

平成28年度 博士論文

《論文題目》

加齢や病気により変質した音声の明瞭度改善に関する研究

平成28年度

14595401

田中 裕人

指導教員 中藤 良久 教授

九州工業大学大学院 工学府

工学専攻 電気電子工学領域

第1章	序論	1
1.1	研究背景	2
1.1.1	音声コミュニケーションの重要性	2
1.1.2	高齢者とのコミュニケーションの現状と課題	3
1.1.3	喉頭摘出者とのコミュニケーションの現状と課題	4
1.2	本論文の目的	12
1.2.1	本論文の新規性及び有効性	12
1.3	本論文の構成と各章の概要	13
第2章	高齢者音声の明瞭度に関する音響的特徴	15
2.1	高齢者音声データベース	16
2.2	被評革者の選定	17
2.2.1	明瞭度の低い声とは	17
2.2.2	聴取実験による明瞭度の高低の判別	18
2.2.3	被評価話者の選定	20
2.3	高齢者音声の遷移量と調音位置の関係	22
2.3.1	データ解析手法	22
2.3.2	高齢者音声の遷移量	26
2.3.3	遷移量差と調音位置の関係	28
2.4	高齢者音声における F1-F2 平面上の母音間距離	32
2.4.1	データ解析手法	32
2.4.2	グループで見た明瞭度の高低による F1-F2 平面上の平均ホルマント周波数	35
2.4.3	低明瞭度の高齢者音声における話者ごとの平均ホルマント周波数	37
第3章	ホルマント面積拡張による高齢者音声の明瞭度改善方法の検討	44
3.1	ホルマント周波数シフトによる明瞭度改善アルゴリズムの概要	45
3.2	ホルマント周波数のシフト量について	52
3.3	シェッフェの対比較法を用いた聴取実験による明瞭度改善効果の検証	54

3.3.1	データ解析方法	54
3.3.2	シフト量が既知の場合における補正効果	57
3.3.3	シフト量が未知の場合における補正効果	61
第4章	電気式人工喉頭音声の音質改善システム	64
4.1	電気式人工喉頭音声改善システムの提案	65
4.2	LPC 分析合成に基づく音質改善方法の予備実験	67
4.2.1	実験内容	67
4.2.2	実験結果	68
第5章	種々の音源による電気式人工喉頭の音質改善	72
5.1	電気式人工喉頭と擬似生体音源による音声の比較実験	73
5.1.1	擬似声帯音源を用いた音声の生成	73
5.1.2	音声収録	75
5.1.3	聴取実験条件	76
5.1.4	シェッフエの一対比較による主観評価	77
5.1.5	ケプストラム分析による客観評価	79
5.2	LPC 残差波における抽出音源の違いによる音韻性への影響	86
5.2.1	音声収録	86
5.2.2	聴取実験条件	86
5.2.3	シェッフエの一対比較による主観評価	87
第6章	スペクトル包絡のマッピングに基づく音質改善方法の検討	94
6.1	スペクトル包絡のマッピング概要	95
6.2	スペクトル補正音声の作成	96
6.2.1	GMM による音声モデル作成	96
6.2.2	GMM モデル間距離を用いたマッピング	97
6.2.3	LPC 分析合成による補正音声の生成	97
6.3	音声収録	99
6.4	客観評価による比較	100
6.4.1	5 母音のスペクトル包絡のマッピング	100

6.4.2	未知語へのスペクトル包絡のマッピング	106
第7章	結論	114
7.1	本論文のまとめ	115
7.2	今後の課題	117

第 1 章 序論

1. 1 研究背景

1. 1. 1 音声コミュニケーションの重要性

人間は、人間同士でコミュニケーションを取ることによって情報のやり取りを行っている。人類学者の Hall は、コミュニケーションを①相互作用、②連携、③生存、④性、⑤なわばり、⑥時間性、⑦学習、⑧遊び、⑨防御、⑩開発という 10 個の第一次情報体系で分類し、言葉によるコミュニケーションは、①相互作用に属すると述べている[1]。

人間の言葉を構成する聴覚信号は、①叫び、②発語、③発文の 3 つのレベルに従って組み立てられた信号形式をとる。叫びは、情動変化情報の伝達に用いられることの多い信号形式であり、言葉にあたる意味を持たない。発語は、命名や命令の伝達において最も重要な役割を持つ信号形式で、言葉の聴覚信号単位の伝達を担っている。この際言葉とは、単語単位のことであり、受信者の脳には、その聴覚信号で表される概念に関連した様々な記憶を受信と同時に想起される。発文とは、文という信号形態を用いる方法であり、文という形態をとることで、各単語の表す概念の相互関係を明らかにし、行為の記述や、判断・推理の伝達を可能にする信号形態である[2]。動物も様々な感覚チャンネルを用いてコミュニケーションを行うが、人間のコミュニケーションは、単語を一定の法則に基づいて文を構成することにより、行為の記述や、判断・推理を深層構造のレベルまでの情報を相互交換するという高度なコミュニケーションを行っている。これより、人間の音声によるコミュニケーションとは、脳にある意思等を文という形態で聴覚信号により別の脳に伝える、これを相互に繰り返すことである。言語情報のみならず、音声には、発話意図、感情といった情報も含まれている。感情情報は、音響的特徴量的には音声の声帯情報、強さ、持続時間、ホルマントなどの変化に表れると言われている[3]。音声は、たくさんの情報を持っており、これを用いて対人コミュニケーションを行うことは、非常に高度なコミュニケーションである。

平成 18 年に文部科学省は、対人コミュニケーションと社会的適応能力や脳の発達に相関があると報告しており[4]、音声によるコミュニケーションは、社会的適応能力の向上、脳の発達に繋がると考えられる。

1.1.2 高齢者とのコミュニケーションの現状と課題

現在、世界的規模で急速に高齢化社会になっており、中でも日本は高齢化率（65歳以上人口が全人口に占める割合）が2015年で26.7%に達し、更に2025年には、65歳以上人口は3500万を超え、高齢化率は30%以上になると予想されている[5]。このような超高齢化社会が進む中で、高齢者が安心して生活を営み、社会参加できるようにすることは急務である。また、近年レジャーや仕事にアクティブに取り組む高齢者が多く見られるようになってきている。このような方々にとって音声によるコミュニケーションは、本人のみならず家族や周囲の人にとって大変有益であると考えられる。

高齢者との音声言語によるコミュニケーションをすることにおいて、加齢による発声器官の衰えによる発声音声の劣化が生じ、多くの場合において、高齢者音声は成人の音声と比べ、聞き取りづらくなる。高齢者と円滑にコミュニケーションを行うためには、高齢者音声を明瞭化する必要がある。高齢者音声は聴覚的に、「明瞭性が低い」、「遅い声」、「しゃがれ」の3つに分類されると言われている[6]。これまでに、発話の音韻的な明瞭性の向上や、かすれ声の明瞭化といった研究がされている[7]が、未だ、完全に声の明瞭化が行うことはできていない。高齢者音声に関する分析として、30年以上前から行われているが[8-10]、学術的見地からの定性的な解析にとどまっている。また、高齢者の発声器官に関する医学的見地からの報告もあるが[11,12]、高齢者音声の明瞭性を低下させる原因がなにか、あるいは音響的にどのような特徴があるか、といった工学的見地からの基礎的な研究は少ない。そこで、基礎的研究として高齢者音声にはどのような音響的特徴があるかを明らかにし、高齢者音声を明瞭にすることは、高齢者とのコミュニケーションを円滑にすることが可能になると考える。これは、補聴器や電話器等に応用することに繋がる。

加齢によって聴力が低下したことを補償する研究（補聴器）は多数報告されている。例えば、言葉の聞き取りを改善するための手法に関する研究[13-20]、雑音対策に関する研究[21,22]などがある。中でも、言葉の聞き取りに関しては、成人の音声を聴力の低下した高齢者に対して周波数[13-19]、話速変換[20]により、聞き取りを改善する研究は報告されているが、高齢者音声を聞き取りやすくする研究は見受けられない。

また、音声認識分野において、高齢者音声は認識率が成人と比べ、認識性能が低下すると報告[23]もあり、認識性能改善に向けた研究[24,25]も報告されている。しかし、高齢者音声自体の明瞭度を向上させるという手法を用いた研究は見受けられない。

1.1.3 喉頭摘出者とのコミュニケーションの現状と課題

喉頭癌などの病気によって声帯を失い音声を使ったコミュニケーションを行うことができなくなってしまう喉頭摘出者 (the laryngectomee), または無喉頭者 (alaryngeal patient) と呼ばれる人がいる。喉頭癌というのは60歳以上になるとかかりやすくなる癌で、比較的治りやすいが症状が進んでしまった場合や再発してしまった場合は喉頭を摘出しなければならない。その喉頭摘出術により、呼吸器官系と消化器官系が交叉している喉頭部を摘出し、呼吸を行うために肺と直結した気管孔を全頸部下方に設けなければならない。このため呼吸器官と消化器官は分断されることとなる (図 1.1)。

そのため術後の障害として、発声ができない、気管孔に水が入ると危険なので入浴の制限がある、食道と気管孔が完全に分断されているので食事に難がある、などがあげられる。

喉頭摘出者に対するアンケート調査[26]では、

- ・看護師や家族以外と話をするのが大変であり気が進まない
- ・自分の要求を伝えられたとしても日常会話は減少するだろう
- ・声が出せないもどかしさを理解して欲しい

などの声を出せないことに対する不満が多々あげられていた。これより、音声を失ってしまった人は声を取り戻したいと考えている事がわかる。

このような喉頭摘出者数は日本で12,000人、世界で600,000人程度と推定される。喉頭摘出者の年齢層は高齢者が多く、高齢人口が増加すると患者数が漸増している。これからの高齢化社会ではますます喉頭摘出者が増えるのではないかと想像される[27]。

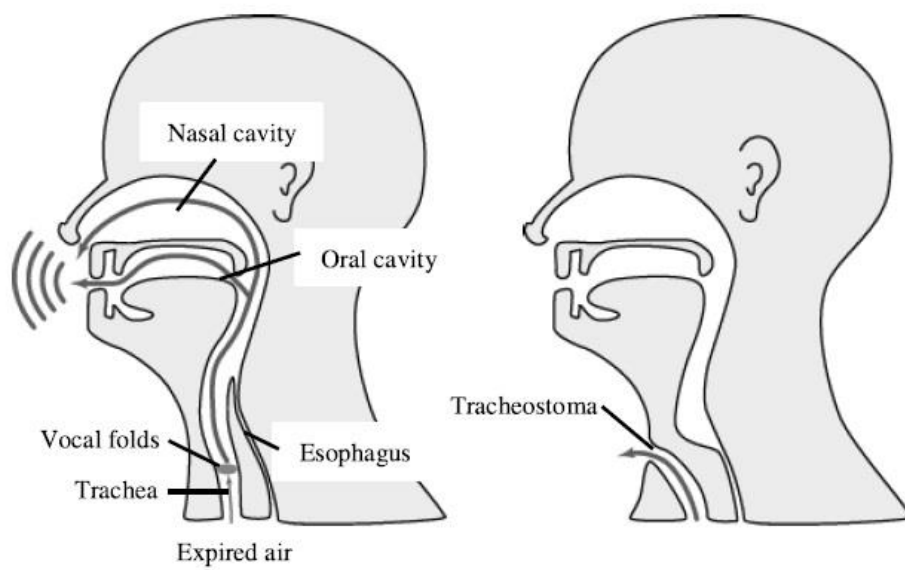


図 1.1 肺からの呼気の流れ（左は健聴者，右は喉頭摘出者）（[28]より転載）

発声器官の構造

我々人間の生活と切っても切れない存在である音声は、図 1.2 の構造で発声している。腹筋が横隔膜を押し上げることによって、肺から押し出された空気は、気管を通ったあと、喉頭の声門 (glottis) すなわち左右の声帯 (vocal cords) の間を通る。通常の呼吸の時は、声門は大きく開いているが声を出そうとすると声帯が接近する。この間を肺からの空気が通りぬけようとするため、空気流と声帯との相互作用により、声門が周期的に開閉し、ほぼ規則的な空気の断続が生じる。これは非対称三角波で近似でき、これが音声の音源となる。これを喉頭原音または声帯音源 (glottal source) と呼ぶ。声帯の緊張が大きく、かつ肺からの空気圧が高いと、声帯の振動周期 (基本周期, fundamental period) が短くなって、音源の音の高さが高くなり、逆の時は低くなる。基本周期の逆数を基本周波数 (fundamental frequency) と呼び、声の高さ (ピッチ, pitch) に対応する。

喉頭より上の部分は声道 (vocal tract) と呼ばれ、顎、舌、口唇などを動かすことによって、種々の形に変化する。これによって音源波に音色が付与され/a/や/o/の音声となる。鼻腔は軟口蓋を持ち上げることにより、声道から遮断される。種々の言語音を発声するために声道の形を調整することを、調音 (articulation) と呼び、声道の各部分の動きを調音運動と呼ぶ。調音に用いられている声道の各部分を調音器官 (articulatory organ) と呼び、その中で、舌、口唇、口蓋帆のように自由に動けるものを特に調音器 (articulator)、調音によって生ずる声道の狭めの位置を調音点あるいは調音位置 (place of articulation) と呼ぶ。調音によって種々の音色が付与されるのは、伝達特性が変わり、共鳴作用によって周波数的にエネルギーの強弱が生ずるためである。

喉頭摘出者は喉頭、声帯が無いいため発声の元となる声帯音源を作り出すことができない。調音器官はあるものの声帯振動が無いいため発声することができなくなる。

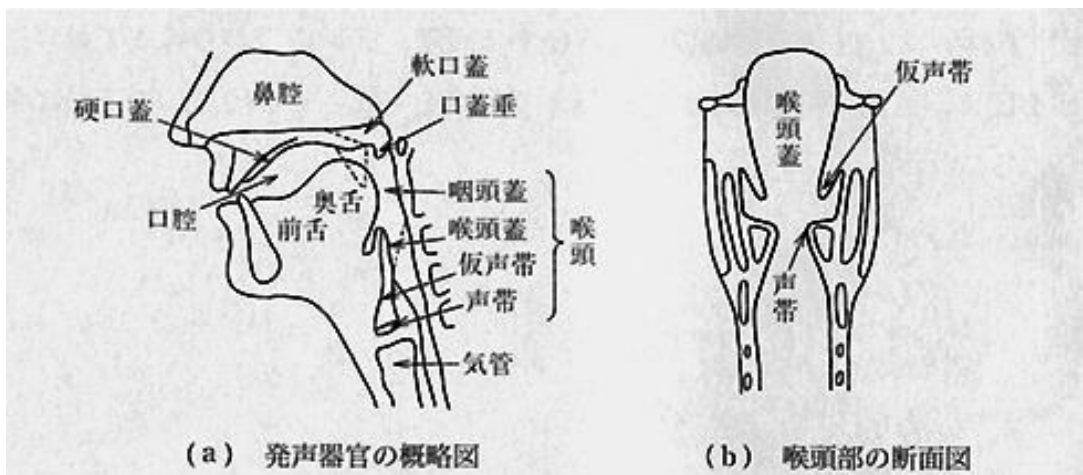


図 1.2 発声器官[29]より転載

代替発声法

喉頭摘出者は、調音はできるものの、発音の元となる声帯音源が無いために普通の方法で発声することができない。そのため、声帯音源の代わりとなる音源を生成し発声しなければならない。そのための代替発声法は主に電気式人工喉頭（以下 EL）を用いる方法、笛式人工喉頭を用いる方法、T-E シヤント発声法、食道発声法等がある。

食道発声法とは、いわゆる「げっぷ」を利用する方法で、食道の中に空気をため、貯めた空気を吐き出すときの音を利用する方法である。ゲップ音が口から出てくるときに、声道の形が変わることによって音声となる。

長所として、器具を使用しないので手が塞がることがない、熟達すると普通の音声のような発声ができる、などがあげられる。短所として、この方法は獲得が難しく個人差はあるが練習期間は6ヶ月から1年もかかる、音声の明瞭度は熟練者と未熟者では大きな差がある、訓練が大変であるので高齢化すると獲得がより難しくなる、などが挙げられる。

T-E シヤント発声法(traphea-esophageal speech production method)とは、Tは器官(trachea)、Eは食道(esophagus)の意味である。このTとEの部分に小さなシヤントすなわち穴を開けて、気管孔を指で塞ぐと穴から空気が食道の方に行き、この部分に仮声門(psoudg lottis)と呼ばれる声門が振動して音が発生する。これを音源として利用するのがT-E シヤント発声法である。

長所として、肺からの呼気をエネルギー源としているので容易に発音が出来、目立つ器具は使用しない、明瞭な音声で発声できる、ある程度ピッチを付与することが出来る、などが挙げられる。短所として、新たに手術をして体内に器具を埋め込む必要がある、気管孔を片手で塞がなければならない、食道と気管が穴で繋がるので誤嚥の可能性があるのである、などが挙げられる。

笛式人工喉頭を用いる方法とは、タピア(Tapia)式人工喉頭とも呼ばれ、摘出術時に開けた空気孔に装置を装着し、呼気をエネルギー源とし、ゴム膜や金属板を振動させて生成した音を、パイプを介して口腔内に伝達させる方法である。この方法はELが普及する前は最も使用されていた方法である。振動体がゴムの場合は、貼り方、形状、笛式人工喉頭使用者の訓練度合いにより、ピッチ周波数を制御することも可能であるが、一般にはその音声は単調である。

長所として、扱いが簡便である、習得が容易である、音声の抑揚を容易に付けられる、などが挙げられる。短所として、器具が目立つので周囲の目が気になる、音声が単調である、などがあげられる。

EL を用いる方法とは、電気をエネルギーに振動する振動体を振動させ、その振動音を声帯振動音の代わりとして用いる方法である。その装置のことをEL という。最もよく用いられているものは皮膚電動式EL と呼ばれるもので、下顎頸部に振動子を押し当て、皮膚を介して振動音を口腔咽頭に伝達する装置である（図 1.3）。

長所として、扱いが簡便である、習得が容易である、他の代替発声法よりも清潔である、などが挙げられる。短所として、人の音声とは言えない音である、発声できない音素がある、声道を通らずに放射される声が出交るために音声聞き取りにくい、使用時は片手が使えない、などが挙げられる。また、放射線治療などで瘢痕がひどい場合は、振動が声道まで伝わらないため使用することができない。指である程度振動を制御し音身に抑揚をつけるシステムが付属しているものもあるが、会話中に自然に制御するのは難しい。

喉頭摘出者が代替発声法の内訳については図 1.4[29]のようになっており、近年、EL の需要が伸びていることがわかる。これはEL が他の発声法に比べ、習得が容易であり、衛生的であるからと考えられる。よってこれからEL の需要がさらに増加していくと考えられる。

しかし、EL を用いて発生した音声（以下EL 音声）は、上記で述べたように、音質が非常に悪い。そこで、EL 音声を改善する研究がいくつか報告されている[28, 30-38]。鹿野らは、NUM マイクと呼ばれる素子を用いて肉伝導音声を捉え、その音声を通常音声へと変換する手法を提案している[28, 30-32]。中村らは、統計的手法を用いた肉伝導音声を通常音声へと変換する手法について検討している[33, 34]。伊福部、粕谷らは、電気式人工喉頭が発する音源に付与するべきピッチ変化情報について検討している[35-38]。森田らは、声帯の緊張に伴う下咽頭収縮筋の電気インピーダンス変化に着目し、頸部に装着した皮膚電極電位より音声ピッチを推定する手法の検討を行っている[39]。しかし、完全な肉声音声を再現するまでに至っていない。

そこで、基礎的研究としてEL 音声の発声原理にまで遡り、音質劣化の原因を探ることで、EL 音声を明瞭にすることは、EL 使用者とのコミュニケーションを円滑にすることが可能になると考える。これは、補聴器や電話器等に応用することに繋がる。

