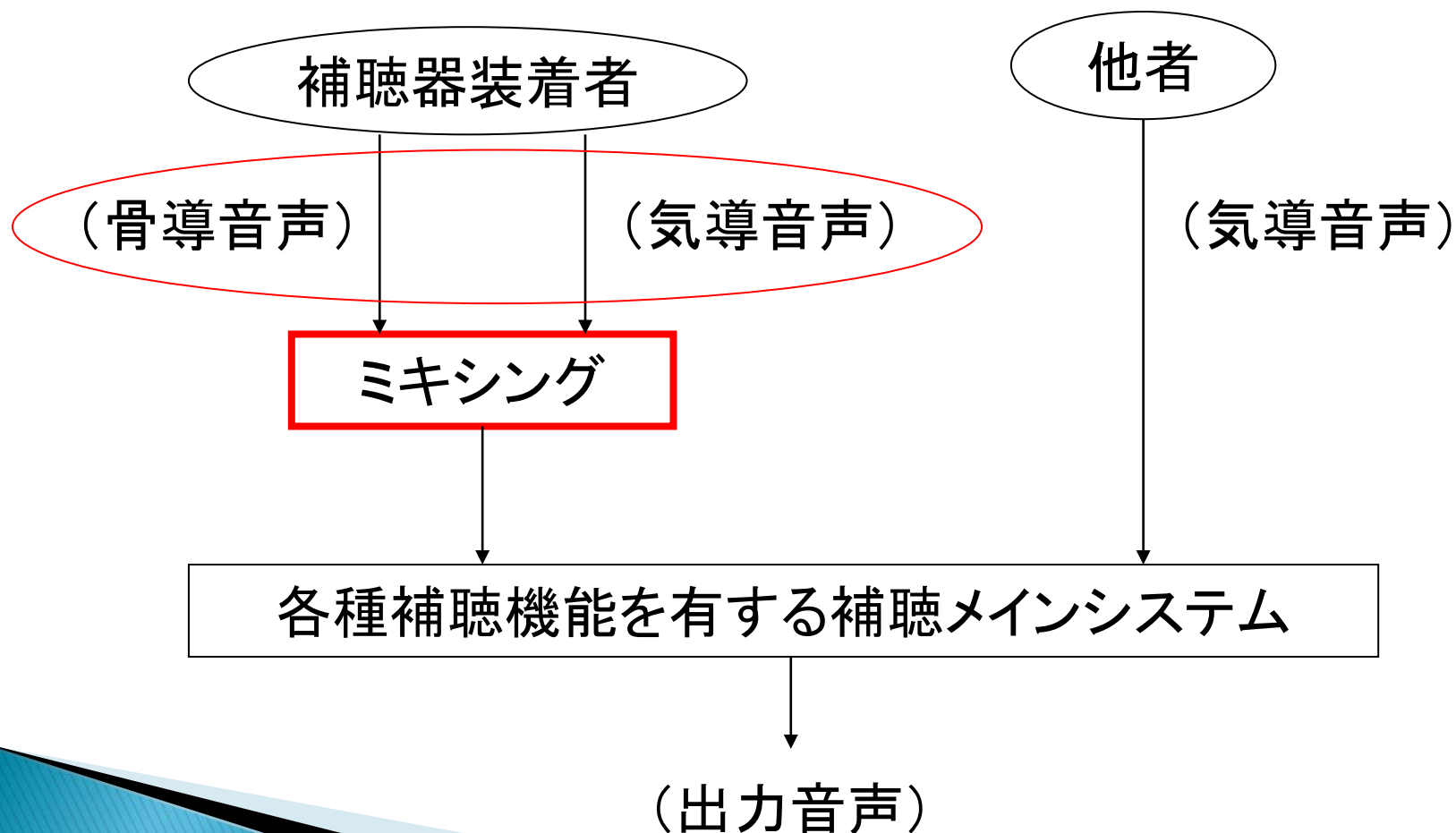
ボイストレーナーのイメージ

骨導音声を併用した ボイストレーニングシステム



PR値: 歌声フォルマント周波数(2.5~4kHz)に含まれるパワーと2.5kHz以下に含まれるパワーの比

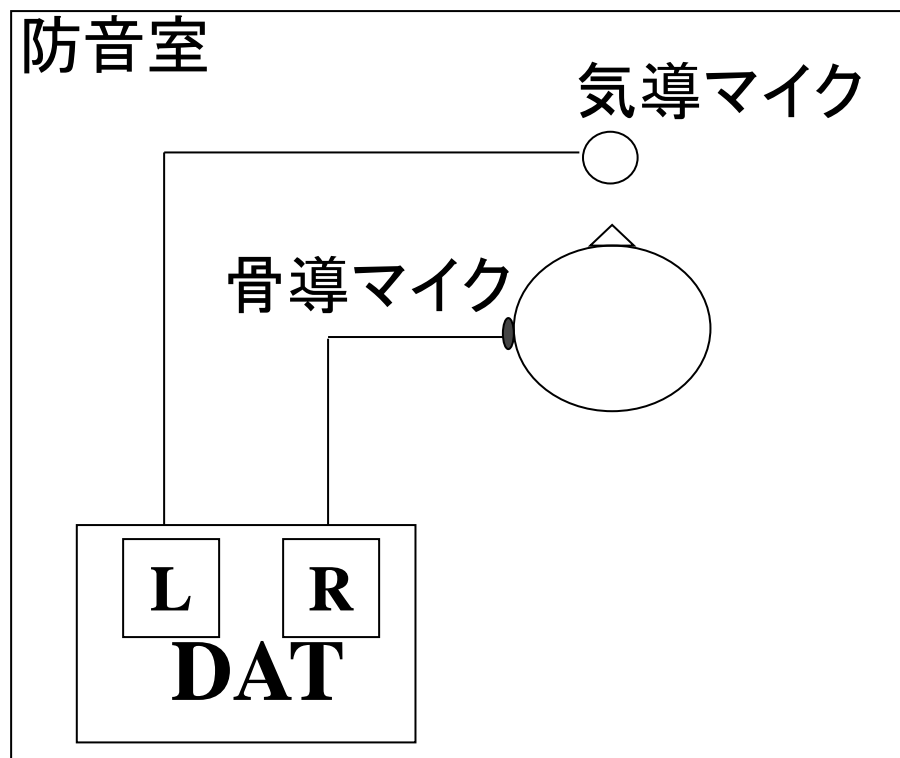
骨導音声を併用した補聴器



収録環境

防音室内で気導音声と骨導音声を同時に収録

サンプリング周波数:48kHz 量子化ビット:16ビット



気導音声の回りこみを防ぐため、骨導マイクの上からイヤーマフを装着

音声データ

- ▶ 20歳～24歳の成人男性62名、成人女性8名、計70名
- ▶ 日本語5母音 (/a/, /i/, /u/, /e/, /o/)
- ▶ サンプリング周波数48kHz
- ▶ 量子化ビット 16ビット
- ▶ 気導マイクの高さ 発声者の鼻の頭の高さ
- ▶ 気導マイクと口唇の距離 10cm(一定)
- ▶ 連続で5回発声