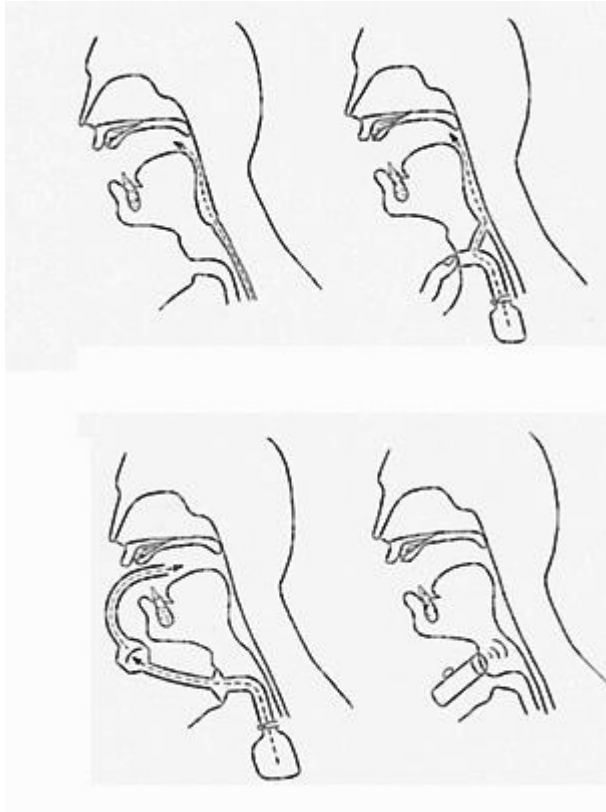


図 1.3 代替発声法（左上が食道発声法，右上がシャント発声法，左下が笛式発声法，右下が EL 発声法）（[28]より転載）

表 1.1 種々の代替発声法の特徴 ([28]より作成)

	駆動力	振動体	習得のしやすさ	手術の有無	音質	調節			使い勝手
						強度	連続	ピッチ	
食道発声	空気	粘膜	×	×	△	△	×	×	○
T-Eshunt	呼吸	粘膜	△	○	△	△	○	×	△
笛式人工喉頭	呼気	ゴム膜	○	×	△	○	○	△	×
		金属板				○	○	×	×
電気式人工喉頭	電気		○		×	×	○	×	×
径皮形		ブザー		×					
埋込み		スピーカ		○					
パイプ挿入式		スピーカ		×					

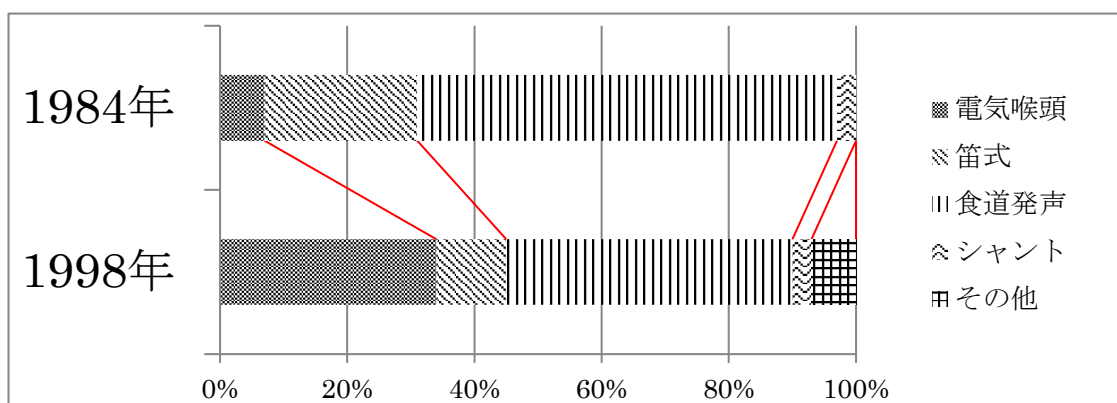


図 1.4 種々の代替発声法の使用率 ([29]のデータより作成)

1.2 本論文の目的

本研究では、世界的にますます増加する高齢者への支援を目的に、高齢者音声や電気式人工喉頭を用いて発声した音声等の劣化音声を対象とし、まず劣化要因や音響的特徴を明らかにし、次に音声の明瞭性を向上させるアルゴリズムの作成を目的とする。これにより音声によるコミュニケーションをより円滑にすることができる。また、本研究の劣化音声の明瞭化は、音声による円滑なコミュニケーションの実現できる装置の一部として活用でき、安定した音声コミュニケーションの提供が可能となる。これにより、音声による対人コミュニケーションを盛んにすることにより、QoL (Quality of Life) の向上や、脳の活性化に繋がると考えられる。

他にも、駅構内等のアナウンス音声の改善や、補聴器の音声処理部分の性能改善、日本人にとって聞き取りやすい英語への変換などにも適応できる。また、音声の明瞭化という分野のみならず、音声認識の性能の更なる改善や、声質変換等にも応用できる。さらには、医療分野における発展も想定でき、例えば、音声から声帯異常検知システム等の実現の可能性もある。

1.2.1 本論文の新規性及び有効性

本研究の特徴は、それぞれの劣化音声の生成過程まで遡り、高齢化による発声器官の劣化がもたらす様々な音響的な変動を定量的に分析し、静的特徴と動的特徴の各々最適なモデル化に基づき、明瞭な音声へと変換することである。これにより劣化音声の明瞭性を改善し、さらに音声認識に用いることで大幅な認識性能の改善を図るものである。これまで、高齢者音声や代替発声音声の音響的特徴を様々な角度から定量的に分析し、明瞭性の改善まで踏み込んで検討された報告は見当たらない。本稿では、日本語の発音において最も重要である母音に着目し特に検討を行う。

1.3 本論文の構成と各章の概要

本論文における各章の内容を以下に示す。本論文は、「高齢者音声に関する研究」、「電気式人工喉頭音声に関する研究」の2つに大別でき、「高齢者音声に関する研究」は第2章～第3章、「電気式人工喉頭音声に関する研究」は第4章～第6章である。

第1章 序論

序論であり、研究の背景、本研究の目的、研究の概要について述べる。

第2章 高齢者音声の明瞭度に関する音響的特徴

本章では、高齢者音声を明瞭度という観点に着目して、聴感印象である明瞭度の高低と物理量である音響的特徴量との相関について、ホルマント周波数の解析を行う。

第3章 ホルマント面積拡張による高齢者音声の明瞭度改善方法の検討

本章では、明瞭度の低い高齢者音声のホルマント周波数の特徴を用いて、実際に低明瞭な高齢者音声に対して、ホルマント面積を広げるように補正を行うアルゴリズムを提案し、高齢者音声の明瞭度の改善を図る。そして、補正を行った音声を聴取実験により、明瞭度改善効果の検証を行う。

第4章 電気式人工喉頭音声の音質改善システム

本章では、電気式人工喉頭音声の問題である「音声不明瞭である」、「電気式人工喉頭の発する振動音自体がノイズになってしまう」、という問題を解決するための電気式人工喉頭の発声方法を基に、新しい発声支援システムの提案を行う。

第5章 種々の音源による電気式人工喉頭の音質改善

本章では、電気式人工喉頭音声の問題である「音声不明瞭である」という問題に着目し、音源である電気式人工喉頭の発する信号を改善することによる音質改善を図り、効果の検証を行う。

第6章 スペクトル包絡のマッピングに基づく音質改善方法の検討

本章では、電気式人工喉頭音声の問題である「音声不明瞭である」という問題に着

目し，音声の包絡情報を，GMMによるマッピングに基づく音質改善方法を提案し，効果の検証を行う．

第7章 結論

結論であり，本研究の成果と今後の課題について述べる．

第2章 高齢者音声の明瞭度に関する音響的特徴

2.1 高齢者音声データベース

本研究では、音声の明瞭度に関して高齢者音声の音響的特徴の差異を解明し、高齢者音声の明瞭性を改善することが目的である。高齢者音声の特徴量を詳細に解析するには、各音韻での特徴を分析する必要がある。そこで、電子技術総合研究所で選定された543単語に関して孤立単語音声を収録した音声データを使用する（付録1）。音声の収録は24kHz サンプリングの12ビットまたは16ビットでデジタル録音されている。表2.1に本研究で使用する高齢者音声データベースの各年代人数を示す。

表2. 1 音声データベースの各年代別人数

年齢	20~29	30~39	40~49	50~59	60~69	70~79	80~	合計
男性 [人]	19	4	10	13	16	18	2	82
女性 [人]	21	6	4	4	18	17	0	70

2.2 被評価話者の選定

2.2.1 明瞭度の低い声とは

高齢者音声は聴感的印象として、「明瞭度が低い声」、「遅い声」、「しゃがれた声」であると言われている[6]。本研究では、高齢者の音声を聞き取りづらくしている最も大きな要因として、「明瞭度が低い声」に着目して、高齢者音声の音響的特徴を解析する。

「明瞭度の低い声」とは、発声中の音素の調音点に達しないまま次の音素を発声してしまうため、各音素の調音が曖昧になり一音一音の区切りがはっきりしない音声の事である。音声は、肺から送られてくる呼気流エネルギーを声帯などの発声器官や、顎関節や舌などの調音器官の複雑な動きによって音響エネルギーに変換し生成される。まず、肺からの呼気が声帯を通過する際に、発声音素が有声音ならば声帯が振動し、無声音ならば振動しない。その後、調音器官で種々の音色が付与される。例えば、“あ”と発声する際には“あ”に対応する舌の位置や口の形がある。これを調音点といい、正しく調音点に達した発声をする事できちんとした音色が付与された音声が発せられる。すなわち、明瞭度の低い声を発声している話者は、調音器官が正常に動いていない事が考えられる。特に高齢者は加齢による筋力の衰えによって調音器官の動きが鈍くなり、正しく調音点に達した発声が難しくなると考えられる。

2.2.2 聴取実験による明瞭度の高低の判別

声の明瞭度に関する高齢者音声の音響的差異を解析するために、評価話者として先行研究[10]で行われた聴取実験によって選定された「明瞭度の低い高齢者」、「明瞭度の高い高齢者」を用いることにする。

先行研究[10]において、評価話者を選定するために、まず聴取実験によって、声の明瞭度度合いの定量化を行っている。表 2.2 にその聴取実験の概要を示す。聴取実験では、男性 10 名、女性 10 名の計 20 名の 20 代成人の被験者に高齢者音声を提示し、評価話者ごとに「5:非常に明瞭度が低い」「4:やや明瞭度が低い」「3: どちらとも言えない」「2:やや明瞭度が高い」「1:非常に明瞭度」の 5 段階評価を付与してもらっている。被評価話者は、60～89 歳の男性 36 名、女性 35 名の高齢者である。提示音声は音素バランスのとれた 543 単語の音声データから高齢者ができるだけ正常に発声している単語を 50 単語選択し、1 秒間隔で連結させたものを評価話者ごとに用意している。聴取実験では、各被験者の判断誤差の影響を低減するため、計 5 回の判断を求めている。集計の際には誤差が大きい 1 回目の評価結果は棄却し、残り 4 回の評価結果に対する 20% 調整平均値を評定値として使用する。それを全被験者で平均化し、評価話者ごとに聴感的な明瞭度度合いを算出している。この聴取実験により高齢者男女計 71 名の音声データに対して聴感的な明瞭度度合いが話者ごとに付加された。なお、評価値が大きいほど、その話者は明瞭度が低いと判断された事になる。

表 2.2 聴取実験の概要

評価方法		明瞭度について 1(高い)～5(低い)の 5 段階評価
評価話者	年齢	60～89 歳
	人数	男性 36 名 女性 35 名
被験者	年齢	20 代
	人数	男性 10 名 女性 10 名
評価回数		5 回
提示音声		音素バランスのとれた 543 単語から 選択した 50 語 (表 2.3) を連結したもの

表 2.3 聴取実験で使用した 50 単語

いいき	あやうい	ぶんうん	えうるし	かぎやく
ういういしい	にえあがる	やにようしょ う	ねみだれがみ	はつひかげ
だぼはぜ	じよそんだん ぴ	うでたてふせ	ぎやくゆにゆ う	ほほえましい
ふしょぞん	ぶあいきょう	ひとづきあい	うじゃうじゃ	でかせぎ
さんわおん	おあいそ	ねっしゃびよ う	がでんいんす い	げんあん
うよきよくせ つ	すっぽぬける	みやくみやく	ししゃごにゆ う	ぼつしゆみ
はいかつりよ う	あおうなばら	ぼちやぼちや	ひえしょう	によやしや
ぎよぎょう	ずひょう	かおいろ	ふによい	おうらさま
ふるいおとす	うおいちば	まえいわい	てんやわんや	やぶへび
きみやく	そりやく	どだい	ひゃひゅひよ	ふひつよう

2.2.3 被評価話者の選定

前節の聴取実験の結果が図2.1である。(a)は高齢者男性、(b)は高齢者女性のものである。縦軸は付加された聴感的な明瞭度度合いで、値が大きくなるほど明瞭度が低い音声であることを示している。横軸は評価話者(高齢者)で、左から明瞭度が高い順に並び替えてある。図2.1より、明瞭度評価値の上位6名の話者を「低明瞭度の高齢者」グループ、下位6名を「高明瞭度の高齢者」グループとして抽出した。ここで、高齢者女性は、高齢者男性に比べて比較的声の明瞭性が高いため、本研究では、評価話者として、高齢者男性における「低明瞭度の高齢者」6名、「高明瞭どの高齢者」6名を用いて高齢者音声の解析を行う。

2.3 高齢者音声の遷移量と調音位置の関係

高齢者音声の聴感印象である明瞭度が低い要因として、加齢による口腔まわりの筋肉の劣化によって、舌や口の形を正しい調音位置まで動かしていないのではないかと考察されている[9-10]。そこでは、明瞭度の低い高齢者音声と、明瞭度の高い高齢者音声では、物理量である遷移量に差があると報告されているが、明確に正しい調音位置までと遷移量との相関までは示されていない。そこでこの節では、物理量である「明瞭度と関係がある遷移量」と、「発音するための調音位置」と相関があるのかを調査する。

2.3.1 データ解析手法

本節の解析に用いる遷移量とは、先行研究[10]で用いられている手法である。概念図を図2.2に示す。この手法は、音声認識の音声のマッチング等に用いられている2つの音素におけるケプストラム係数のユークリッド距離を用いたものであり、ケプストラム係数とは、ケプストラム分析によって求められる値である。そこでまず、ケプストラム分析について述べる。

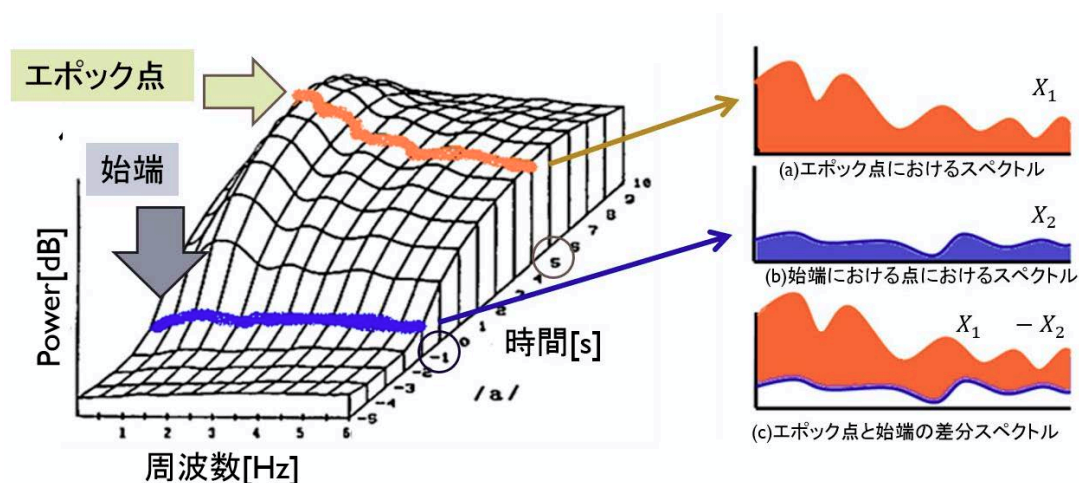


図 2.2 遷移量概念図

ケプストラム分析

ケプストラム分析は音声スペクトルの分析法の一つで、信号のモデル化を行わないノンパラメトリック分析法である。特徴としてはスペクトル包絡と微細構造がケフレンシ領域において近似的に分離することが出来る点である[40]。以下でケプストラム分析について説明する。

信号を $S_{(n)}$ ，そのフーリエ変換を $S_{(e^{j\omega})}$ とすると，ケプストラム(cepstrum)は

$$C_{(m)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |S_{(e^{j\omega})}| e^{jm\omega} d\omega \quad (2.1)$$

と定義される。

m はケフレンシ(quefrensy)と呼ばれ，時間の単位を持ち，音声信号のような実信号に対しては，振幅スペクトルは偶関数であるため，

$$|S_{(e^{j\omega})}| = |S_{(e^{-j\omega})}| \quad (2.2)$$

より

$$C_{(m)} = C_{(-m)} \quad (2.3)$$

が成り立つため，対数振幅スペクトルはケプストラムを用いて次式で表すことが出来る。

$$\log |S_{(e^{j\omega})}| = C_{(0)} + \sum_{m=-\infty}^{\infty} 2C_{(m)} \cos m\omega \quad (2.4)$$

信号 $S_{(n)}$ が二つの信号 $h_{(n)}$ と $e_{(n)}$ の畳み込みで表現される場合，そのフーリエ変換はそれぞれの信号のフーリエ変換の積 $H_{(e^{j\omega})}E_{(e^{j\omega})}$ で与えられる。

この場合の対数振幅スペクトルは

$$\log |S_{(e^{j\omega})}| = \log |H_{(e^{j\omega})}| + \log |E_{(e^{j\omega})}| \quad (2.5)$$

となる。

ケプストラムはそれぞれの信号のケプストラムの和で与えられるため，

$$C_{(m)} = C_{h(m)} + C_{e(m)} \quad (2.6)$$

ここで，二つの信号のケプストラムの分布に重なりがなければ，それぞれの成分が存在するケフレンシ領域のケプストラムを取り出すことで，二つの信号の対数振幅スペクトルを分離することが出来る[41]。このように，ケフレンシ軸上でケプストラムの処理を行うことをリフタリング(liftering)と呼び，この時の処理に用いるフィルタをリフ

タ (lifter) と呼ぶ。

音声の生成過程は、声帯による音源信号に、声道と口からなる声道フィルタの出力としてモデル化することが出来る。20ms 以下の短時間区間内であれば、音源信号や声道フィルタの特性はほぼ一定とみなすことができる。よって、音声信号 $S_{(n)}$ は音源信号 $e_{(n)}$ と線形システムである声道フィルタのインパルス応答 $h_{(n)}$ が畳み込まれた信号と考えることが出来る。

声道フィルタの振幅特性は、声道の構造より周波数上で緩やかに変化するような特性を有しており、インパルス応答 $h_{(n)}$ のケプストラムは主に低ケフレンシ領域に集まる。また、母音などの有声音の場合、音源として準周期的インパルス列を考えると、音源信号 $e_{(n)}$ のケプストラムは音源の周期に相当するケフレンシの整数倍の位置に集中している。

音声認識等においては、2 音間の音色距離として、低次のケプストラム係数のユークリッド距離を次式により計算し、これを用いて音素間マッチングを図ることが多い。

$$d_{spec}(S, T) = \sum_{k=1}^M (c_k^S - c_k^T)^2 \quad (2.7)$$

遷移量

本稿で用いる遷移量においては、(2.7) 式の平方根を求めたものを用いて算出することとする。

$$d_{spec}(S, T) = \sqrt{\sum_{k=1}^M (c_k^S - c_k^T)^2} \quad (2.8)$$

なお、語中の遷移量については隣接する 2 つの音素のエポック点をそれぞれ c^S , c^T とし、語頭の遷移量については発話前の無音区間点を c^T と先頭音素のエポック点を c^S とした。用いる音声データ 24kHz サンプリングに対して、FFT サイズは 512, ケプストラム次数は 30 次までとした。