聴取実験方法

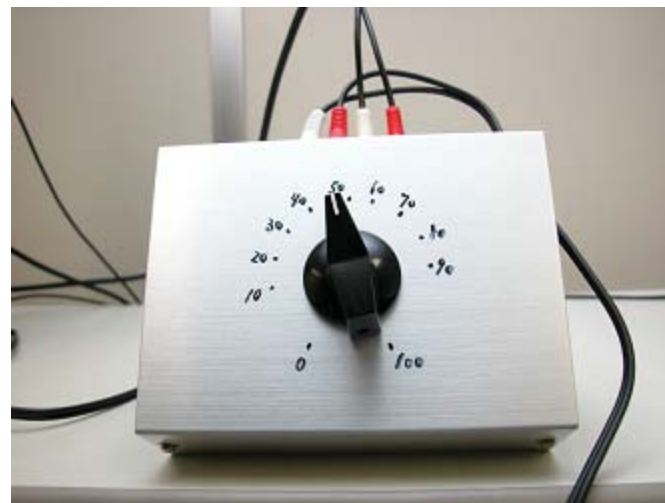
- ▶ 気導音声と骨導音声を最大振幅で規格化
- ▶ 左右のチャンネルの振幅比をクロスフェーダーで変化させて再生(被験者と音源との距離:50cm)
- ▶ 被験者は同じ母音を発声しながら、自己聴取音との聴感上の違いが最小となるときの気導音声の割合を10%刻みで読み取る
- ▶ 被験者はブラインド状態で自らの聴力のみで聞き取りを行う(クロスフェーダーを操作するのは第3者)

日付を変えて3回行った

聴取実験風景



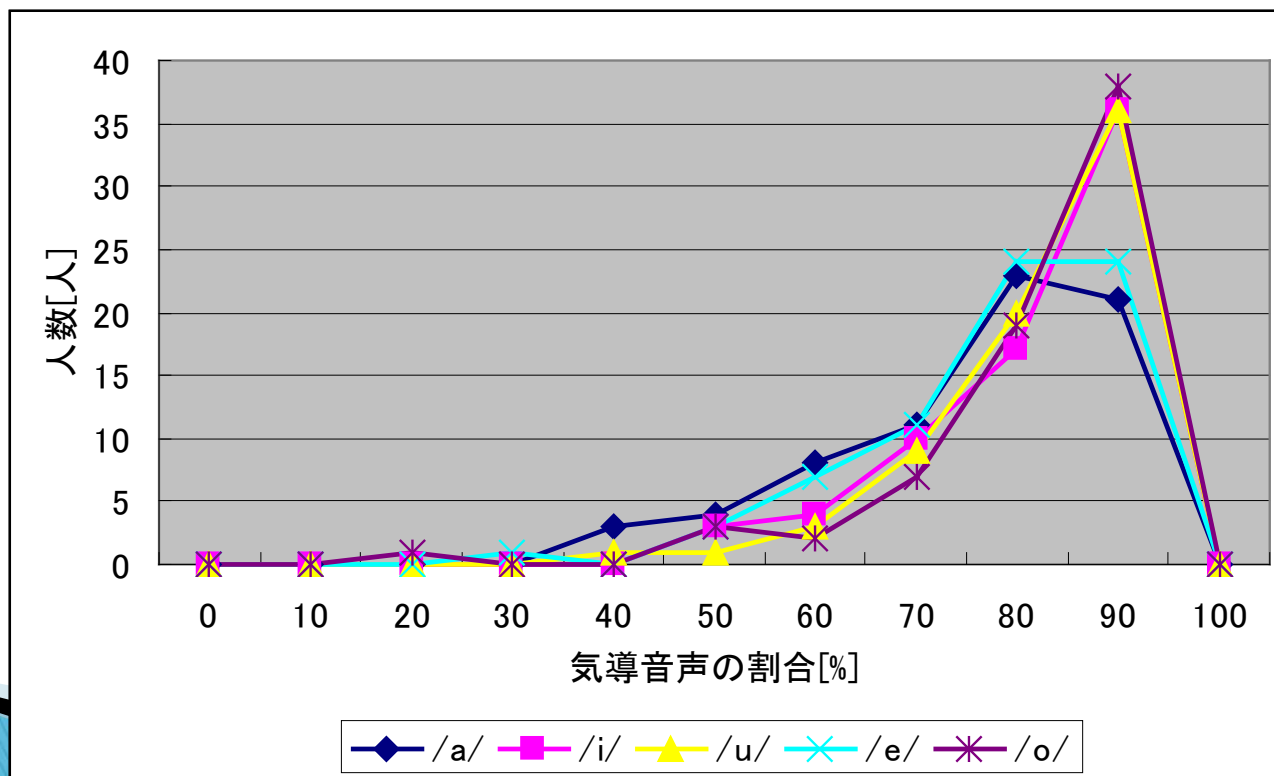
聴取実験風景



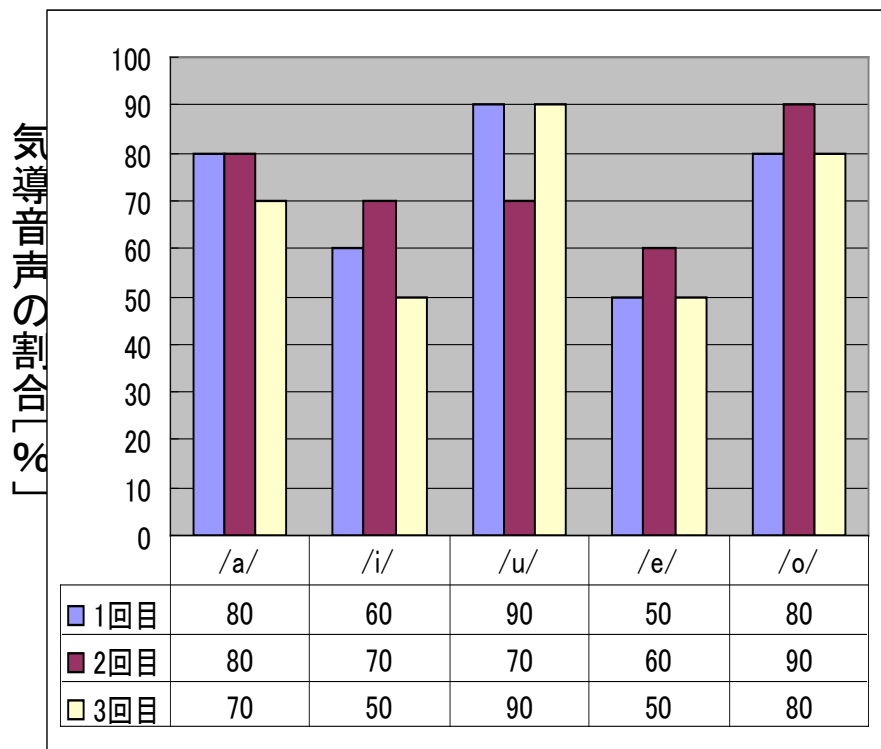
クロスフェーダー

実験結果：平均値と度数分布図

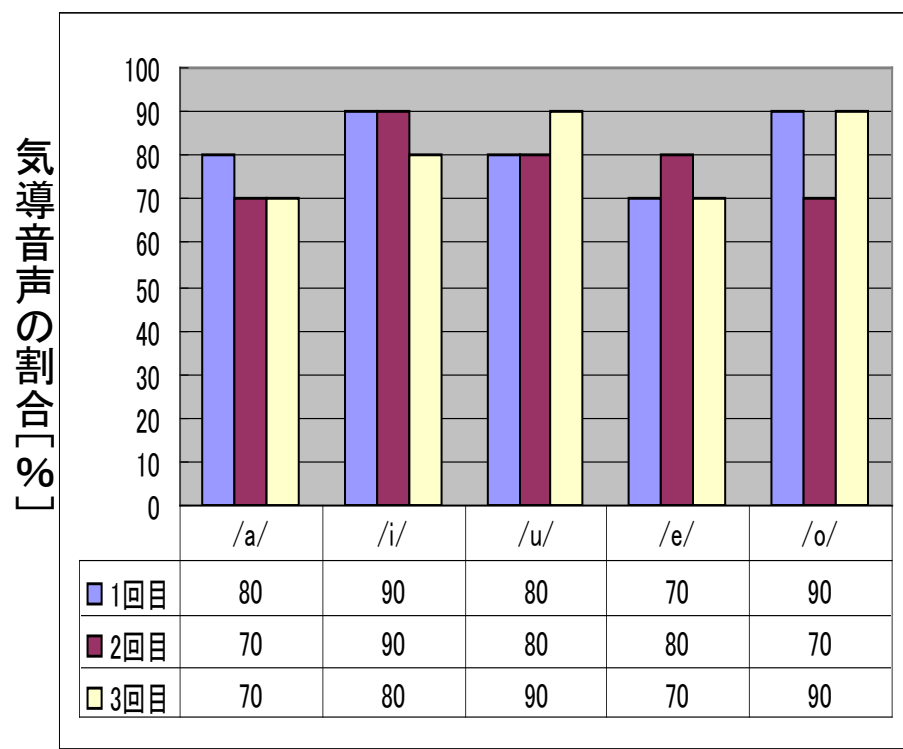
日本語母音	/a/	/i/	/u/	/e/	/o/
気導音声の割合[%]	73	79	79	75	79



実験結果：発声者2名

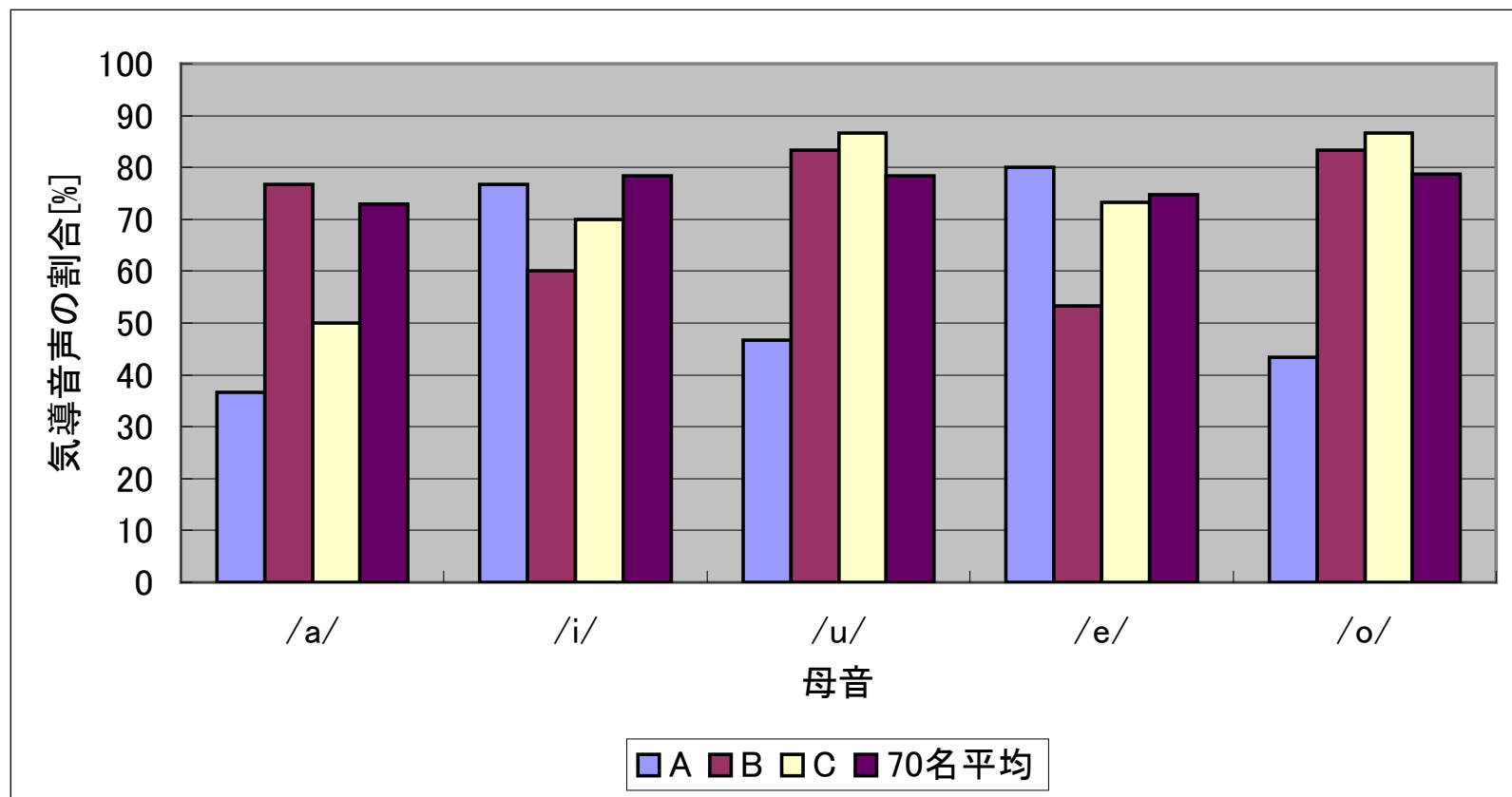


被験者No51



被験者No52

実験結果：発声者3名と平均値



まとめ

- ▶ 発声時、発声者は自分自身の気導音声と骨導音声を同時に聞いていることがわかった
- ▶ 気導音声の割合は10～90%に分布していることがわかった
- ▶ 同一話者でのばらつきは少なく、再現性があることがわかった
- ▶ 同じ発声者でも母音によって気導音声の割合が異なることがわかった
- ▶ 同じ母音でも発声者によって気導音声の割合が異なることがわかった