2.3.2 高齢者音声の遷移量

2.2.3 で選出した 6 名の明瞭度の高い高齢者音声と、6 名の明瞭度の低い高齢者音声について、語中、語頭に分けて遷移量を求める。表 2.4 に遷移量の算出に用いた音声データについて示す。

表 2.4 遷移量の算出に用いた音声データ

対象話者	高齢者	明瞭度が高い 男性 6人 明瞭度が低い 男性 6人
対象音声	音声データベース 各話者 5 4 3 単語	

語中における高齢者音声の遷移量

表 2.4 の音声データを用いて、各単語の語中における先行音素ごとに算出した遷移量を、「高明瞭度高齢者音声」、「低明瞭度高齢者音声」の 2 つの話者グループごとに平均する。各母音における語中遷移量を図 2.3 に示す。

「高明瞭度高齢者音声」と「低明瞭度高齢者音声」の遷移量を比較すると、全ての母音で「高明瞭度」よりも「低明瞭度」の遷移量が小さくなっていることがわかる。特に /a/, /i/, /o/ が大きく低下していることがわかる。一方で、/u/ や /e/ があまり低下していないことがわかる。

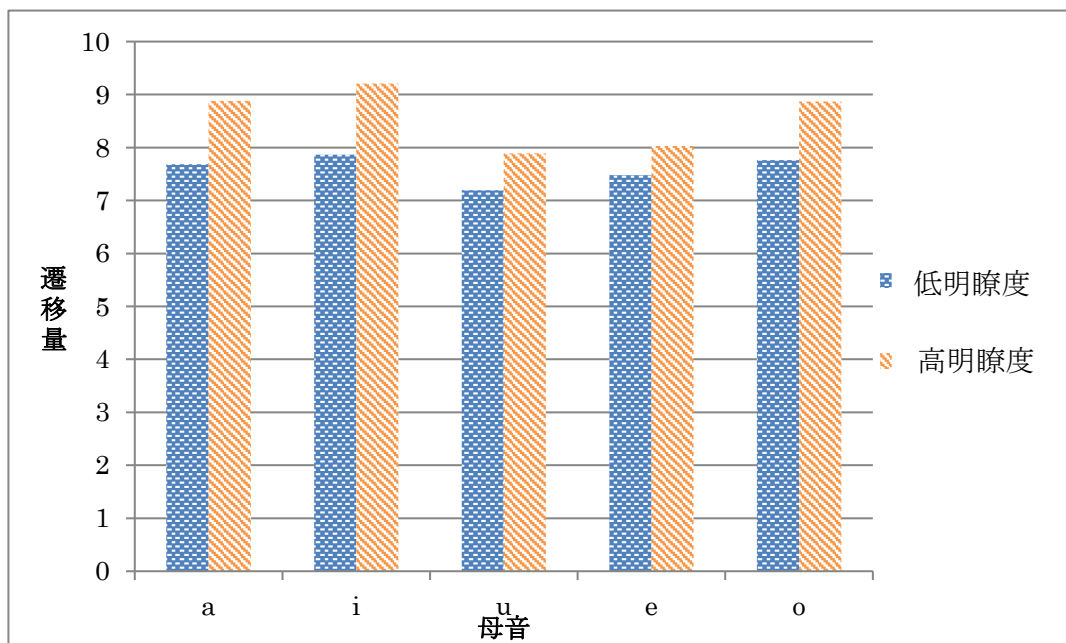


図 2.3 各母音における語中の遷移量

語頭における高齢者音声の遷移量

表 2.4 のデータを用いて、各単語の先頭音素ごとに算出した遷移量を、「高明瞭度高齢者音声」、「低明瞭度高齢者音声」の2つの話者グループで平均する。各母音における語頭遷移量を図 2.4 に示す。

「高明瞭度高齢者音声」と「低明瞭度高齢者音声」を比較すると、語中と同様に、全ての母音で「高明瞭度高齢者音声」よりも「低明瞭度高齢者音声」の遷移量が小さくなっており、特に/i/や/e/が大きく低下している。一方で、/u/があまり低下していないことがわかる。

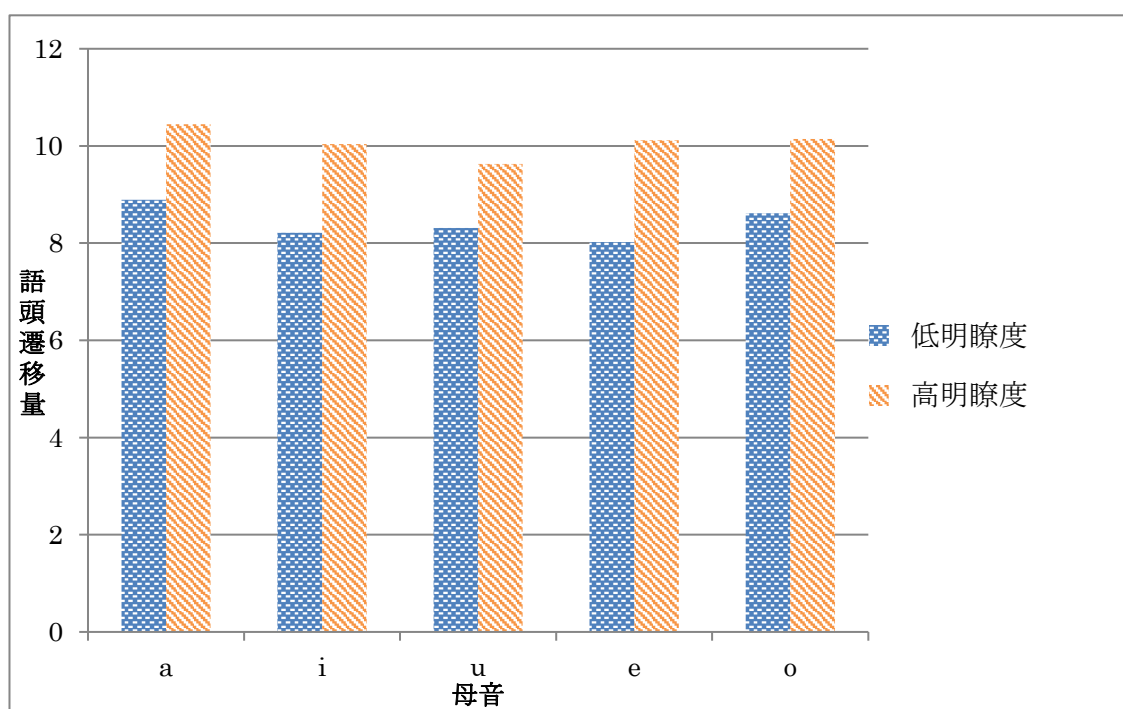


図 2.4 各母音における語頭の遷移量

2.3.3 遷移量差と調音位置の関係

ここでは、2.3.2で求めた遷移量が実際の発声動作と相関があるのかを調べるために、高明瞭度音声と低明瞭度音声の差分値（遷移量差）を求め、調音位置との相関を検討する。

遷移量差

各母音の遷移量の低下の度合いを確認するために、「高明瞭度」、「低明瞭度」の話者グループの差分を、「高明瞭度」を基準とし、「低明瞭度」の遷移量から「高明瞭度」の遷移量を引いた差を遷移量差として算出する。その結果を図2.5、図2.6に示す。

「高明瞭度」と「低明瞭度」との間の遷移量差において、図2.5より、語中では/a/、/i/が負の方向に特に大きくなっている。反対に、/u/、/e/は比較的遷移量差が負の方向に小さいことが確認できる。図2.6より、語頭においては、/i/と/e/の遷移量差が小さい。反対に、/u/は比較的遷移量差が負の方向に小さいことが確認できる。これは、語頭、語中に限らず、口や舌を大きく動かす必要のある母音において、遷移量差が大きくなっていると考えられる。

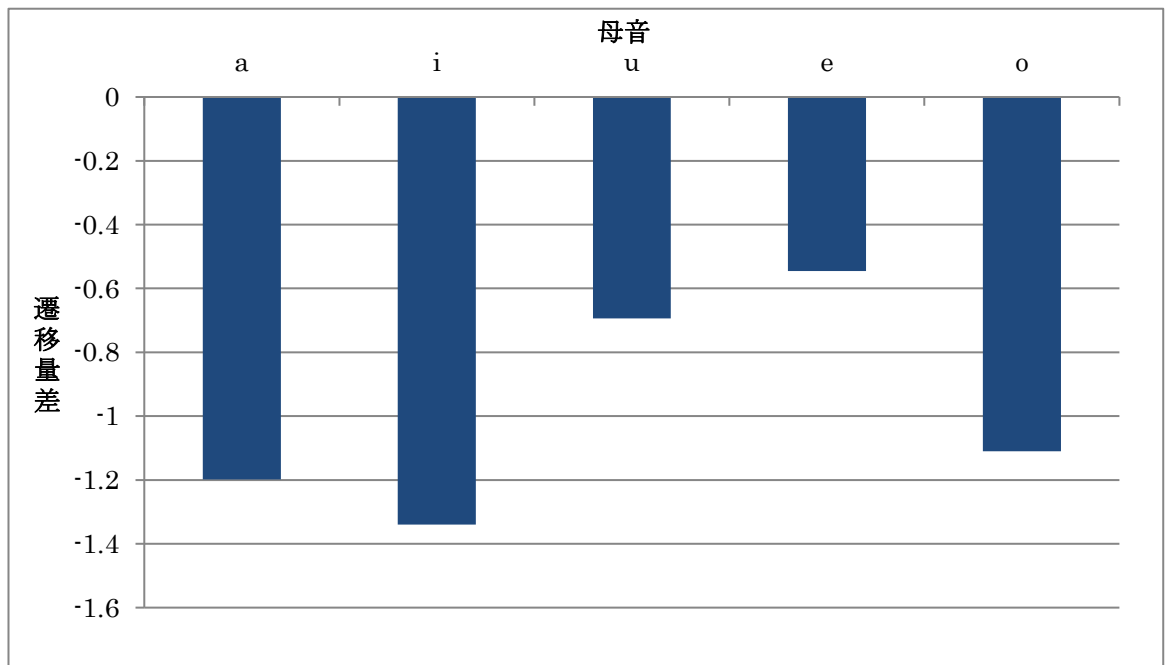


図 2.5 語中における遷移量差

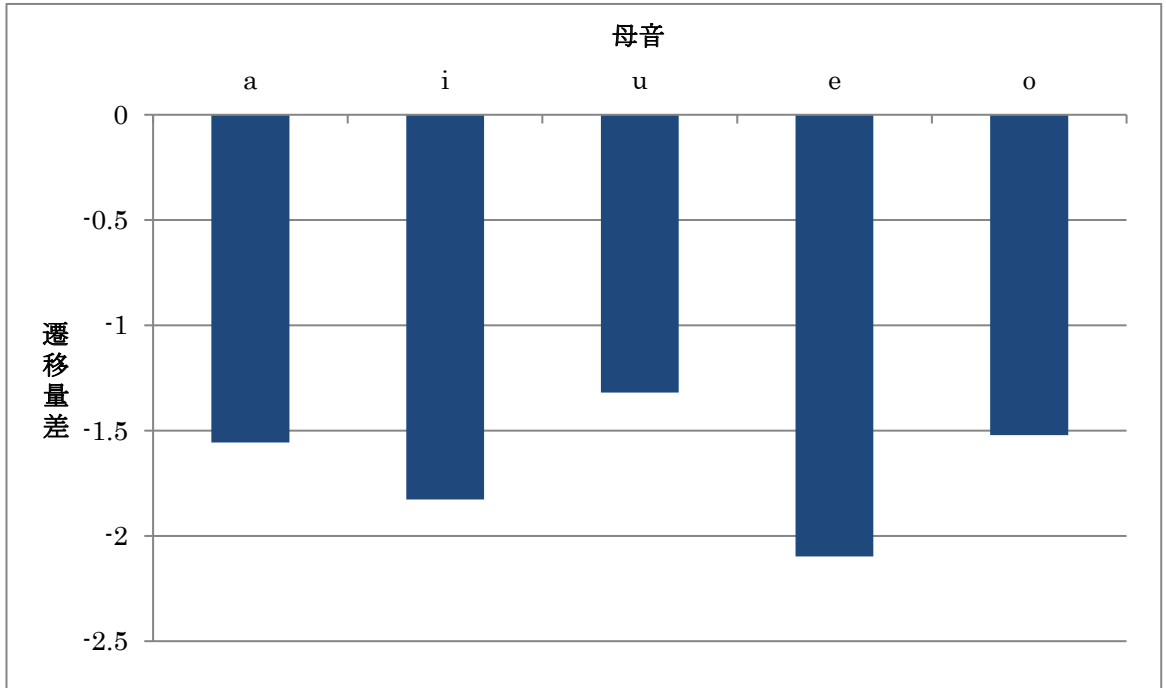


図 2.6 語頭における遷移量差

母音の調音位置

日本語における母音は、/a, i, u, e, o/の5種類であり、有声音（声帯振動を伴う音声）に分類される[42]。母音を発声する際の舌の位置と口の開きに対応した表を表 2.5 に示す。表 2.5 より、/a/は口を最も大きく開けて発音し、/i/は最も前の位置に舌をおき発音していることがわかる。

表 2.5 各母音における口や舌の調音位置（[40]より作成）

	前 ← 舌の位置 → 後			
狭 ↑ 口の 開き ↓ 広	/i/			/u/
		/e/		/o/
			/a/	

遷移量差と調音位置の関係

「高明瞭度音声」と「低明瞭度音声」から求めた遷移量差と、発声時の口の開きや舌の位置との対応関係を調査する。そこで、遷移量差の絶対値を半径とし、表 2.5 の調音位置にプロットし整理したものを、語中と語頭で分け、それぞれ図 2.7, 2.8 に示す。図 2.7 の語中における遷移量差と調音位置の関係をみると、舌を前の方に調音する /i/ と、口を大きく開ける /a/ の遷移量差の遷移量差が大きいことがわかる。図 2.8 の語頭における遷移量差と調音位置の関係をみると、口や舌を大きく動かす /i/, /e/, /a/ の遷移量差の絶対値が大きいことがわかる。これより、低明瞭度の高齢者の遷移量が小さいのは、口や舌をあまり大きく動かすことができず、正しく発音できていないことが要因であることが確認できた。






	前	←	舌の位置	→	後
狭 ↑	/i/ 				/u/ 
口の 開 き			/e/ 		/o/ 
↓ 広			/a/ 		

図 2.7 語中における遷移量差と調音位置



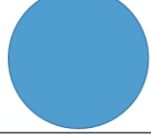


	前	←	舌の位置	→	後
狭 ↑	/i/ 				/u/ 
口の 開 き			/e/ 		/o/ 
↓ 広			/a/ 		

図 2.8 語頭における遷移量差と調音位置

2.4 高齢者音声における F1-F2 平面上の母音間距離

2.3 より、高齢者は加齢による筋力が衰えにより、口や舌をうまく調音できていないことが要因で明瞭度が低くなっていることがわかった。そこで、遷移量に対応する音響的特徴として、口や舌の動きに大きく関係するホルマント周波数について算出することで、高齢者音声の明瞭度について解析を行う。

2.4.1 データ解析手法

本節では、音声の声道情報と関係のあるホルマント周波数について解析を行う。ホルマント周波数とは、音声波形のスペクトル包絡上でみられる複数のピークのことであり、ピーク周波数の低い方から、第1ホルマント (F1)、第2ホルマント (F2)、第3ホルマント (F3) という [41]。音声工学では一般的に、F1 は口の開き具合、F2 は舌の位置に相関があると言われている。ホルマント周波数は、音声波形を線形予測分析 (LPC 分析) によって求められる包絡情報から求めることができる。そこで、線形予測分析について述べた後に、本節で用いるホルマント周波数の計算方法について述べる。

線形予測分析 (LPC 分析)

線形予測とは、離散信号の将来の値をそれまでの標本値の線型和の結合として予測する数学的操作である。線形予測を用いて分析を行うことを線形予測分析 (linear predictive coding : LPC) という。音声に線形予測分析を行うことにより、音声の包絡情報が得られ、この包絡情報は声道特性フィルタとみなすことが出来る。

音声波を周期 $\Delta T[s]$ で標本化した信号を $\{x_t\}$ (t : 整数) とし、現在の標本値 x_t と、これに隣接する過去の p 個の標本化値との間に、次のような線形 1 次結合が成り立つと仮定する。

$$x_t + \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} = \varepsilon_t \quad (2.9)$$

ただし、 $\{\varepsilon_t\}$ は平均値 0、分散 σ^2 のお互いに無相関な確率変数である。ここで

$$\hat{x}_t = \sum (a_i x_{t-i}) \quad (2.10)$$

と置けば、(2.10) 式は

$$x_t = \hat{x}_t + \varepsilon_t \quad (2.11)$$

と書ける. この式は, 現在の標本値 x_t が \hat{x}_t によって, すなわち過去の標本値の線形結合によって予測できることを示している. 言い換えると, 式(2.11)は線形予測係数 (linear predictor coefficient) を $\{\alpha_i\}$ とする線形予測モデルであり, ε_t は線形予測残差 (residual error) ということができる.

ここで,

$$\text{線形予測フィルタ} : F(z) = - \sum_{i=1}^p \alpha_i z^{-i} \quad (2.12)$$

を定義し, $\hat{X}(z), X(z), E(z)$ をそれぞれ $\hat{x}_t, x_t, \varepsilon_t$ の z 変換とすれば,

$$\hat{X}(z) = F(z)X(z) \quad (2.13)$$

$$X(z) = (1 - F(z))^{-1} E(z) \quad (2.14)$$

と書ける. これより, z 変換領域での線形予測モデルのブロック図は, 図2.9のようになる.

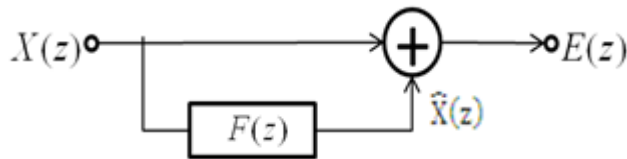


図 2.9 z 変換領域での線形予測モデルのブロック図

音声に線形予測モデルを適用すること, すなわち線形予測分析は, 線形予測フィルタまたは逆フィルタの係数 $\{\alpha_i\}$ を調整して, 出力 ε_t の分散 σ^2 を最小にする方法である. ある定められた区間 $[t_0, t_1]$ での, 標本値 x_t と線形予測値 \hat{x}_t の誤差 ε_t の二乗和が最小になるように (最小二乗推定) 線形予測係数 $\{\alpha_i\} (i = 1, 2, \dots, p)$ を定めることを考えよう. 全二乗誤差を β と置くと,

$$\beta = \sum_{t=t_0}^{t_1} \varepsilon_t^2 = \sum_{t=t_0}^{t_1} \left(\sum_{i=t_0}^p \alpha_i x_{t-i} \right)^2 \quad (2.15)$$

と書ける. ただし, 初期係数である α_0 については1とする. そこで,

$$c_{ij} = \sum_{t=t_0}^{t_1} x_{t-i} x_{t-j} \quad (2.16)$$

と定義すると,

$$\beta = \sum_{i=0}^p \sum_{j=0}^p \alpha_i c_{ij} \alpha_j \quad (2.17)$$

と書ける. これを最小化するためには, $\alpha_j (j = 1, 2, \dots, p)$ に関する偏微分を 0 と置いて解けばよい. これより

$$\frac{d\beta}{d\alpha} = 2 \sum_{i=0}^p \alpha_i c_{ij} = 0 \quad (j = 1, 2, \dots, p) \quad (2.18)$$

が得られ, 予測係数 $\{\alpha_i\}$ は, この p 個の式からなる線形連立 1 次方程式を解いて求めることができる. $c_{ij} (i = 0, 1, 2, \dots, p; j = 1, 2, \dots, p)$ は式 (2.16) によって標本点から計算されるので, 標本点 x_i は, $t_0 - p$ から t_1 までが必要である. この p を LPC 次数といい, LPC 次数の値が包絡のピークの数に相当する [40].