考察

- ▶ 同じ発声者でも母音によって、同じ母音でも発声者によって異なる原因は、母音によって口の開き方が異なることやホルマント周波数の違いによる影響と考えられる

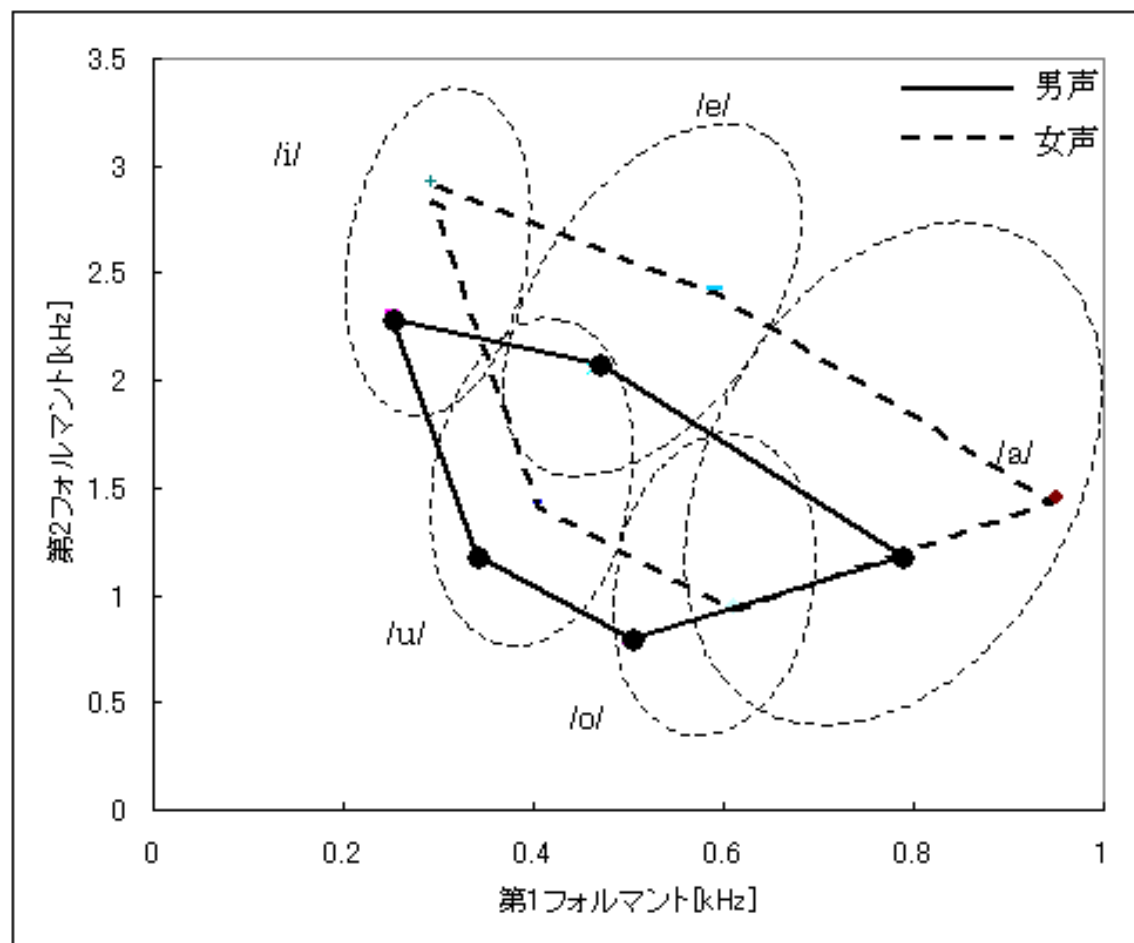
今後の課題

- ▶ 骨導音声収録方法の検討
 - 骨導マイクの特性を考慮
 - 本人の受聴している骨導音を収録できているか？
- ▶ 子音、会話音声に対する割合調査実験
- ▶ 難聴者に対する割合調査実験

参考資料：口の開きと舌の位置

舌の位置 口の開き	前舌	中舌	後舌
狭	i	u
半狭	e	o
半広
広	a

参考資料:F1—F2図



嗚声復元法の検討

福井大学大学院工学研究科
情報・メディア工学専攻

谷口秀次

はじめに

- 病変声帯に起因する嘎声(しわがれ声)
⇒ 講演や会話時などに聞きづらい



嘎声を正常音声に復元するようなシステムがあると便利

- 現在開発されている音声入力システム
⇒ 正常な音声を対象としている



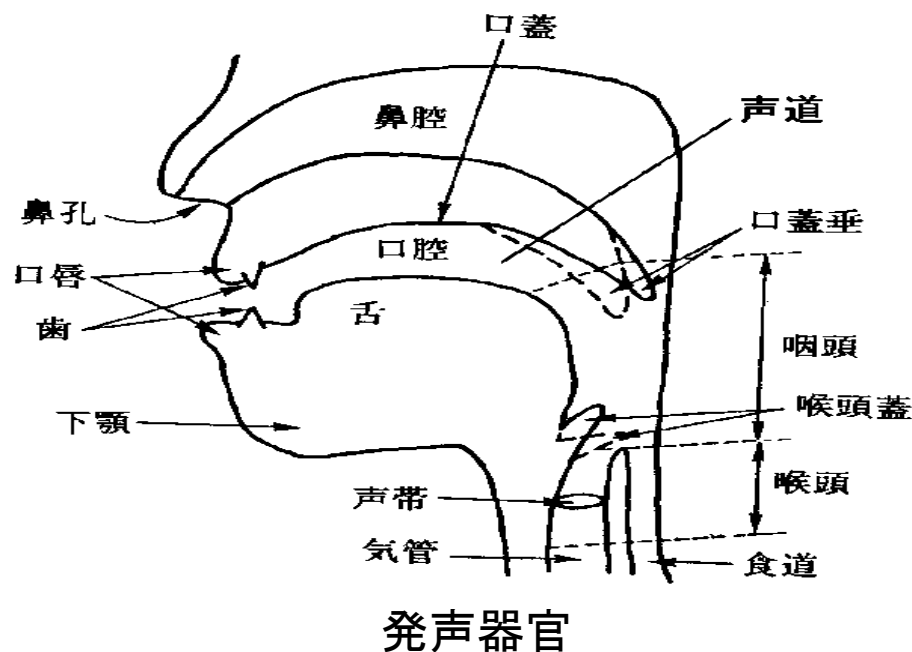
発声障害者にも対応できる音声入力システムの開発が望まれる

研究の目的

- 嚙声の中でも比較的症例が多く、音響的特徴の比較的明確な粗糙性嚙声(ガラガラ声)に注目し、その嚙声をより正常な音声に復元するシステム:嚙声復元システムを開発する。

嗄声とは

- 病変声帯に起因して生成される音声で、一般にはしわがれ声、専門用語では嗄声と呼ばれている
- 60歳代で30%、80歳代で50%が嗄声



嗄声の原因と特徴

- 声帯にポリープや癌などの病変が生じたり、声帯振動を制御する反回神経の麻痺などが起きると、声門の一部あるいは全体に閉鎖不全が生じたり、声帯振動が不規則になるという現象が見られる場合が多い。
- これらによって音声に気息性雑音が混じったり、音声波形の振幅や基本周期が不規則に大きく揺らぐという現象が現れる。

嗄声の種類と評価尺度

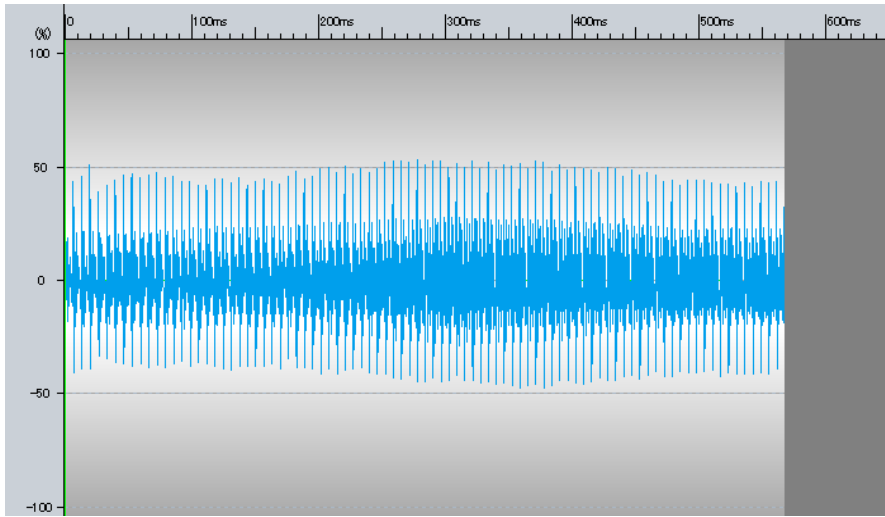
嗄声は4つの声質(聴覚的印象)を持つ

- 粗糙性(R, rough) : ガラガラ声
- 気息性(B, breathy) : 息もれ、かすれ声
- 無力性(A, asthenic) : 弱々しい声
- 努力性(S, strained) : 無理した声、途切れ声