ホルマント周波数

本稿で扱うホルマント周波数は, (2.18) 式より得られる係数の値によって求められるスペクトルのピーク点をホルマント周波数とし, 低次から F1, F2 として扱う.

2.4.2 グループでみた明瞭度の高低による F1-F2 平面上の平均ホルマント周波数

明瞭度の高低の違いによるホルマント周波数の違いの全体的な傾向について述べる。図 2.10, 2.11 は、表 2.3 に示した音声データから、単語の語中・語頭における各母音の第 1, 2 ホルマント周波数をそれぞれ算出し、各話者グループで平均したものである。

図 2.9, 図 2.10 より、声の明瞭度が低い高齢者は、母音/u/以外の母音の第 1, 2 ホルマントが母音/u/に収束しており、5 母音が形成する五角形の面積（ホルマント面積）が小さいことがわかる。これは、2.2.3 でも述べたように、筋力の低下から、口や舌の動きが鈍くなり、口や舌の動きの小さい/u/に、他の母音の声道の形が近づくためだと考えられる。また、語中と語頭を比較すると、前後の音韻の影響を受けている「語中」の方が、ホルマント面積が狭くなっていることがわかる。これは、語中の音声は、前後の音素への移り変わりがより緩慢になってしまうためではないかと考察する。

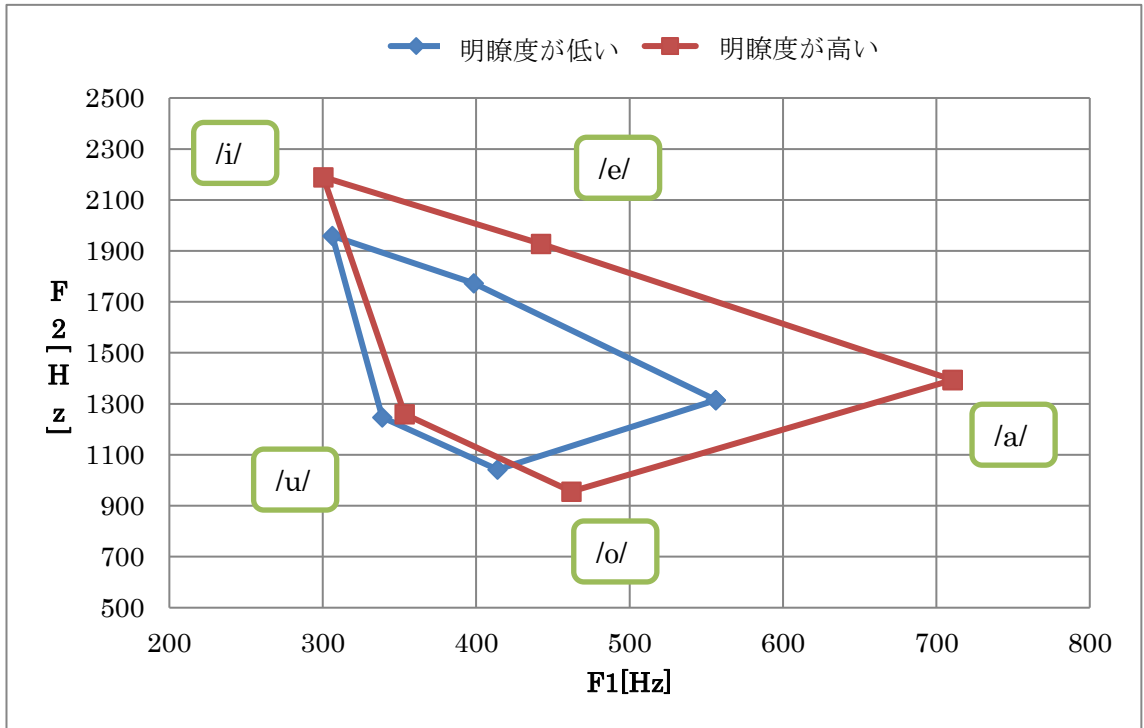


図 2.10 F1-F2 平面上における 5 母音の分布 (語中)

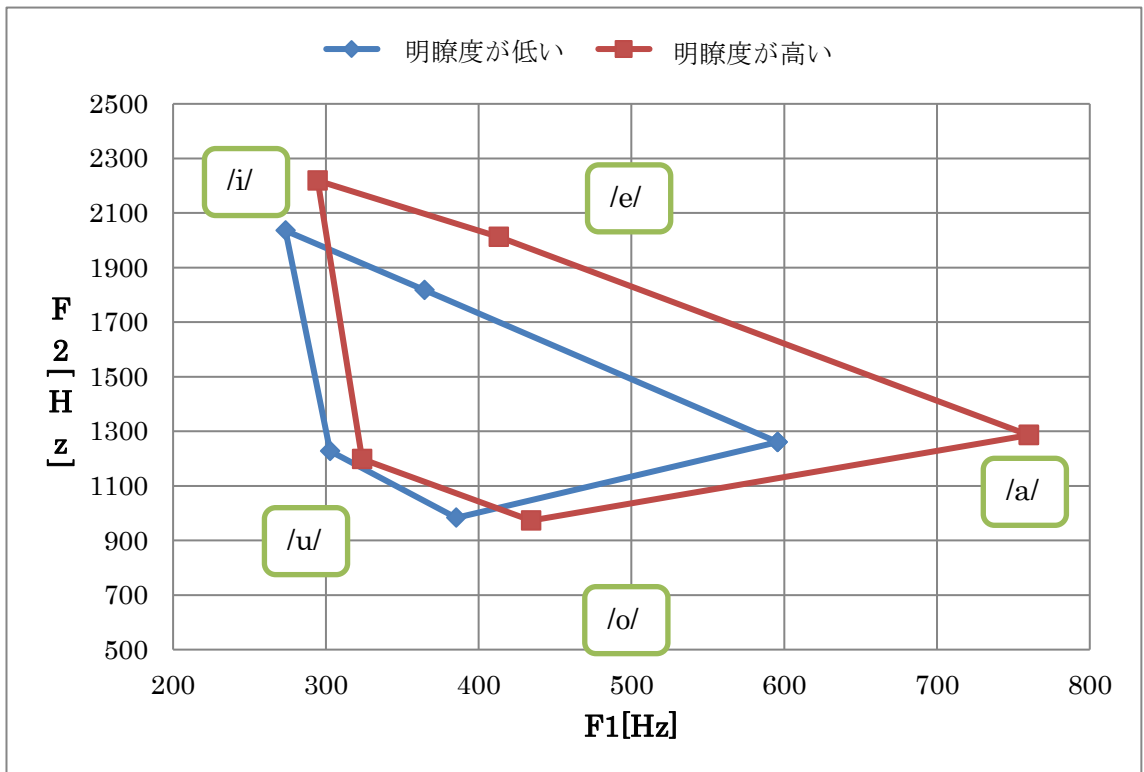


図 2.11 F1-F2 平面上における 5 母音の分布 (語頭)

2.4.3 低明瞭度の高齢者音声における話者ごとの平均ホルマント周波数

ここでは、個人差を分析するために明瞭度の高い高齢者6名の平均ホルマント周波数に対する、明瞭度の低い高齢者6名各々の平均ホルマント周波数を、語頭・語中に分け、それぞれを図2.12～2.23に示す。

図2.12～2.23より、明瞭度の低い高齢者の/u/を中心として、母音/u/以外の各母音が明瞭度の高い高齢者から明瞭度の低い高齢者に向けて、母音/u/以外の各母音が収束していることがわかる。また、声の明瞭度がF1、F2に与える影響もそれぞれ異なっており、話者番号120、125の話者は、特にF1において、明瞭度の高い話者に比べて大きく収束している。一方で、話者番号108、116の話者は、特にF2において、明瞭度の高い話者に比べて大きく収束している。これより、高齢者音声の明瞭度を改善するためにホルマント周波数を補正する場合には、話者ごとに補正する値を変える必要があると考えられる。

また、語中と語頭を比較してみると、同じ話者でも特性が異なることがわかる。これより、語中・語中それぞれで補正方法を変える必要があると考えられる。

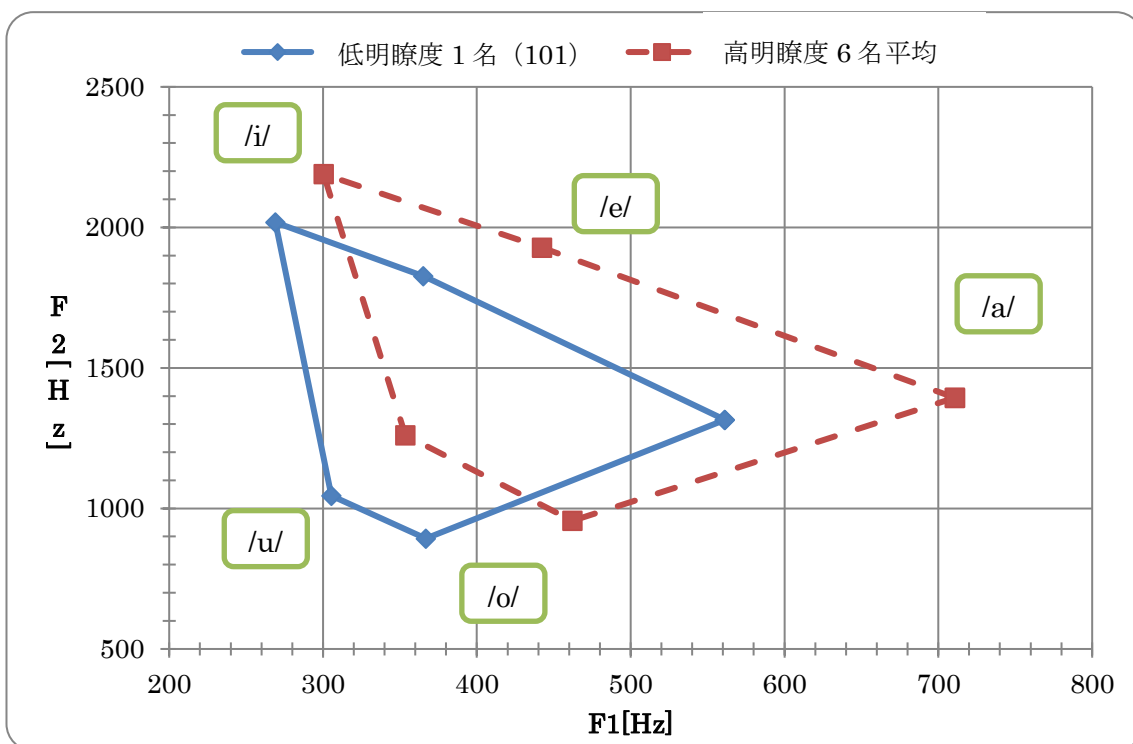


図 2.12 明瞭度の低い高齢者（話者番号：101）の平均ホルマント周波数（語中）

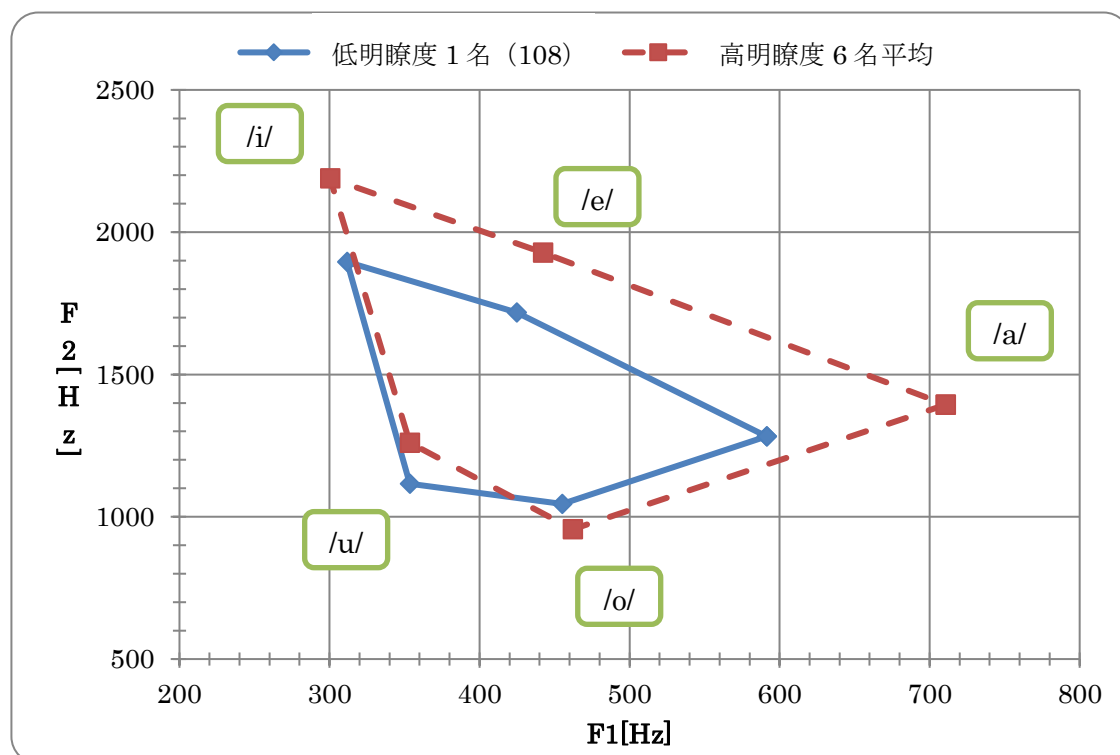


図 2.13 明瞭度の低い高齢者（話者番号：108）の平均ホルマント周波数（語中）

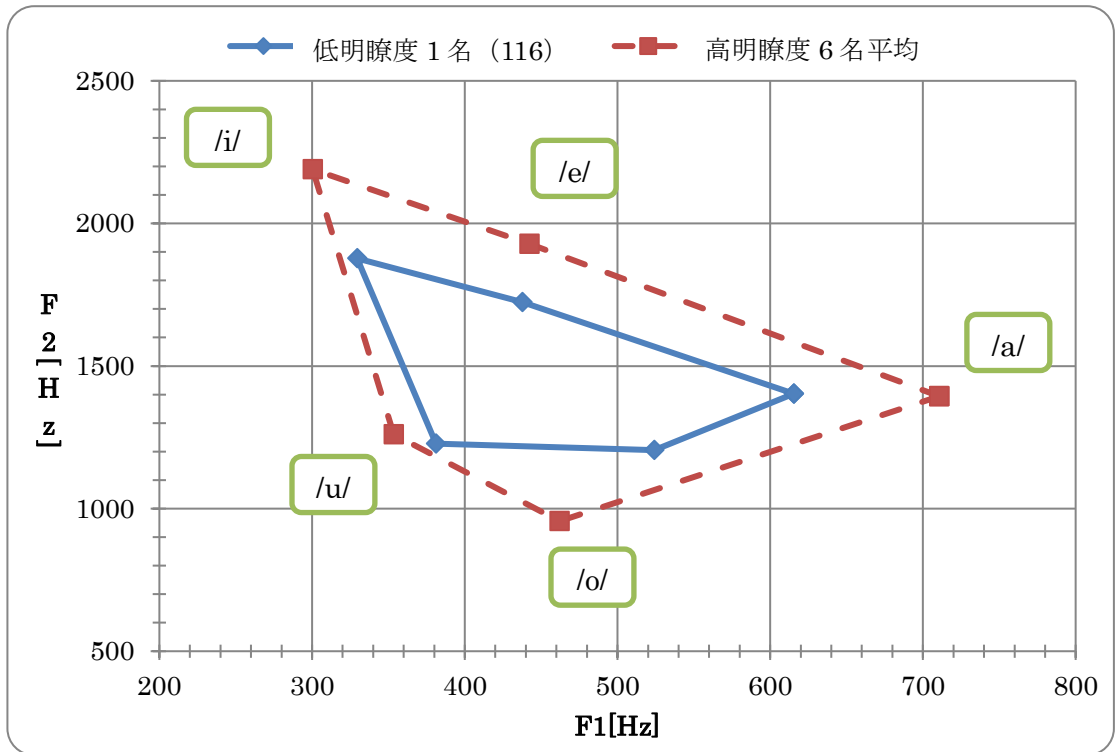


図 2.14 明瞭度の低い高齢者（話者番号：116）の平均ホルマント周波数（語中）

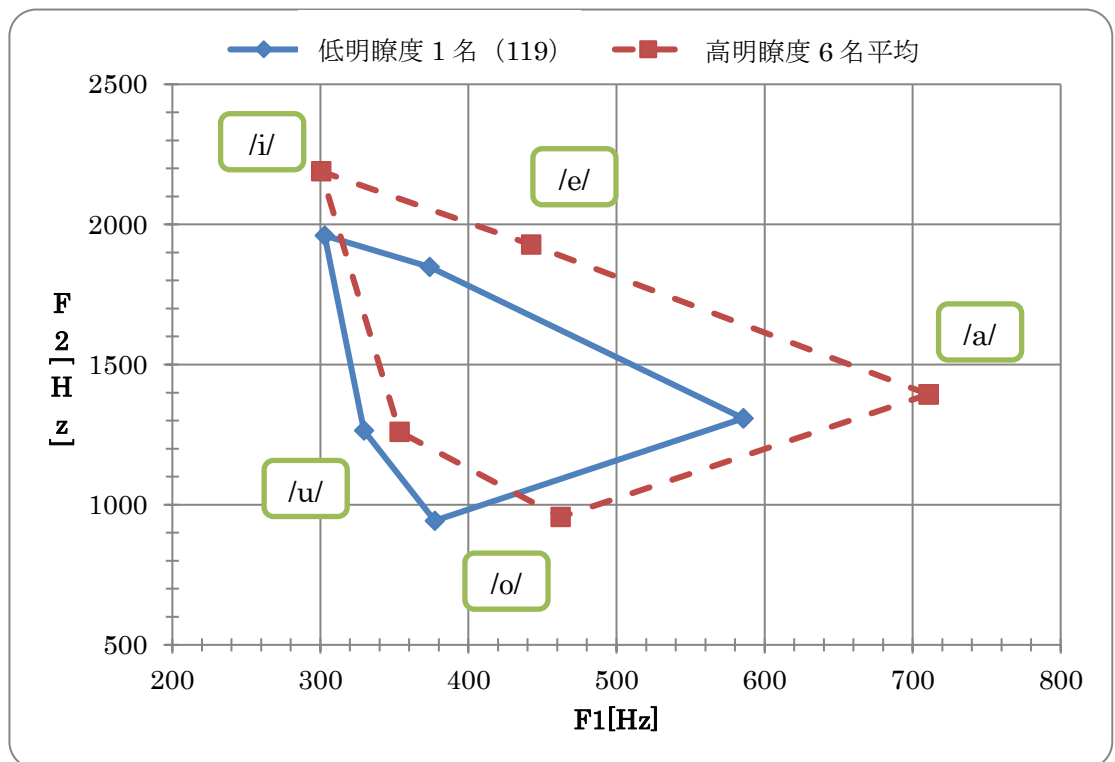


図 2.15 明瞭度の低い高齢者（話者番号：119）の平均ホルマント周波数（語中）

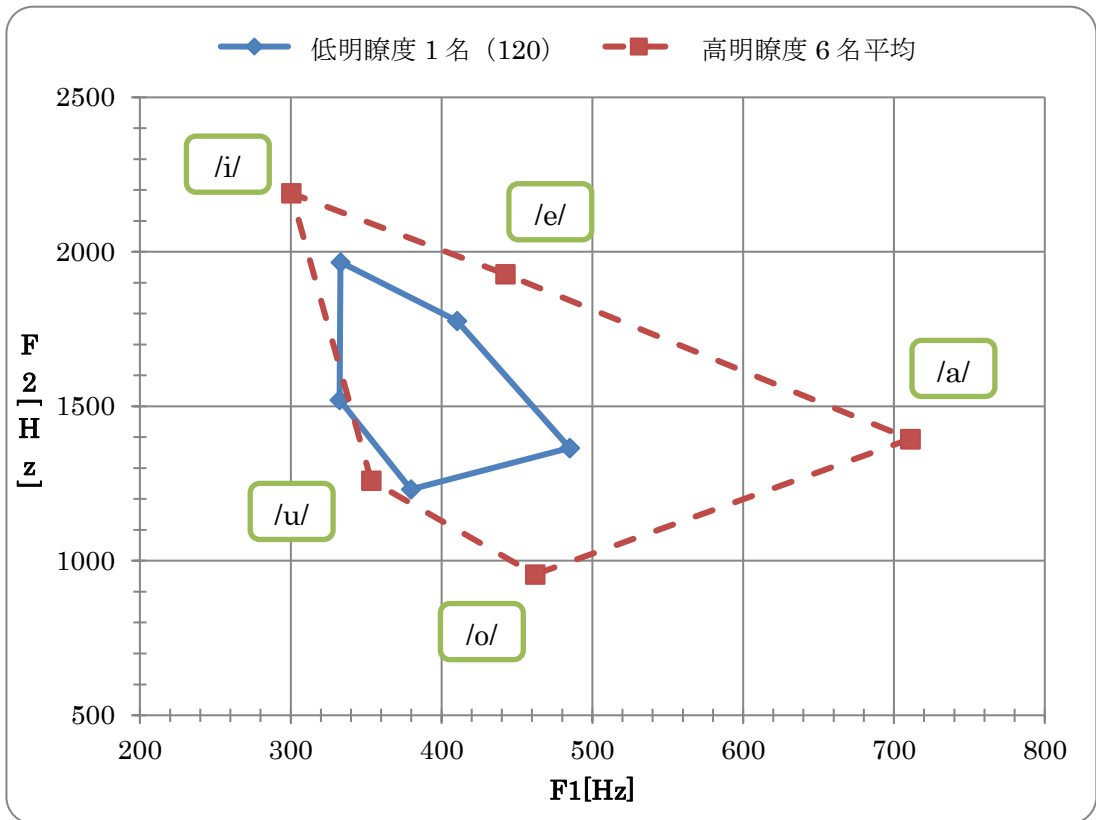


図 2.16 明瞭度の低い高齢者（話者番号：120）の平均ホルマント周波数（語中）

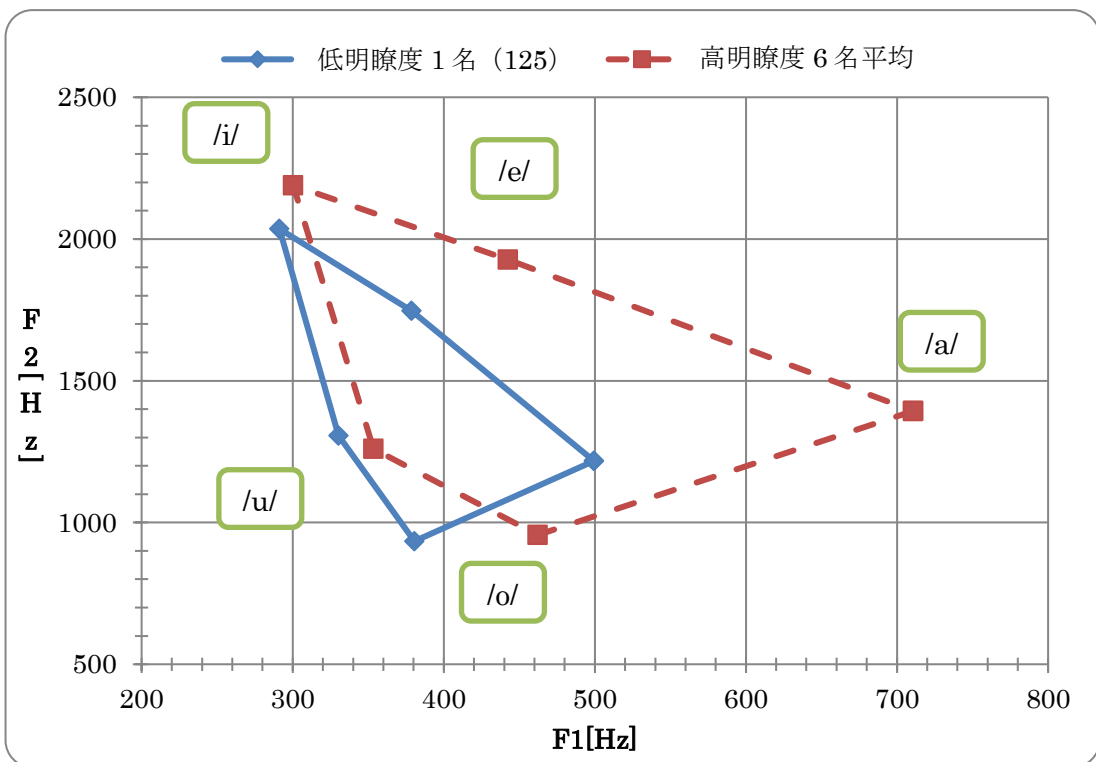


図 2.17 明瞭度の低い高齢者（話者番号：125）の平均ホルマント周波数（語中）

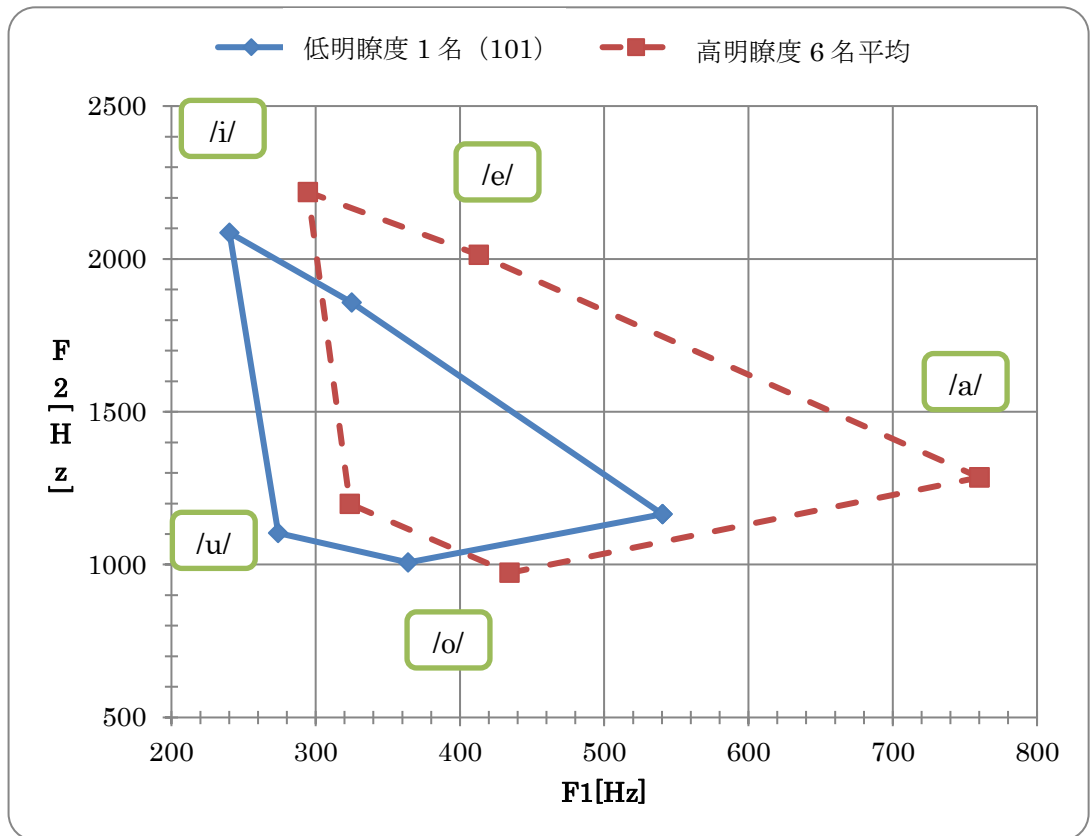


図 2.18 明瞭度の低い高齢者（話者番号：101）の平均ホルマント周波数（語頭）

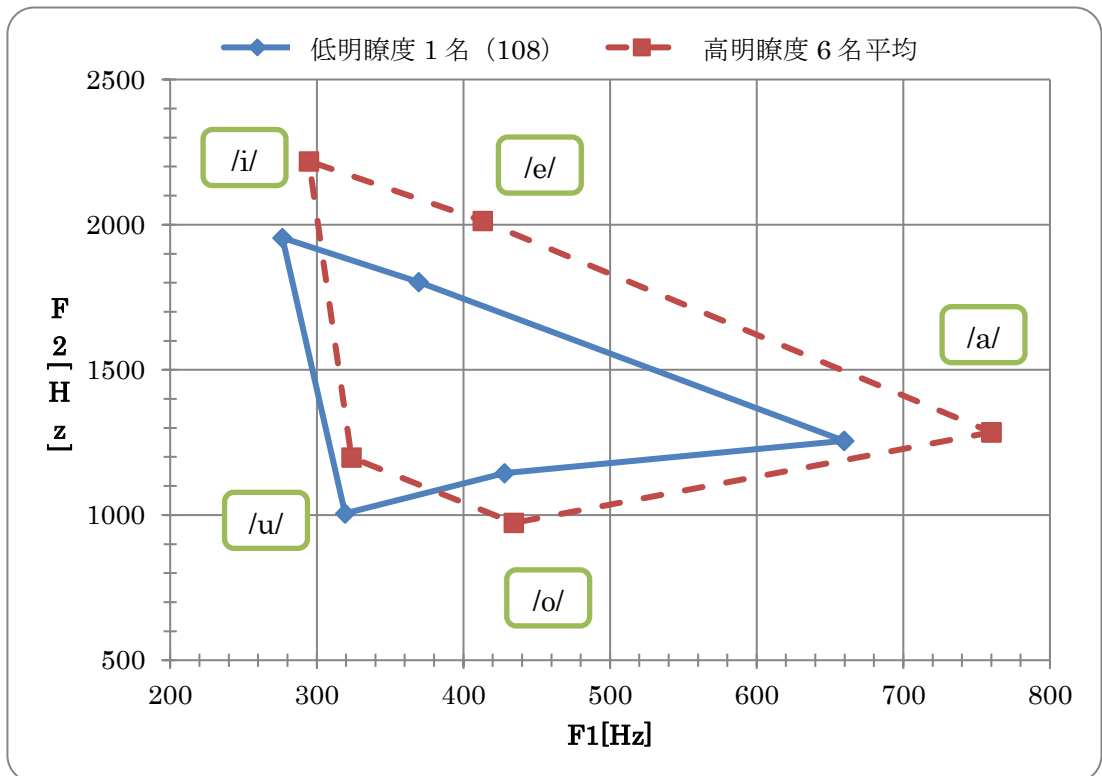


図 2.19 明瞭度の低い高齢者（話者番号：108）の平均ホルマント周波数（語頭）

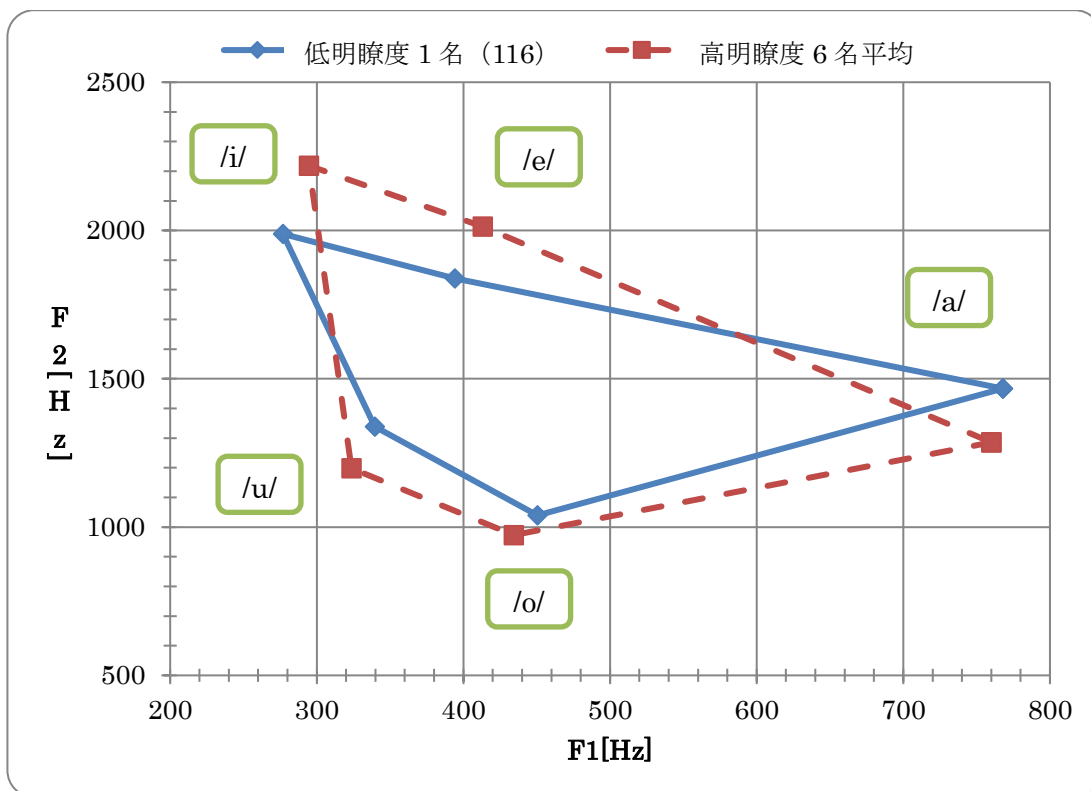


図 2.20 明瞭度の低い高齢者（話者番号：116）の平均ホルマント周波数（語頭）

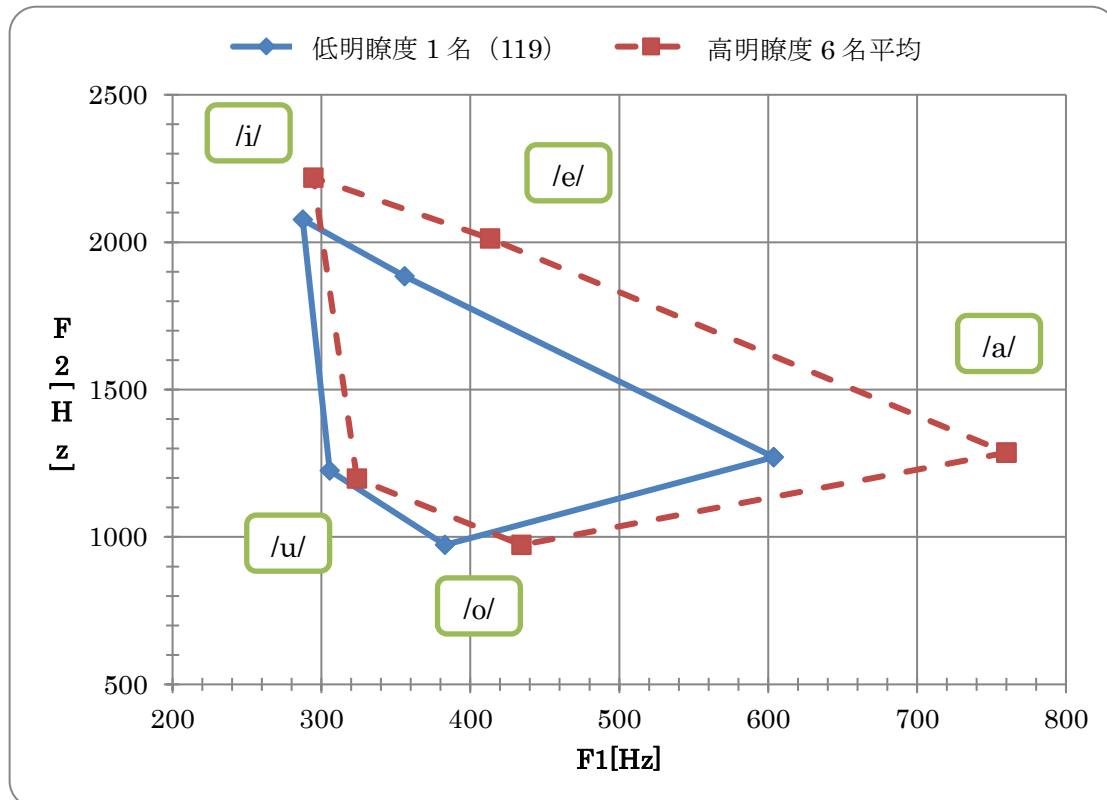


図 2.21 明瞭度の低い高齢者（話者番号：119）の平均ホルマント周波数（語頭）

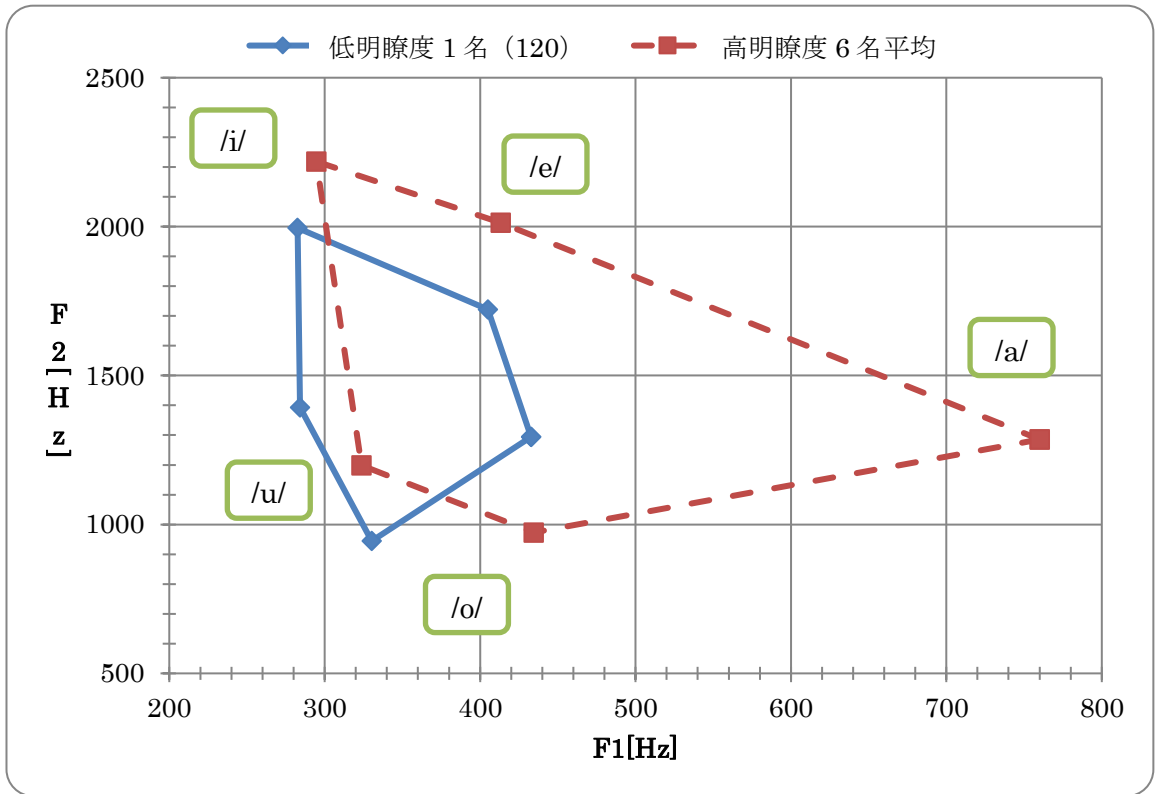


図 2.22 明瞭度の低い高齢者（話者番号：120）の平均ホルマント周波数（語頭）

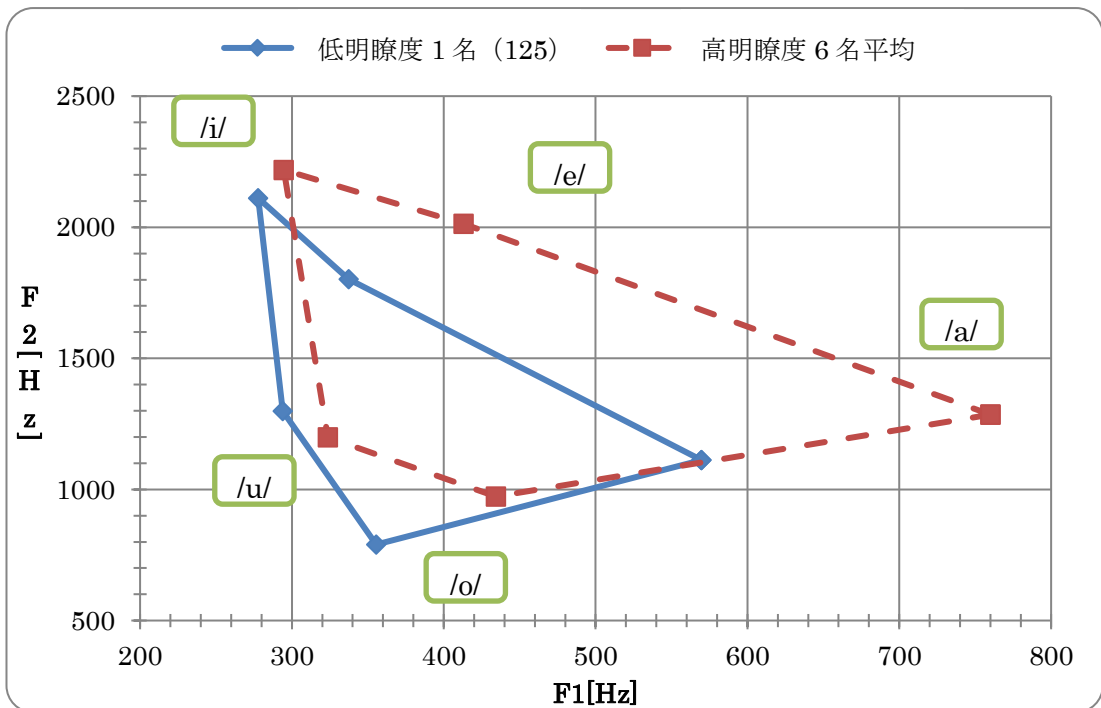


図 2.23 明瞭度の低い高齢者（話者番号：125）の平均ホルマント周波数（語頭）

第3章 ホルモンと面積拡張による高齢者音声の

明瞭度改善方法の検討

前章において、明瞭度別に高齢者音声を分析したところ、低明瞭度話者は、口の動きが緩慢になることにより、口や舌を正しい調音位置で発音することができず、ホルマント面積が小さくなることがわかった。そこで本章では、明瞭度明瞭度の低い高齢者音声のホルマント周波数の特徴を用いて、実際に低明瞭な高齢者音声に対して、ホルマント面積を広げるように補正を行うアルゴリズムを提案し、高齢者音声の明瞭度の改善を図る。そして、補正を行った音声を聴取実験により、明瞭度改善効果の検証を行う。

3.1 ホルマント周波数シフトによる明瞭度改善アルゴリズムの概要

図 3.1 に、提案するホルマント周波数を補正するためのアルゴリズムを示す。このアルゴリズムは、入力音声信号を声帯信号と声道信号に分け、声道信号より求められるホルマント周波数を補正し、元の声帯信号と掛け合わせることによって補正音声を作成する方法である。

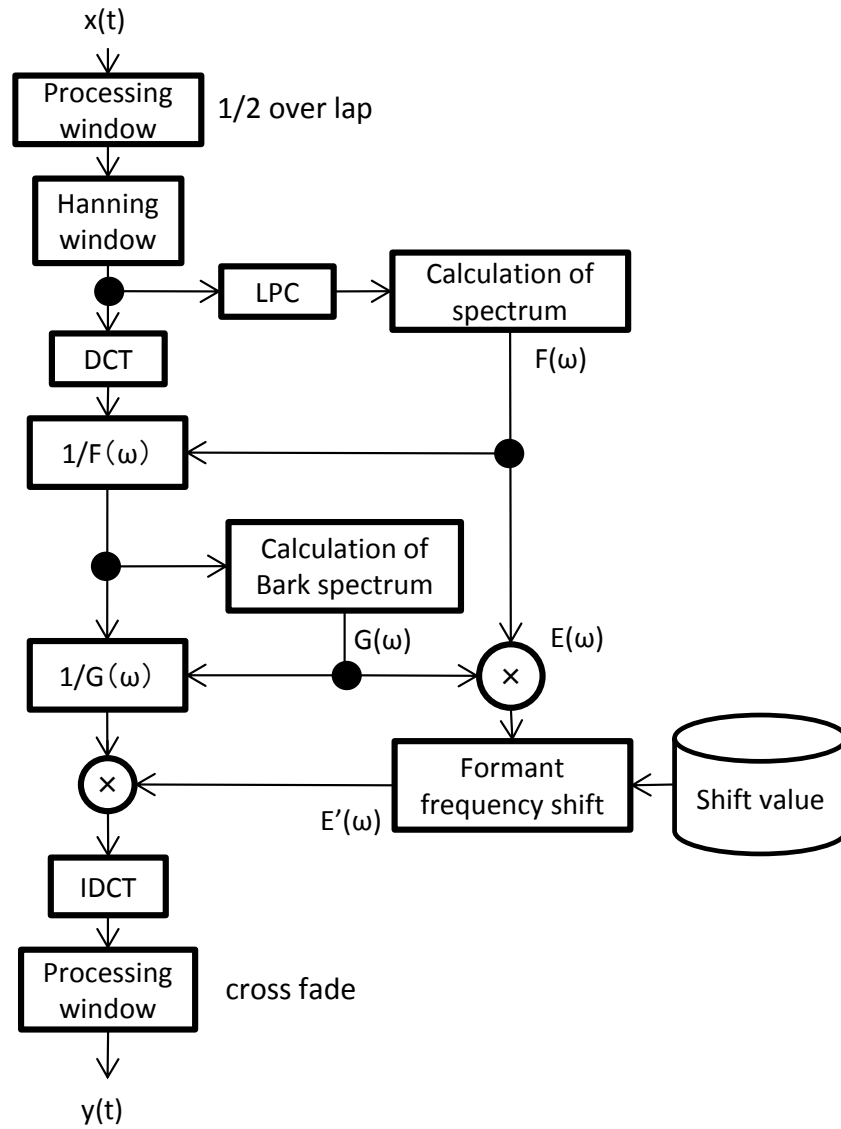


図 3.1, ホルマント周波数を補正するための全体のフローチャート

この処理方法は大きく分けて、以下の5つの処理で構成されている。

- (1) 離散コサイン変換による時間信号から周波数領域への変換
- (2) LPC分析によって求められる包絡による(1)の信号の平坦化
- (3) バーク尺度を用いた(2)のさらなる平坦化
- (4) (2), (3)で平坦化の用いたスペクトルをかけ合わせたものに対しての包絡の補正
- (5) (3), (4)を掛け合わせ、逆離散コサイン変換することによる補正音声の作成

次に、各処理の内容についてそれぞれ述べる。

- (1) 離散コサイン変換による時間領域から周波数領域への変換

入力信号 $x(\omega)$ を、離散コサイン変換によって周波数領域に変換することで $X(\omega)$ を作成する。

$$X(\omega) = \sqrt{\frac{2}{N}} C(\omega) \sum_{t=0}^{N-1} x(t) \cos\left[\frac{(2t+1)\omega\pi}{2N}\right] \quad (3.1)$$

ただし、 $C(\omega)$ は次式である。

$$C(\omega) = \begin{cases} \frac{1}{2} & \omega = 0 \\ 1 & \omega \neq 0 \end{cases} \quad (3.2)$$

- (2) LPC分析によって求められる包絡によるスペクトルの平坦化

一方で、入力信号をLPC分析によって包絡情報 $F(\omega)$ を求める。そこで求めた包絡情報である $F(\omega)$ を用いて $X(\omega)$ を割ることで信号 $A(\omega)$ を求める。

$$A(\omega) = \frac{X(\omega)}{F(\omega)} \quad (3.3)$$

(3) バーク尺度を用いた (2) で求められたスペクトルの更なる平坦化
次に、(2) で平坦化された $A(\omega)$ を、更に音韻性を取り除くために、バーク尺度と呼ばれる聴覚フィルタ [43] を用いて変換した $G(\omega)$ を求める。バーク尺度とは、聴覚の周波数分解能に関する臨界帯域幅の観測から導かれたもので、尺度の範囲は、20, 100, 200, 300, 400, 510, 630, 770, 920, 1080, 1270, 1480, 1720, 2000, 2320, 2700, 3150, 3700, 4400, 5300, 6400, 7700, 9500, 12000, 15500Hz の 24 帯域で表される。また、周波数 f に対するバーク尺度は、(3.3) 式で求めることができる。これは、人の聴覚は低い周波数帯域ほど周波数分解能が高く、高い周波数帯域ほど周波数分解能が低いという特性を活かした臨界帯域になっている。