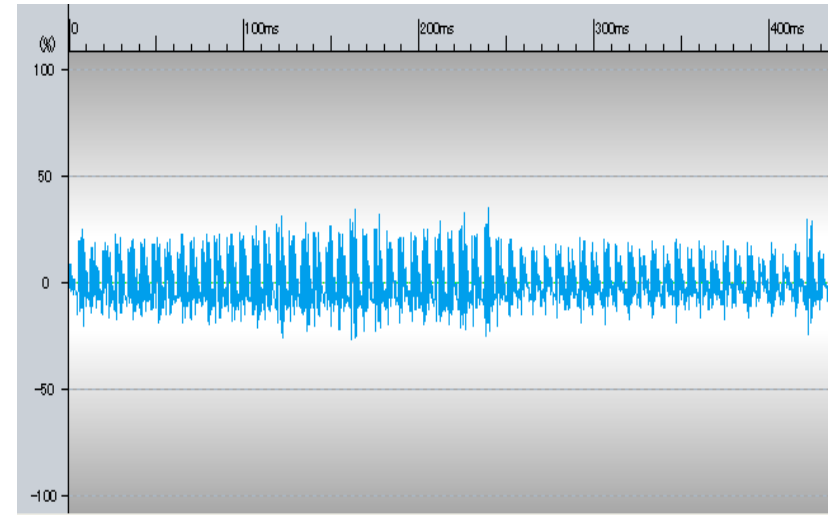
嗄声の聴覚的印象評価: GRBAS尺度により行われる

GRBAS尺度: 音声異常度を測る尺度

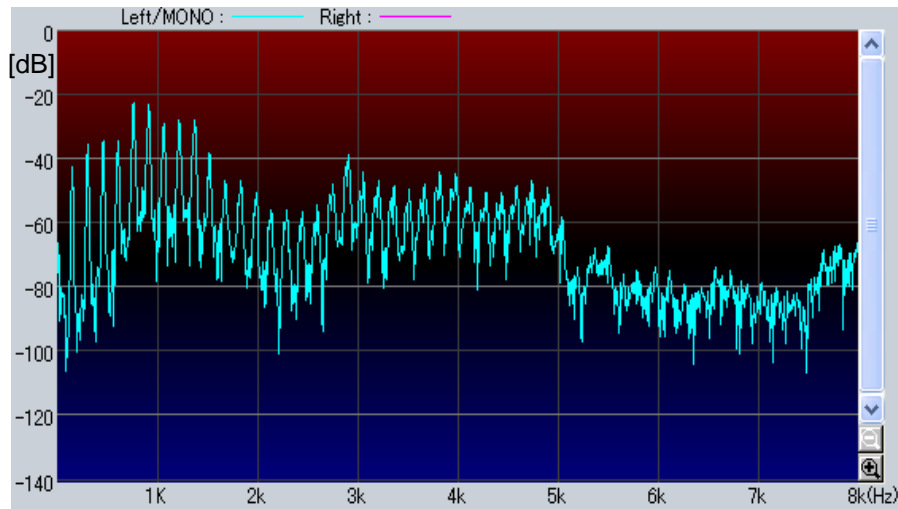
G(総合的嗄声度、Grade)、R、B、A、Sを
0, 1, 2, 3の4段階で評価



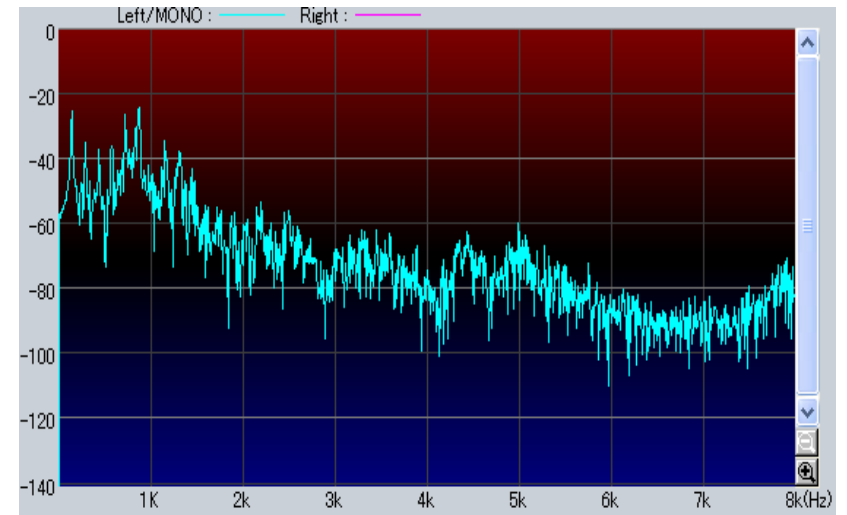
正常音声波形



粗糙性嘔声波形



正常音声波形のスペクトル



粗糙性嘔声波形のスペクトル

復元方式

スペクトル分析合成方式

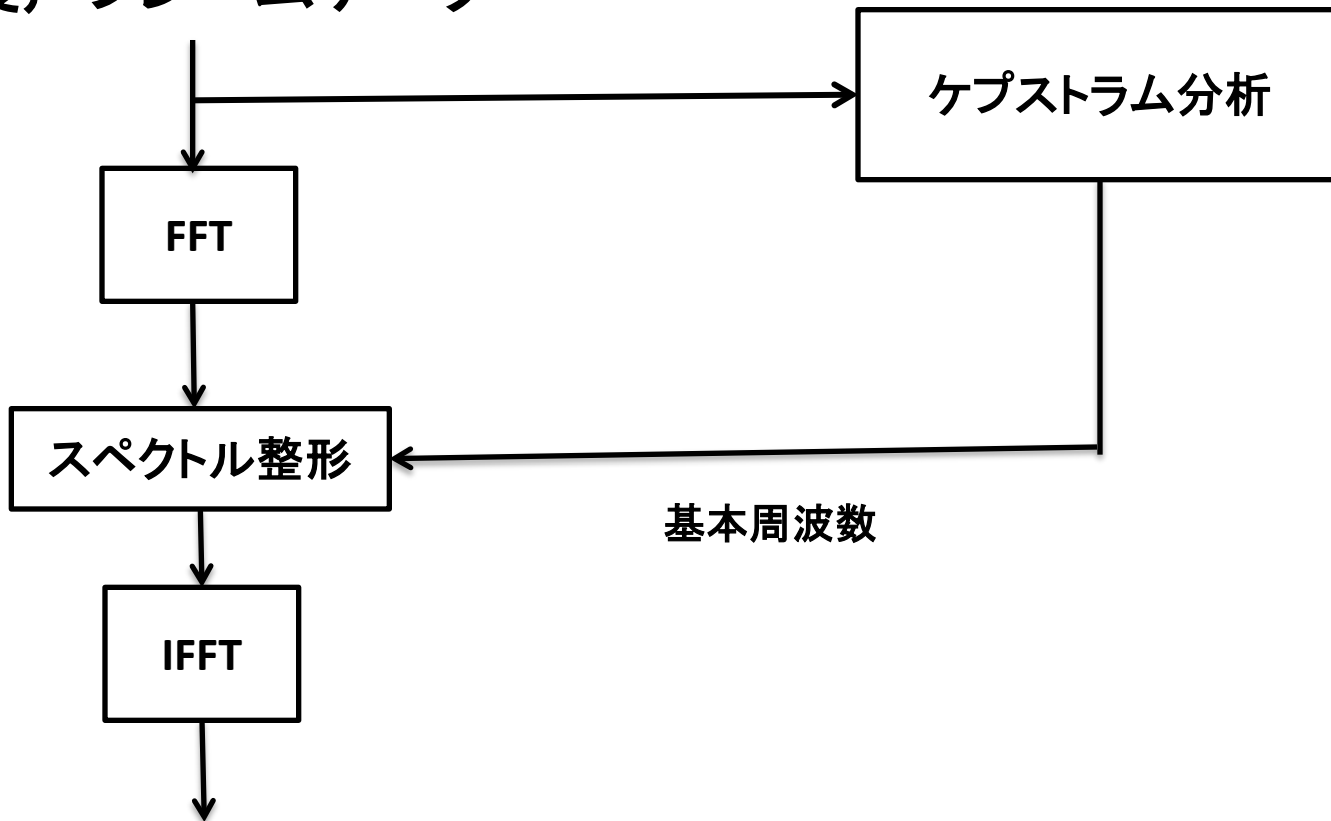
⇒ 嘆声の振幅 (A) や基本周波数 (f_p) の
過度な揺らぎを減らすことができる