$$\text{BarkScale} = 13 \tan^{-1}(0.00076f) + 3.5 \tan^{-1}\left(\left(\frac{f}{7500}\right)^2\right) \quad (3.4)$$

(3.3) 式の f は、 $A(\omega)$ の帯域幅における周波数を表しており、この式によって変換された周波数帯域までの値を一律に扱うことで、 $G(\omega)$ を算出する。

$$\begin{aligned} G(\omega) = & \frac{1}{t_1} \sum_{n=1}^{t_1} A(n) + \frac{1}{t_2 - t_1} \sum_{n=t_1+1}^{t_2} A(n) \\ & + \frac{1}{t_3 - t_2} \sum_{n=t_2+1}^{t_3} A(n) \\ & + \dots \\ & + \frac{1}{t_{24} - t_{23}} \sum_{n=t_{23}+1}^{t_{24}} A(n) \end{aligned} \quad (3.5)$$

(t_1, t_2, \dots, t_{24}) は、24 ある臨界帯域までの周波数に含まれる個数を表しており、その区間の値を平均したものをそれぞれ足し合わせたものを $G(\omega)$ として扱う。(3.4) 式によって求めた $G(\omega)$ を用いて、次式によって更にスペクトルの平坦化を行う。

$$A'(\omega) = \frac{A(\omega)}{G(\omega)} \quad (3.6)$$

- (4) (2), (3) で平坦化に用いたスペクトルをかけ合わせたスペクトルに対してホルマント周波数の補正

次に、包絡の平坦化に用いた $G(\omega)$ と $F(\omega)$ をかけ合わせた $E(\omega)$ を計算する。これを、今回は音声包絡として扱う。

$$E(\omega) = G(\omega) \cdot A(\omega) \quad (3.7)$$

(3.7) 式を用いて計算した $E(\omega)$ に対して、ホルマント周波数を求め、明瞭度を改善する補正を行う。補正方法の処理フローを図 3.2 に示す。まず、 $E(\omega)$ から、ピーク点とボトム点を算出する。算出したピーク点を低周波数の方から $F1, F2$ と判断する。他のホルマント周波数に影響を与えないために、ボトム点は固定したまま、ピーク点まわりのみの補正を、ピーク点から前後のボトム点までのデータを伸縮することによって行う。補正のイメージ図を図 3.3 に示す。この処理によって求めた包絡情報を $E'(\omega)$ とする。この補正には、補正值 (Shift Value) を用いて、その値分のみをシフトすることによって補正を行う。補正量については次節で検討する。

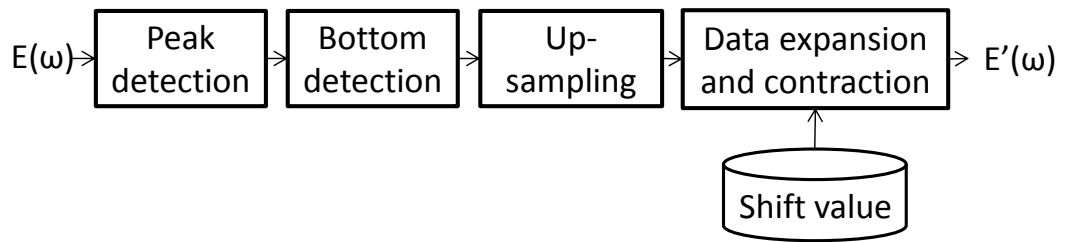


図 3.2 ホルモン周波数シフトブロックにおける処理フロー

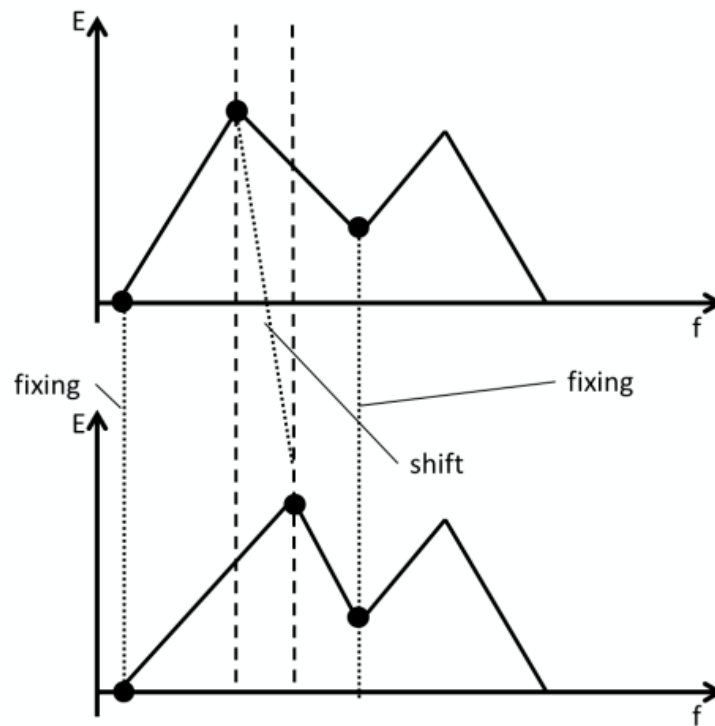


図 3.3 データ伸縮によるホルモン補正処理イメージ

(5) (3) と (4) のスペクトルをかけ合わせて逆離散コサイン変換することによる補正音声の作成

補正を行った音声包絡情報である $E'(\omega)$ と、2重で平坦化された $A'(\omega)$ を掛け合わせることで補正音声スペクトル $Y(\omega)$ を求める。

$$Y(\omega) = E'(\omega) \cdot A'(\omega) \quad (3.7)$$

(3.7) 式によって求めた $Y(\omega)$ を逆離散コサイン変換することにより、補正音声 $y(\omega)$ を作成する。

$$y(\omega) = \sqrt{\frac{2}{N}} \sum_{t=0}^{N-1} C(\omega) Y(\omega) \cos \left[\frac{(2t+1)\omega\pi}{2N} \right] \quad (3.9)$$

ただし、 $C(\omega)$ は次式である。

$$C_{\omega} = \begin{cases} \frac{1}{2} & \omega = 0 \\ 1 & \omega \neq 0 \end{cases} \quad (3.10)$$

このように (1) ~ (5) の5つのブロックを用いて、ホルマント周波数をシフトすることにより、高齢者音声の明瞭化を図る。上記の処理はフレーム単位で行うのだが、母音区間を選択する必要がある。本稿においては、高齢者音声データベースにあるラベル情報に記載されている情報を基に母音区間のみの補正を行う。また、急に補正を加えると、音声の滑らかさがなくなってしまうため、前後3フレームを用いて、線形的に徐々に補正を行うこととした。

次節で、どの程度ホルマント周波数をシフトするのかについて述べる。

3.2 ホルマント周波数のシフト量について

ここでは、3.1 節で述べた補正アルゴリズムに用いる補正量について検討する。2 章で分析を行った結果、個人に合わせたホルマント周波数の補正を行うこと、また、F1-F2 平面上における 5 母音が構成する五角形を、/u/を軸に面積を広げるように補正を行うことが有効であると考えられる。

そこで本稿では、以下の 2 つの補正方法について効果の検証を行う。

- (1) 補正を行う低明瞭度話者 1 名のホルマント周波数を、高明瞭度話者の平均ホルマント周波数にシフトする方法（補正値が既知の場合）

補正を行う低明瞭話者のホルマント周波数に合わせた値で補正を行う方法である。この方法は、補正を行う話者の音声の平均ホルマント周波数が既知である場合において有効であり、理想的な補正を行うことができると予想される。補正イメージを図 3.4 に示す。

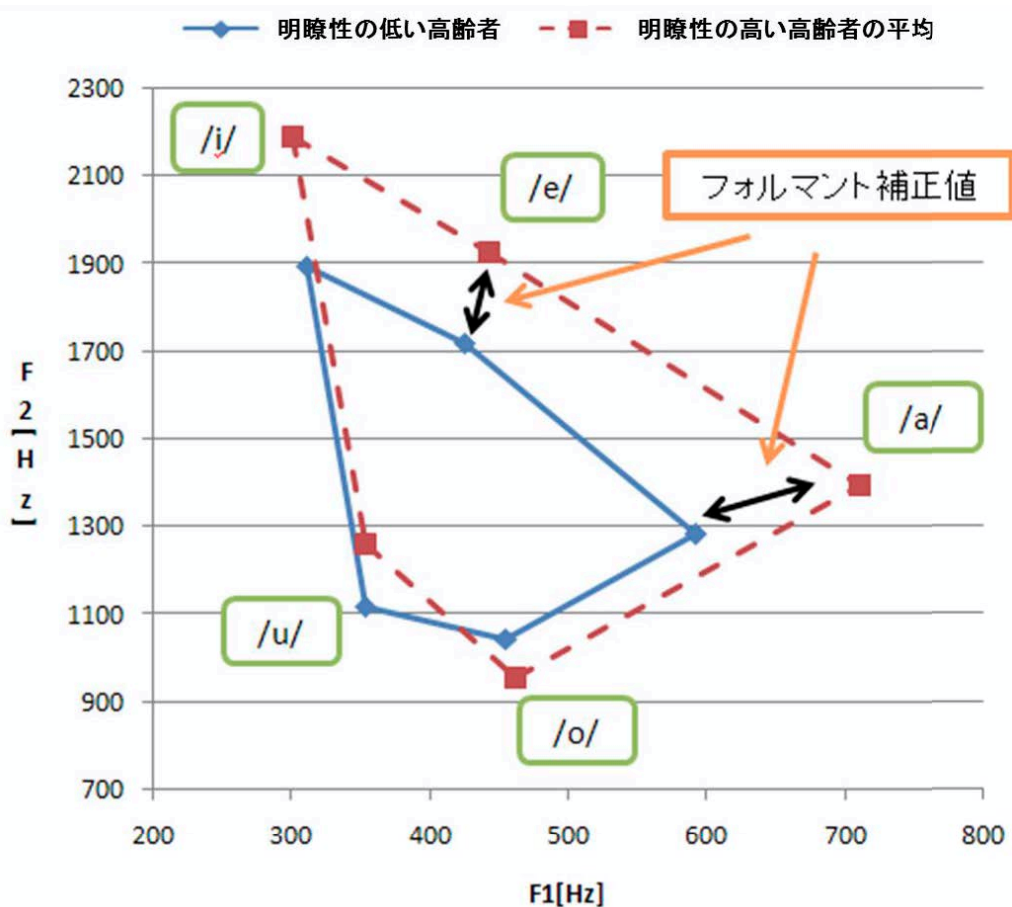


図 3.4 ホルマント補正値の算出方法の概要

(2) 補正を行う低明瞭度話者のホルマント周波数を，ホルマント面積を拡張するように一定倍率広げる方法（補正値が未知の場合）

補正を行う低明瞭度話者のホルマント周波数が構成する F1-F2 平面上の五角形について一定倍率で補正を行う方法である．この方法は，補正を行う低明瞭度話者にホルマント周波数に依存しないため，未知の場合においては有効と予想される．補正イメージを図 3.5 に示す．

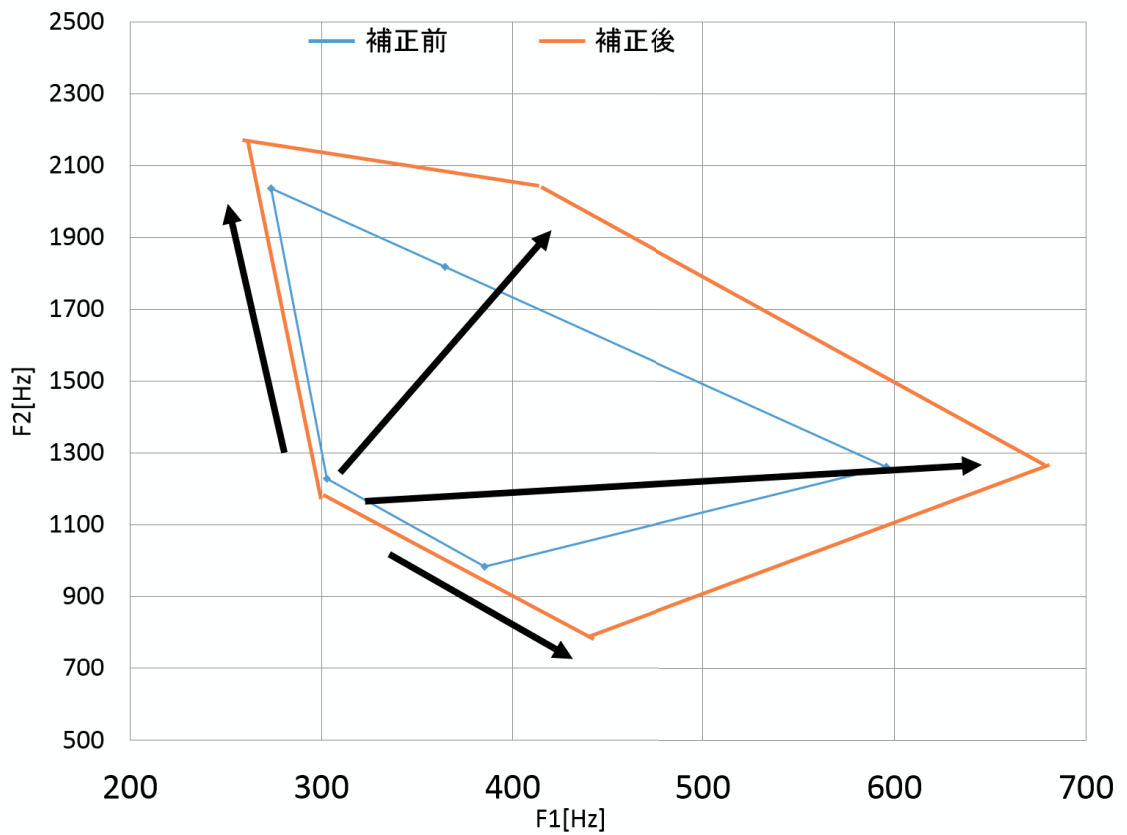


図 3.5 ホルマント周波数を倍率にする場合のイメージ図

3.3 シェッフエの対比較法を用いた聴取実験による明瞭度改善効果の検証

ここでは、3.1 で述べた方法を用いて高齢者音声の明瞭度改善に効果があるのかを、3.2 で述べた、(1) 補正值が既知の場合、(2) 補正值が未知の場合の2つの場合における検証を、聴取実験により検証する。聴取実験には、シェッフエの対比較法を用いた解析を行い、主観評価による検証を行う。

3.3.1 データ解析方法

今回の主観評価実験により得られた結果は、シェッフエの対比較法を用いて分析を行う。まず、分析方法について説明する。

シェッフエの対比較法

シェッフエの対比較法では、試料間の順位だけでなく、試料間の差の程度を知ることができる。本実験では被験者が少数であり、さらに、提示された音声に対する判断の往復を許しているため、シェッフエの「中屋変法」を用いて、解析を行った[44]。

以下ではシェッフエの対比較法「中屋変法」について説明する。

被験者 k が、試料 A_i と A_j をこの順序で比較したときの採点を x_{ijk} とするとき、その採点の内容を次のような構造で考える。

$$x_{ijk} = (\alpha_i - \alpha_j) + (\alpha_{ik} - \alpha_{jk}) + \gamma_{ij} + \varepsilon_{ijk} \quad (3.1)$$

各記号の意味は表 3.1 に示す。

この構造では、それぞれの母数の推定値は以下の式によって与えられる。

$$\text{平均嗜好度} : \alpha_i = \frac{1}{tN} x_{i \cdot \cdot} \quad (3.2)$$

$$\text{嗜好度の個人差} : \alpha_{ik} = \frac{1}{t} x_{i \cdot k} - \alpha_i \quad (3.3)$$

$$\text{組み合わせの効果} : \gamma_{ij} = \frac{1}{N} x_{ij \cdot} - (\alpha_i - \alpha_j) \quad (3.4)$$

各効果の有意性を示すために、分散分析を行う。

各効果の平方和 S は以下の式により求まる。

$$S_\alpha = \frac{1}{tN} \sum_i x_{i \cdot \cdot}^2 \quad (3.5)$$

$$S_{\alpha(B)} = \frac{1}{t} \sum_k \sum_i x_{i \cdot k}^2 - S_{\alpha} \quad (3.6)$$

$$S_{\gamma} = \frac{1}{N} \sum_i \sum_{j>i} x_{ij \cdot}^2 - S_{\alpha} \quad (3.7)$$

$$S_e = S_T - S_{\alpha} - S_{\alpha(B)} - S_{\gamma} \quad (3.8)$$

$$S_T = \sum_k \sum_i \sum_{j>i} x_{ijk}^2 \quad (3.9)$$

これら S の意味とそれぞれの自由度を表 3.2 に示す. 式(3.5)から式(3.9)によって計算した平方和と, 表 3.2 に示した自由度を用いて, 各要因の不偏分散を算出する. そして, 誤差の不偏分散で, 各要因の不偏分散を割った値 F_0 を求め, それが文献[43]の数表 7 の F の値以上であれば有意とみなす.

分散分析の結果により主効果が有意となれば, 次にどの α_i と α_j の間に有意性があるか調べる. そのために, ヤードスティック Y を求める.

$$Y = q \sqrt{\frac{\hat{\delta}^2}{tN}} \quad (3.10)$$

q は, スチューデント化された範囲である. $\hat{\delta}^2$ は誤差の不偏分散の値を用い, $\alpha_i - \alpha_j$ の信頼区間を次式によって求める.

$$(\alpha_i - \alpha_j) - Y \leq \alpha_i - \alpha_j \leq (\alpha_i - \alpha_j) + Y \quad (3.11)$$

この信頼区間が, +側と-側にまたがっている場合, α_i と α_j に有意な差があると言えない. 信頼区間が+側にあるならば, $\alpha_i > \alpha_j$ であり, -側にあるならば, $\alpha_i < \alpha_j$ と考える.

以上の方法に沿って, 各音声間の平均嗜好度の差 ($\alpha_i - \alpha_j$) とその信頼区間の算出を行う.

表 3.1 各記号の意味

記号	意味
t	試料数
i	先に提示された試料.
j	後で提示された試料.
N	被験者の総数
k	被験者の番号. $k = 1, 2, \dots, N$
α_i, α_j	試料 A_i, A_j に対して, 被験者全体がもっている平均的な嗜好度. 解析を容易にするために, $\sqrt{\sum_{i=1}^t \alpha_i} = 0$ と設定する.
α_{ik}, α_{jk}	試料 A_i, A_j に対して, 被験者 k がもっている嗜好度の個人差.
γ_{ij}	組み合わせの効果.
ε_{ijk}	誤差.

表 3.2 各平方和 S の意味と自由度

平方和	意味	自由度
S_α	主効果	t-1
$S_{\alpha(B)}$	主効果×個人	(t-1)(N-1)
S_γ	組み合わせ効果	(t-1)(t-2)/2
S_e	誤差	(t-1)(t-2)(N-1)/2
S_T	総平方和	Nt(t-1)/2

3.3.2 シフト量が既知の場合における補正効果

聴取実験条件

ここでは、補正を行う低明瞭度高齢者の平均ホルマント周波数が既知の場合における、理想値を用いて補正を行った場合の検証結果について述べる。被評価話者として、明瞭度の低い高齢者6名の中から最も声に明瞭度が低いと思われる話者を認知で1名選出した。評価単語は、データベースである543単語の中から2~3種類の母音を含む10種類の単語音声を選択した。表3.3に評価単語を示す。それぞれの単語について、「原音声」、ホルマントを補正してから再合成した「補正音声」、さらに、音声を再合成する際の雑音の明瞭性への影響を低減するためにホルマントを補正せずに再合成した「無補正音声」の3つを用意した。補正音声を作成した条件について表3.4にまとめる。

この被評価話者におけるホルマントの補正值については、図2.12、図2.13から、語中、語頭それぞれにおいて、被評価話者1名の平均F1、F2と明瞭度の高い6名の平均F1、F2をそれぞれ算出し、差分値を補正值とした。表3.5に使用した補正值を示す。

聴取実験の被験者には、成人(20代)10名を用いた。各被験者に各単語の「原音声」、「無補正音声」、「補正音声」の中から選択した2つの音声をA音声、B音声として提示し、明瞭性について「-2:A音声の方が明瞭度が高い」、「-1:A音声の方がやや明瞭度が高い」、「0:同程度」、「1:B音声の方がやや明瞭度が高い」、「2:B音声の方が明瞭度が高い」の5段階で評価してもらう。被験者1人に対して、1単語につき3種類の組み合わせについて評価してもらい、それを単語数分実施した。被験者の判断誤差を低減するために、被験者には計5回の評価を行ってもらい、集計の際には誤差の大きい1回目の評価結果を棄却し、残りの4回の評価結果を使用した。この聴取実験により、相対的な明瞭性の度合いが各音声に付加される。聴取実験は、周りの雑音が音声の聞き取りに悪影響を与えないように、グラスウール製壁面パネルと防音性カーペットに囲まれた簡易防音室内で行った。防音室内には、ノートパソコン1台とヘッドホンを用意した。聴取実験には、Praatをツールとして用いた。聴取実験の際にノートパソコンのディスプレイに表示される画面を図3.6に示す。中央にA、Bと書かれた黄色のボタンを押すと、それぞれ異なる音声が流れる。下部にあるボタンをクリックすることで、-2~2までの5段階評価を行う。下部にある5つのボタンのどれかを選択すると、その下に[OK]ボタンが現れ、そのボタンをクリックすることで次の評価に移ることができる。

表 3.3 評価単語

あおあお	みやくみやく	ぱちぱち	めちやめちや	ぎょくほ
おいおい	ぺこぺこ	ういういしい	ふてぶてしい	いめーじ

表 3.4 補正音声作成条件

音声サンプリング	12kHz
LPC サイズ	16
フレームサイズ	25ms
フレームシフト	10ms

表 3.5 被評価話者（話者番号：108）におけるホマント補正值

	語中		語頭	
	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]
/a/	119	111	100	30
/o/	7	-89	6	-171
/u/	-0.1	144	4	192
/i/	-11	293	18	263
/e/	17	210	43	210

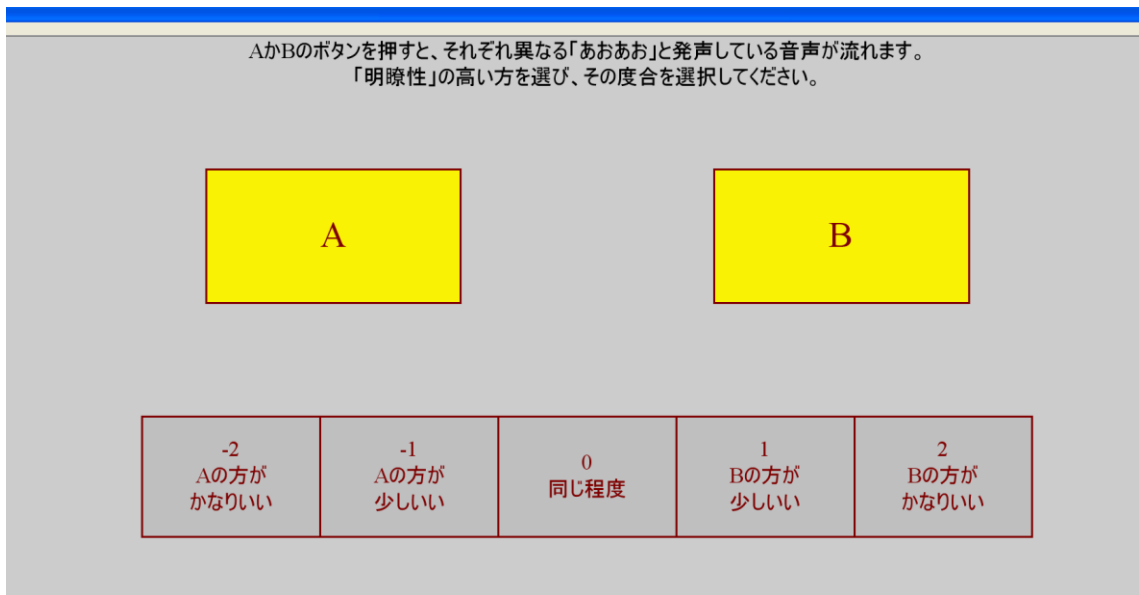


図 3.6 聴取実験画面

実験結果

シェッフェの一対比較の中屋変法によって算出された各音声間の平均嗜好度の差 ($\alpha_i - \alpha_j$) とその 99%信頼度区間を図 3.7 に示す。図 3.7 は、原音声と補正音声の比較結果である。縦軸は平均嗜好度、線は 99%信頼区間であり、正の値は処理後の音声、負の値は処理前の音声の方が明瞭度が高いことを示している。また、99%信頼区間が 0 をまたいでいる場合は、比較に有意性が無いことを意味する。図 3.7 より、10 単語中 10 単語の音声において、平均嗜好度が正の値を示しており、99%信頼区間も 0 をまたいでいない。これから、補正音声と比較すると補正音声の方が明瞭度が高いと評価されていることがわかる。また、各々の評価値についてみると、「ぎよくほ」という単語については他の単語に比べて平均嗜好度が低い。これは、「ぎよくほ」に含まれる母音が /o/, /u/, /o/ の 3 つであり、高明瞭度話者の平均ホルマント周波数と差分値が小さく、補正值も小さかったため、このような結果となったと考えられる。

この聴取実験結果より、提案したホルマント周波数補正方法は、補正值が既知の場合においては、有効であることがわかった。

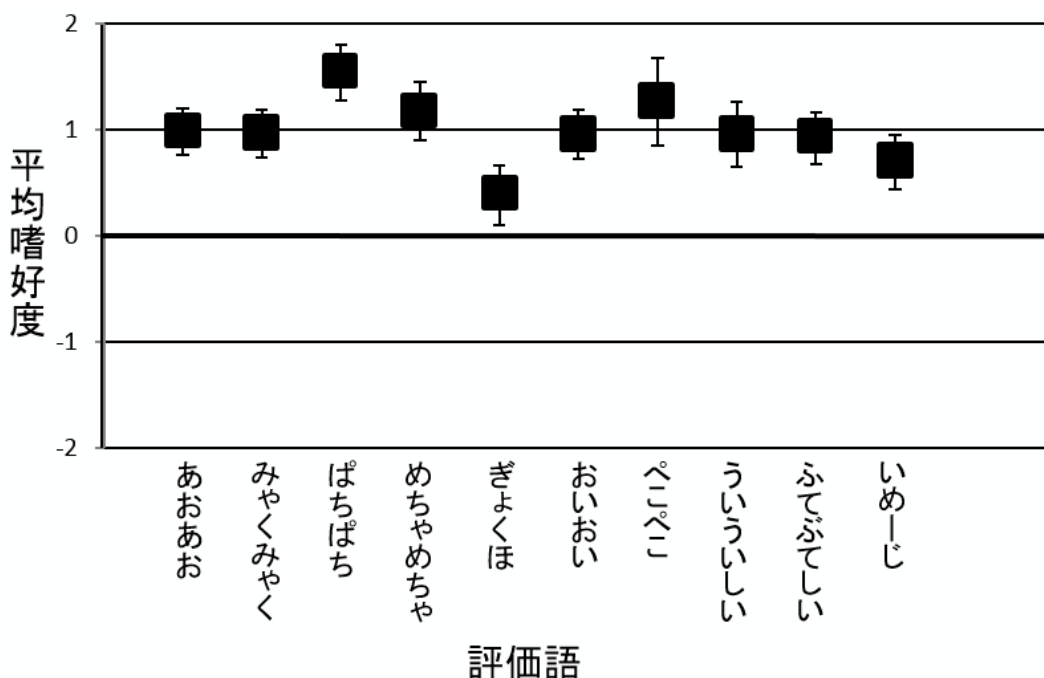


図 3.7 補正值が既知の場合における処理前後の比較 (補正值が差分値の場合)

3.3.3 シフト量が未知の場合における補正効果

ここでは、補正を行う低明瞭度高齢者の平均ホルマント周波数が未知の場合における、倍率を用いて補正を行った場合の検証結果について述べる。被評価話者として、明瞭度の低い高齢者6名の中から明瞭度が低いと思われる話者を認知で2名選出した。評価単語は、データベースである543単語の中から2~3種類の母音を含む4種類の単語音声を選択した。表3.6に評価単語を示す。補正音声を作成する条件は表3.4である。

この被評価話者におけるホルマントの補正值については、F1-F2平面上におけるホルマント周波数が面積を広げるように、原音声のホルマント周波数を表3.7の倍率を用いて補正を行った。この倍率は、予備検討で最も妥当な倍率を検討し決定した。

聴取実験方法は、3.3.2と同様の条件で行った。

表 3.6 評価単語

あおあお	ういういしい	おいおい	いめーじ
------	--------	------	------

表 3.7 ホルマント周波数の補正量（一定倍率）

	F1	F2
/a/	1.1	1.1
/o/	1.2	0.95
/u/	1.0	1.0
/i/	0.9	1.3
/e/	1.2	1.2

実験結果

シェッフェの対比較の中屋変法によって算出された各音声間の平均嗜好度の差 ($\alpha_i - \alpha_j$) とその 99%信頼度区間を図 3.8 に示す。図 3.8 は、原音声と補正音声の比較結果である。縦軸は平均嗜好度、線は 95%信頼区間であり、正の値は処理後の音声、負の値は処理前の音声の方が明瞭度が高いことを示している。また、信頼区間が 0 をまたいでいる場合は、比較に有意性が無いことを意味する。図 3.7 より、低明瞭高齢者(108)の場合、すべての単語において補正前より、補正後の音声の方が明瞭度が高いと評価されている。

また、話者 (101) の場合、「ういういしい」と「おいおい」という単語については、補正後の音声の方が明瞭度が高いと評価されているが、「いめーじ」では、有意性が無く、「あおあお」に関しては、補正前の原音声の方が明瞭度が高いと評価されている。これは、話者 101 においては、母音/ae/については一定倍率による補正効果がみられなかったことがわかる。