フォルマントの形を損なわず、同時にスペクトル包絡の形も維持できるため、音韻性、自然性の保持にも期待できる

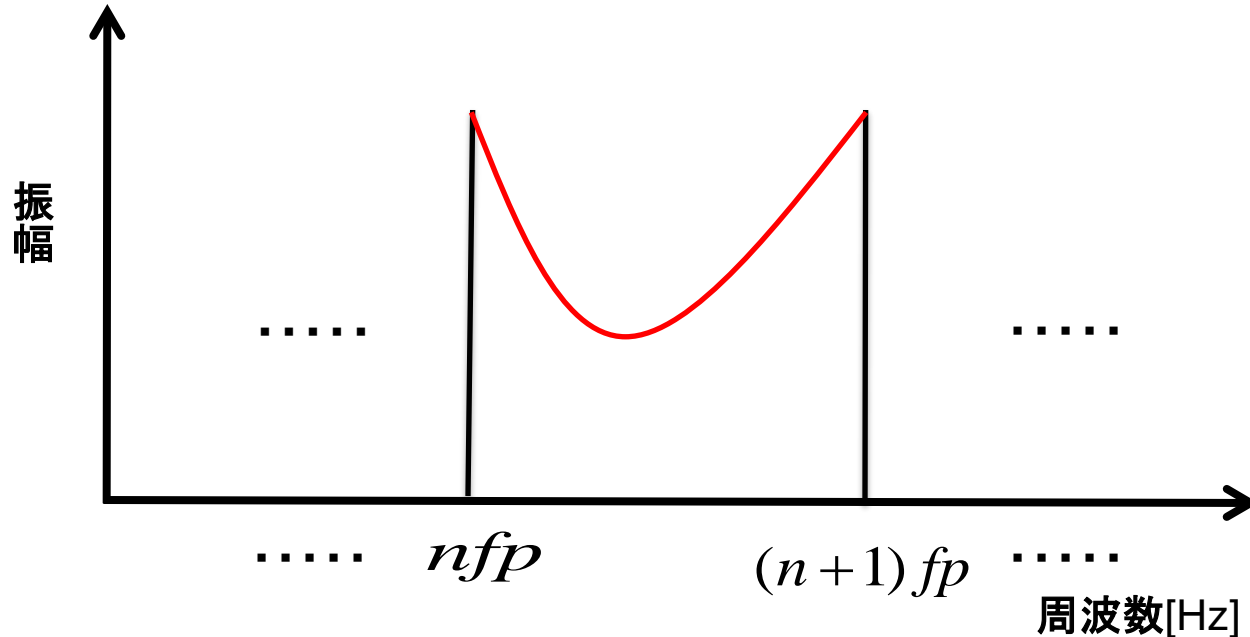
スペクトル分析合成方式の 処理フロー

嘆声フレームデータ



合成波形の接続&加工 → 復元音声

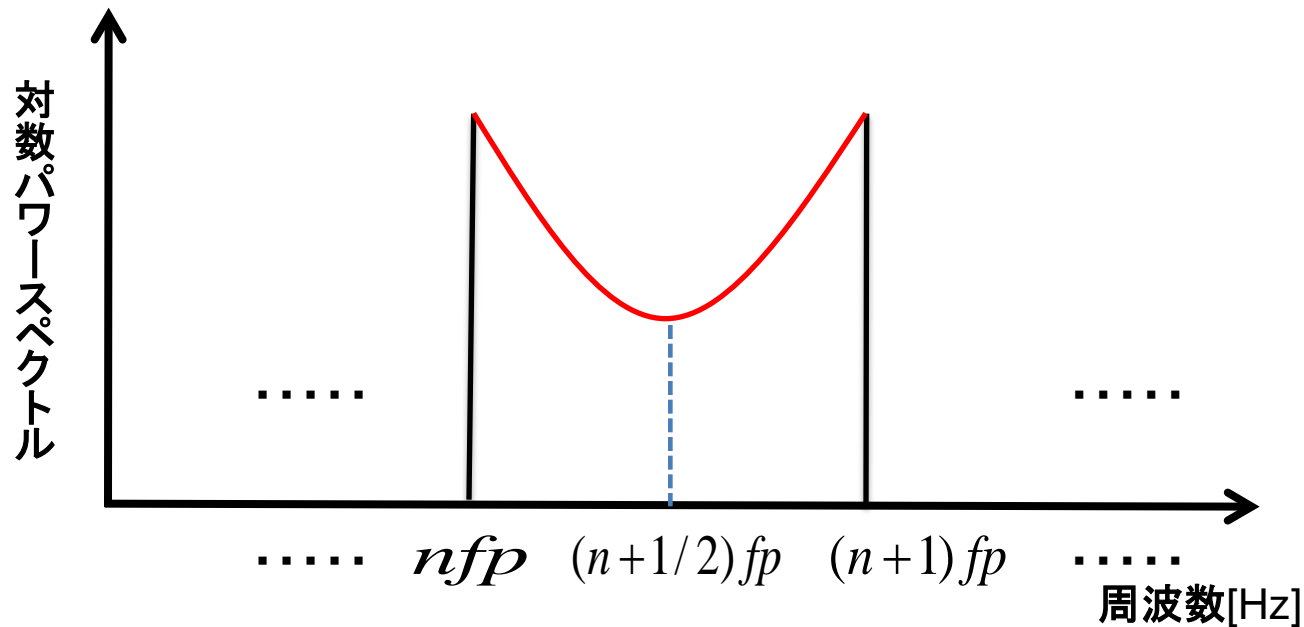
スペクトル整形方式 I



スペクトル整形部でのスペクトル処理イメージ

- ① nfp 間成分を移動平均により滑らかにする
- ② 各 nfp 間の振幅スペクトルの最小値、両隣の nfp の振幅値の3点を通る2次関数を与える

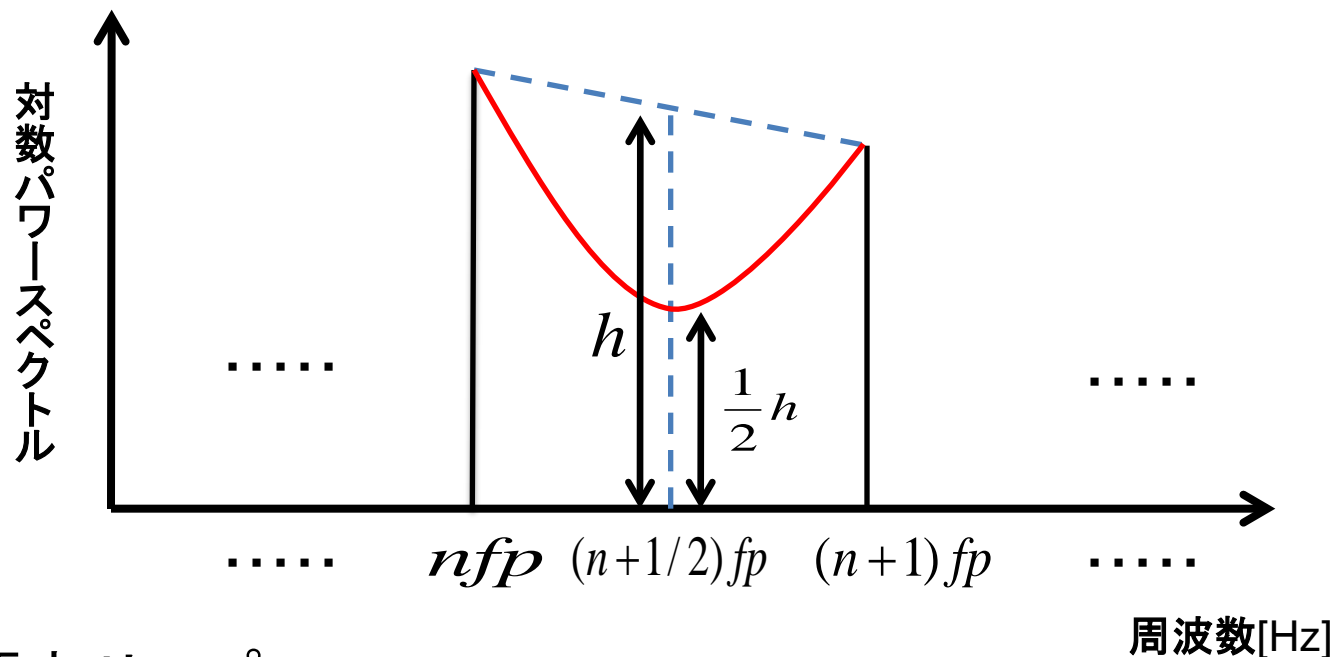
スペクトル整形方式Ⅱ



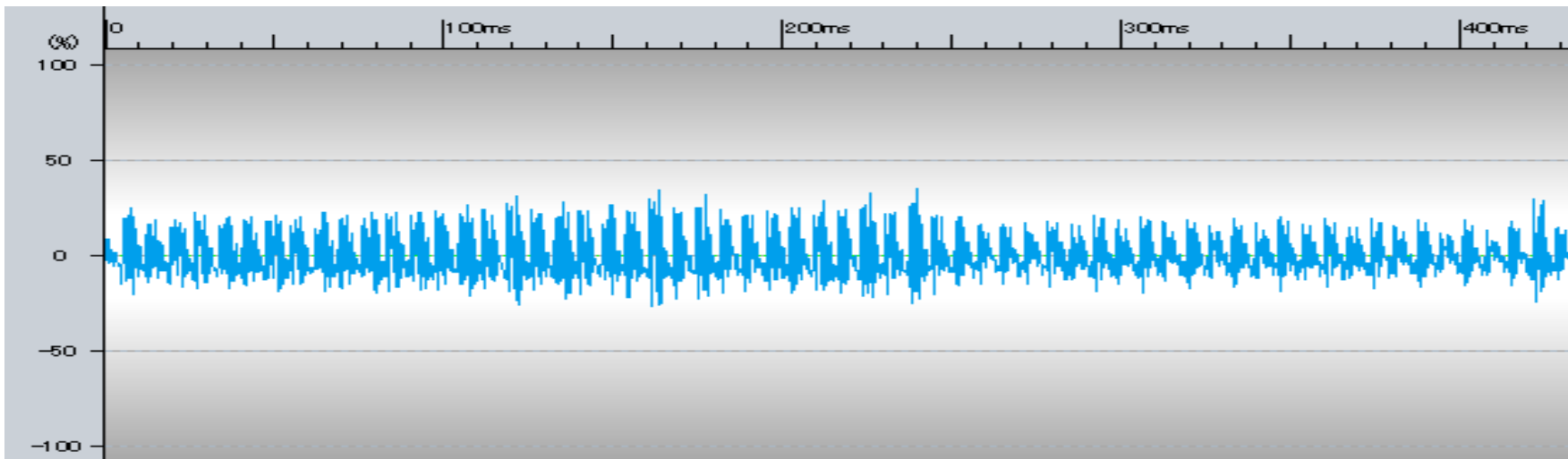
スペクトル整形部でのスペクトル処理イメージ

- ① nfp 間の最小値の周波数位置を nfp 間の中央に定める
- ② 各 nfp 間の対数パワースペクトルの最小値、両隣の nfp 成分の値の3点を通る2次関数を与える

復元実験



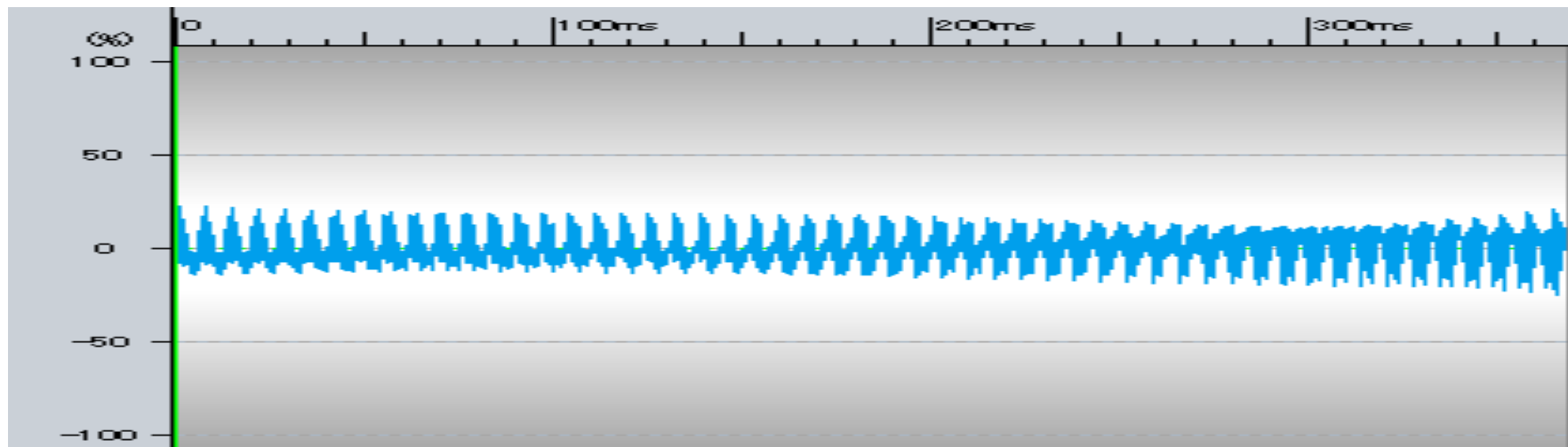
- 嘎声サンプル
嘎声データベース(動画で見る音声障害Ver.1.0”)中の
粗雑性3、または2の5サンプルを使用
- 2次関数の極小点の値(h の長さを $(1/10, 2/10, \dots, 9/10)$ まで)を変化させて音質を調べた



粗糙性嘎声波形

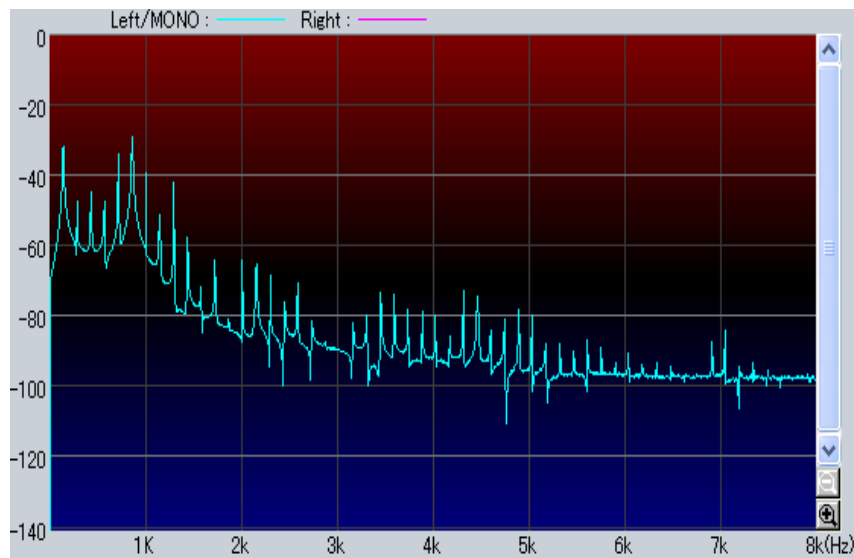


(原嘎声と復元音声)

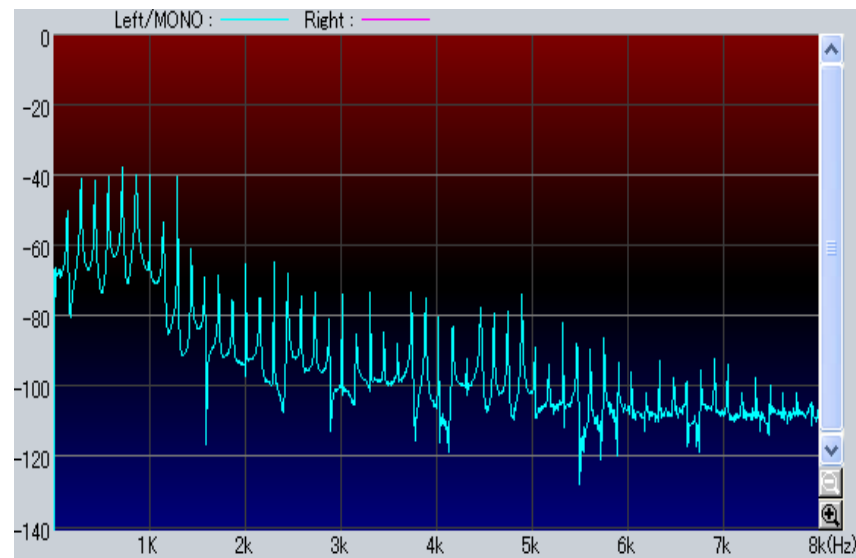


復元波形 (一部)

方式 I と方式 II による スペクトルの違い



方式 I による
復元音声/あ/のスペクトル

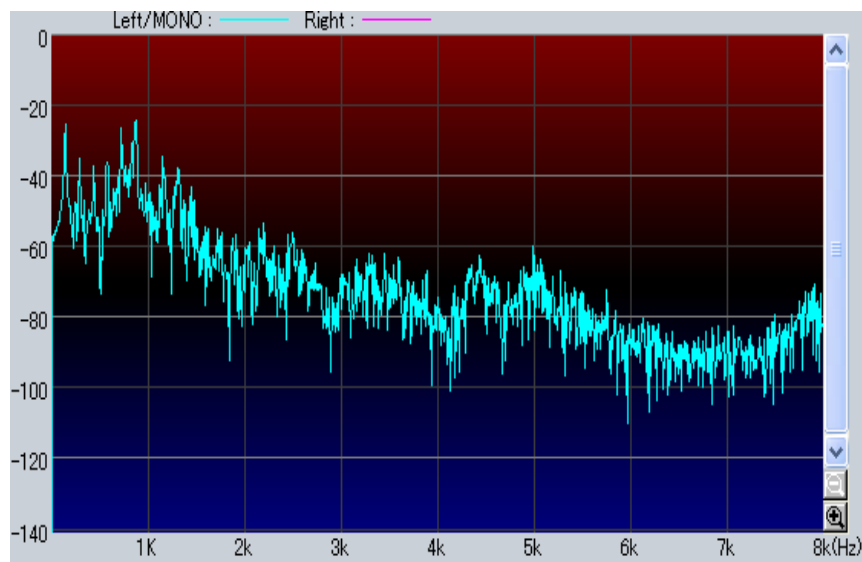


方式 II による
復元音声/あ/のスペクトル



- 方式 II によるスペクトル: 調波構造が高い周波数まで見られる

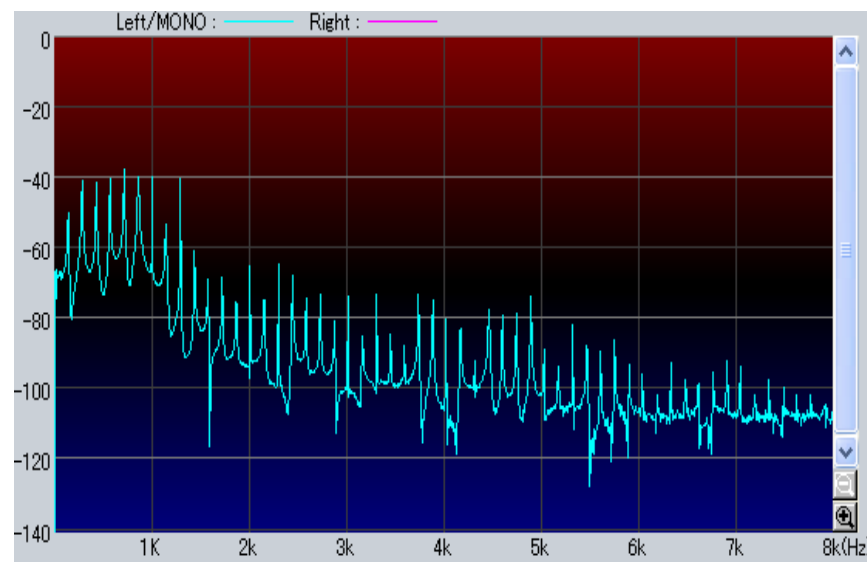
粗糙性嘎声と復元音声の スペクトルの違い



粗糙性嘎声/あ/のスペクトル



- 周期性が乱れている
- 振幅に不規則な乱れがある

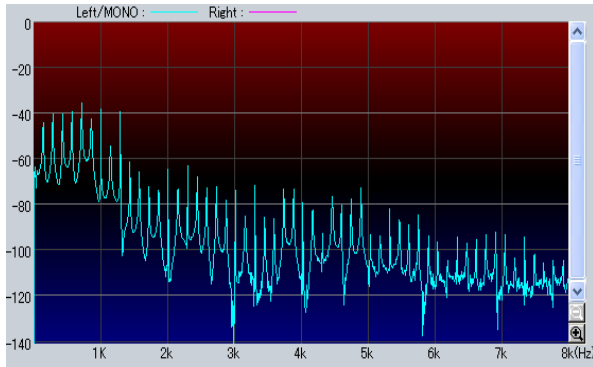


復元音声/あ/のスペクトル

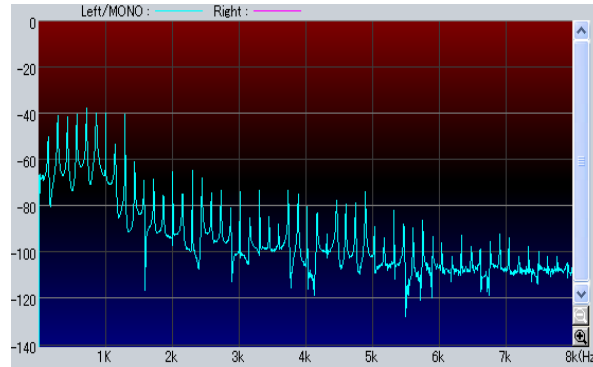


- 振幅の乱れが取り除かれている

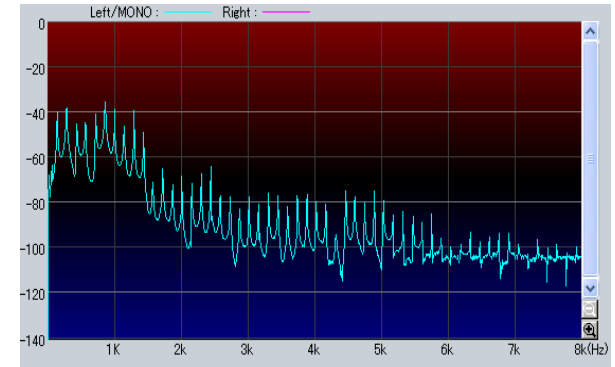
極小点の位置(深さ)を変化 させた場合のスペクトルの違い



復元音声/あ/のスペクトル(深さ $\frac{1}{10}h$)



復元音声/あ/のスペクトル(深さ $\frac{1}{2}h$)



復元音声/あ/のスペクトル(深さ $\frac{9}{10}h$)



- スペクトルに大きな変化はなく、音質もほとんど変わらない

まとめ

- 復元の結果、粗糙性については少し改善されたという評価を得た。
- 音韻性、自然性については改善できなかった。
- 与える2次関数の極小点の位置(深さ)を変化させても復元音声に大きな変化は見られない。

今後の課題

- スペクトルのnfp間成分の与え方の検討
- 基本周波数の高精度抽出法の検討
- 音韻性、自然性を劣化させない本復元方式の改良
- 合成音声波形の接続や波形加工による自然性回復