これは、2 章で検討したホルマント周波数の個人性が影響していると考えられる。話者 108 においては、図 2.13, 図 2.19 をみると、高明瞭度の平均ホルマント周波数の母音/u/と、話者 108 の平均ホルマント周波数の母音/u/は近い点に位置しており、低明瞭高齢話者の平均と同じような傾向を示している。しかし、話者 101 のホルマント周波数の母音/u/と、高明瞭度の平均ホルマント周波数の母音/u/は近い点に位置しておらず、比較すると、話者 101 のホルマント周波数は、全体的に F1, F2 は共に低い方に位置していることがわかる。

このことから、ホルマント周波数を一定倍率補正する方法は、ホルマント周波数の全体的な傾向が、2.4.2 節で述べたような、母音/u/には変化がなく、他の 4 母音が母音/u/に収束しているような傾向を示す場合において有効であると考えられる。

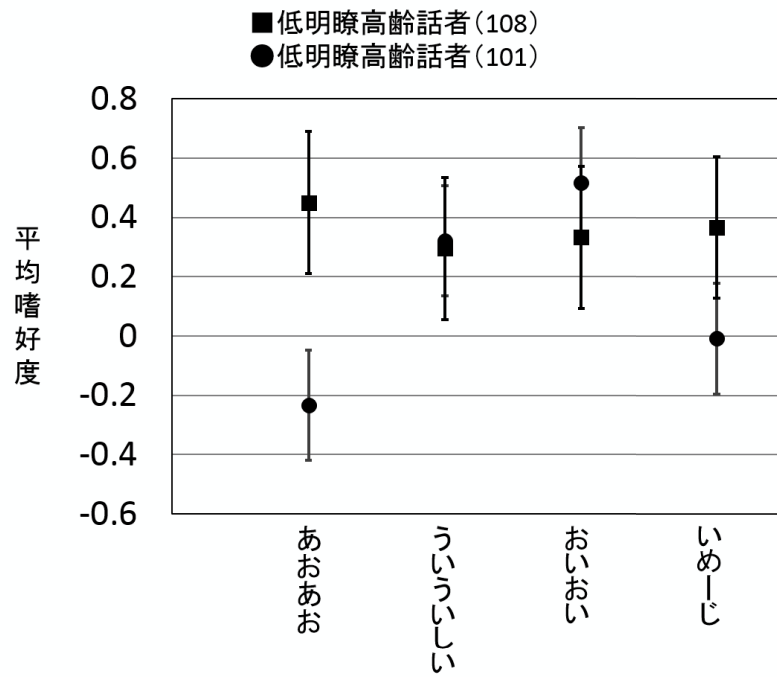


図 3.8 補正値が未知の場合における処理前後の比較（補正値が一定倍率の場合）

第 4 章 電気式人工喉頭の音質改善システム

4.1 電気式人工喉頭音声改善システムの提案

図 4.1 に、本研究で提案する喉頭摘出者を対象にした音声コミュニケーション支援システムの概念図を示す。このシステムでは、実際の声帯音源に近い振動を発生させる音源ユニットを用いる。その音源ユニットを用いて発声された音声信号は、耳孔内マイクを用いて収録される。このようにして得られた音声は、声質変換の入力データとして使用される。音質変換により通常音声に近い音声に変換され、スピーカから出力される。ここでは、使用する音源ユニット、マイクロフォン、音質改善システムに分けて説明する。

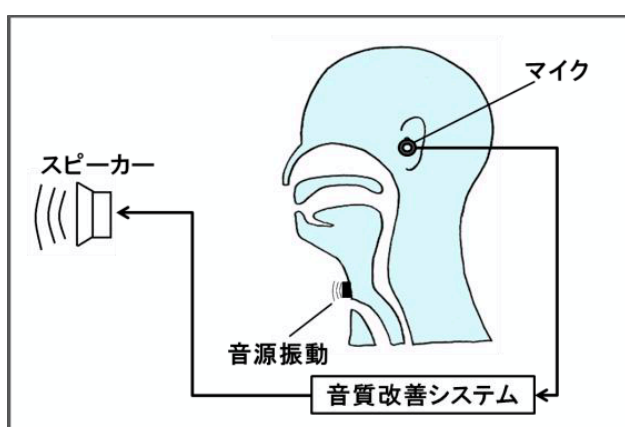


図 4.1 音質改善システムの概念図

- 音源ユニット

本論文では、コバテル社より販売されている振動スピーカ「VR3000」を既存の電気喉頭の代替物として使用する。振動スピーカは、ダイナミックスピーカークの構造を応用しており振動板、磁気回路、スプリングコイル、ハウジングで構成されている。ユニット内部の振動板を震わせて音を発生させる一般的なスピーカとは異なり、本体底面と接触している物を振動版として利用することでスピーカとして機能する。ゆえに、振動スピーカを用いることで理想的な疑似声帯音源を咽喉内に伝えることができる。

- 耳孔内マイクロフォン

声質変換を行う際には、できるだけノイズの含まれていない音声进行处理する必要があ

る。しかし、電気喉頭を用いて発声する際、音声以外のノイズの他にも、電気喉頭の音源の音が周囲に漏れてしまうため、大きなノイズとなってしまふ。そこで、処理に用いる発声音声を収録する場合、周囲雑音の影響を受けにくい収録方法をとらなければならない。

そこで、ノイズの影響を受けにくい体に装着するマイクロフォンの検討を行った。耳孔内マイク、骨伝導マイク、咽喉マイクを用いて音声を収録し、ノイズと音声の SNR を算出した。使用したマイクは表 4.1 にまとめ、結果を図 4.2 に示す。

口元のマイクの結果より、「あ」「え」「お」は発声部のパワーが出ているが、「い」「う」は発声部のパワーが出にくい、ということがわかる。そのパワーが出にくい「い」「う」について、各種のマイクの結果を見ると、耳孔マイクの SNR が、口元の SNR と同等かそれ以上の改善が見られるということがわかる。これより、音源のパワーを小さく使用する時には、耳孔内マイクで音声を収録することが、音源から発するノイズの影響を受けにくいということがわかる。

表 4.1 使用したマイクロフォン

マイクの名称	型番	装着部
耳孔内マイク	NEH-A20B2 (日鉄エレックス)	耳
骨伝導マイク	ECM-TL3 (SONY)	耳
咽喉マイク	DRMC01P (DP)	首

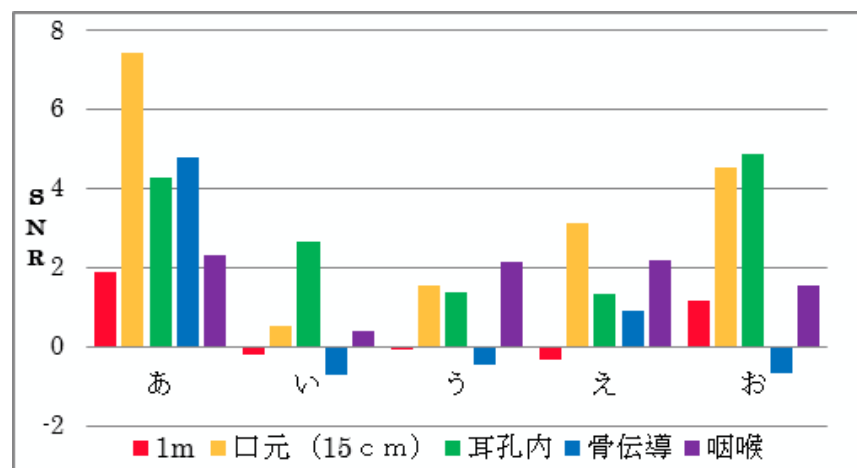


図 4.2 各マイクロフォンの SNR の比較

● 音質改善システム

電気喉頭や振動スピーカといった代替発声機器を用いた音声は、機械的な音声であり、スペクトル包絡に違いがあることがわかっている。この問題を解決するために、代替発声音声のスペクトル包絡を、目的とする通常音声のスペクトル包絡へ変換する。図 4.3 にスペクトル包絡変換の概要を示す。この方法では、音韻バランスのとれた単語を発声した 2 種類の音声を用いて、元の通常音声と疑似声帯音声との間のスペクトル包絡の特徴量を GMM によりモデル化し、モデル間距離を用いて対応関係の学習を行う。この関係を用いて、未知の疑似声帯音声のスペクトル包絡を通常音声のスペクトル包絡へと変換する。

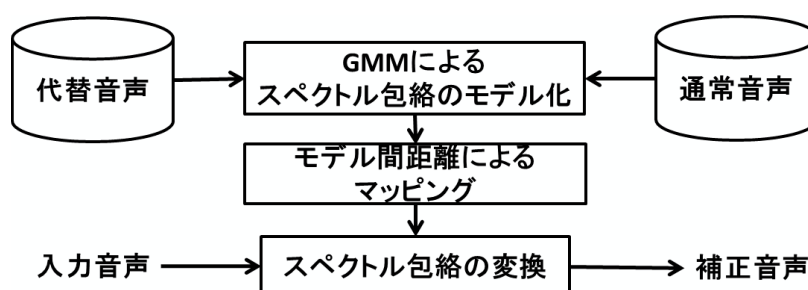


図 4.3 スペクトル変換の概要

4.2 LPC 分析合成に基づく音質改善方法の予備実験

ここでは、LPC 分析合成によるスペクトル包絡の変換により、音質改善効果があるのかを検証するための予備実験を行う。

4.2.1 実験内容

通常音声と代替音声を LPC 分析し、代替音声の音源情報に通常音声のスペクトル包絡である包絡情報を畳み込むことにより、音質が改善するのか評価する。図 4.4 に本実験の概略を示す。代替発声音声は EL と、それぞれの母音から抽出した LPC 残差波音源を、振動スピーカを用いて発声した 5 母音 (/a/, /i/, /u/, /e/, /o/) を用いる。

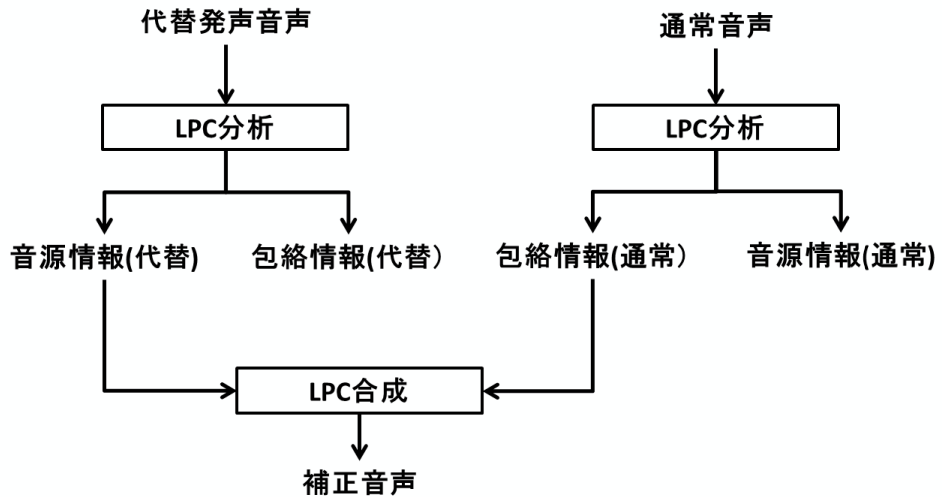


図 4.4 予備実験の概略

4.2.2 実験結果

簡単な音質評価により，低域での音質が改善しており，肉声らしさがましていることがわかった．そこでケプストラム分析を用いて客観評価した．ケプストラム分析については次章で説明する．分析条件は表 4.2 にまとめる．また表 4.3 は，通常音声とのケプストラム距離を算出し，まとめたものである．LPC 残差波音声を LPC と記している．表 4.3 より，LPC 分析合成によるスペクトル包絡の変換の結果，代替音声は通常音声に近い音声に補正されていることがわかる．

ここで，LPC 残差波の変換前後と通常音声のスペクトルを比較する．図 4.5 から図 4.9 に母音毎のスペクトルの比較を示す．図より，スペクトル包絡の変換により変換音声のスペクトル包絡の形状が通常音声に近くなっているのがわかる．特に，低域での改善幅が大きい．

表 4.2 ケプストラム分析時の条件

分析区間	Center 1.0[s]
窓幅	25[ms]
フレームシフト幅	5[ms]
フーリエ変換の次数	1024
フレーム数	195
リフタリング次数	55

表 4.3 通常音声とのケプストラム距離

	/a/	/i/	/u/	/e/	/o/	Sum total
LPC 変換前	0.913899	0.774682	0.94099	1.091547	1.270773	4.991891
LPC 変換後	0.677406	0.733545	0.828524	0.57966	0.924478	3.743613
EL 変換前	1.530725	1.063768	1.61036	0.948571	1.454765	6.608189
EL 変換後	0.705177	0.816065	0.910104	0.540517	0.994911	3.966774

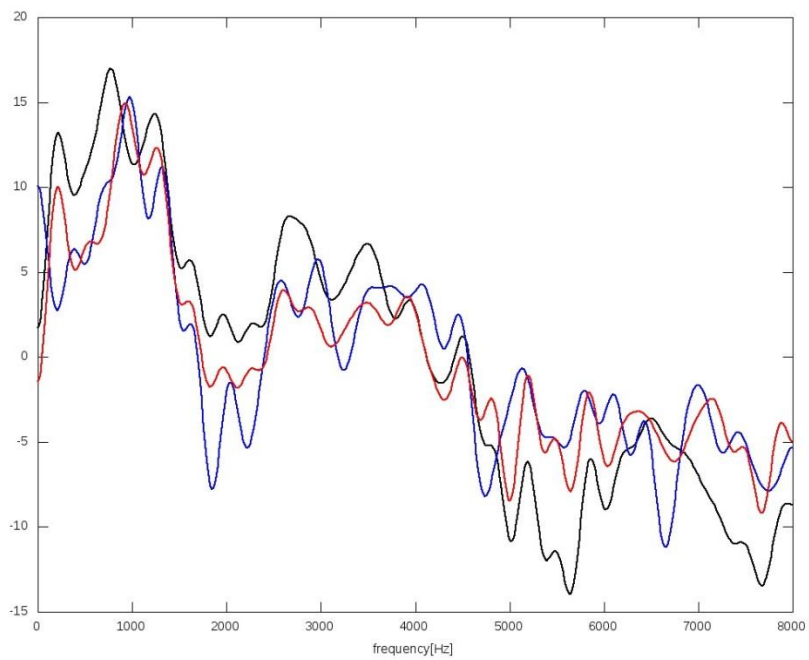


図 4.5 /a/のスペクトル変換結果(黒：通常音声 青：変換前 赤：変換後)

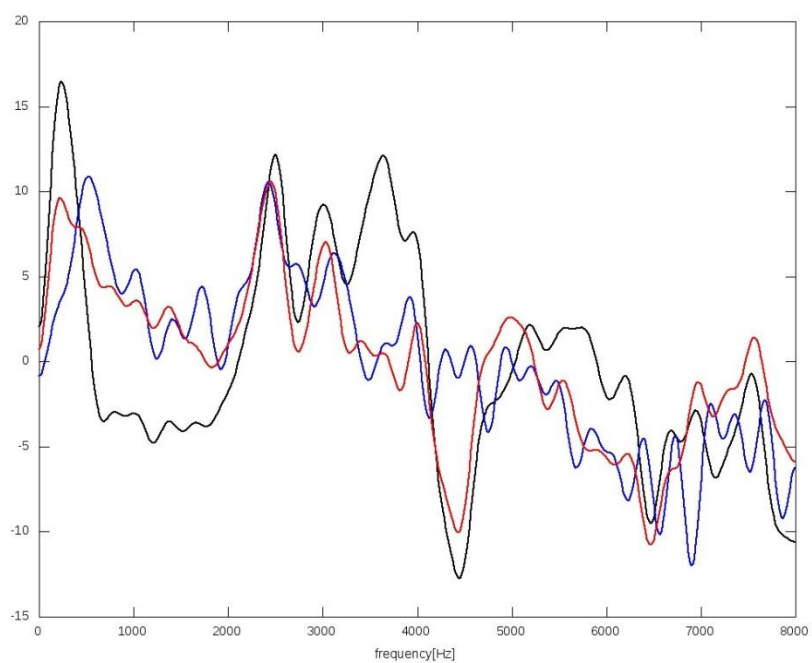


図 4.6 /i/のスペクトル変換結果(黒：通常音声 青：変換前 赤：変換後)

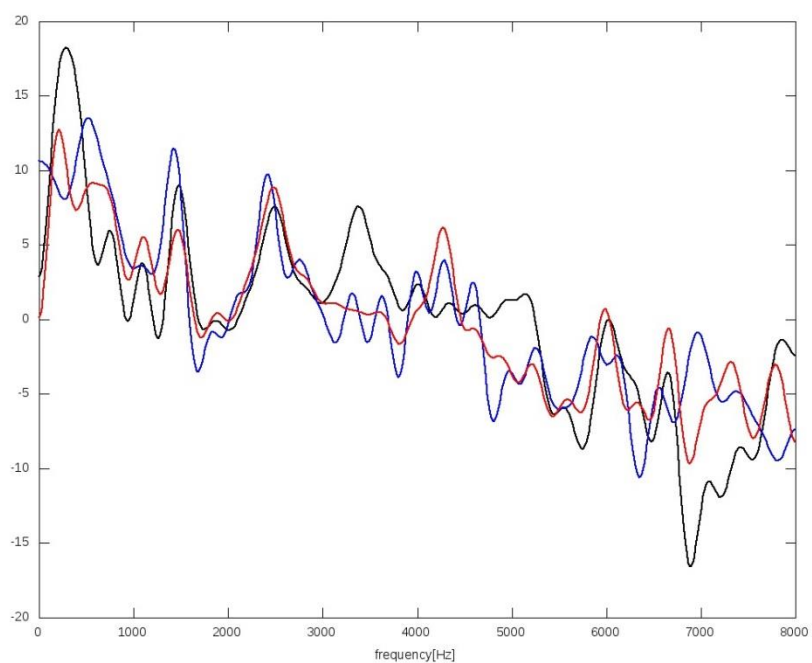


図 4.7 /u/のスペクトル変換結果(黒：通常音声 青：変換前 赤：変換後)

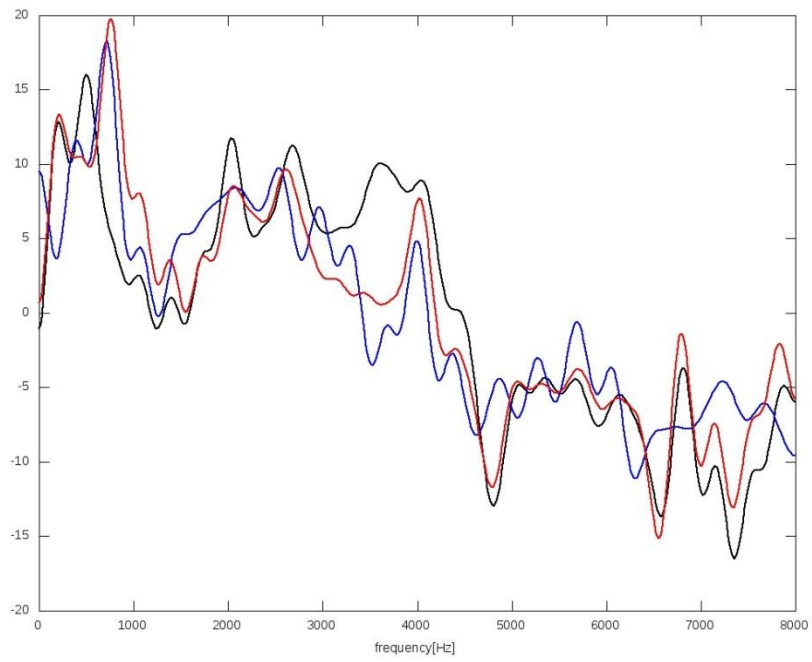


図 4.8 /e/のスペクトル変換結果(黒：通常音声 青：変換前 赤：変換後)

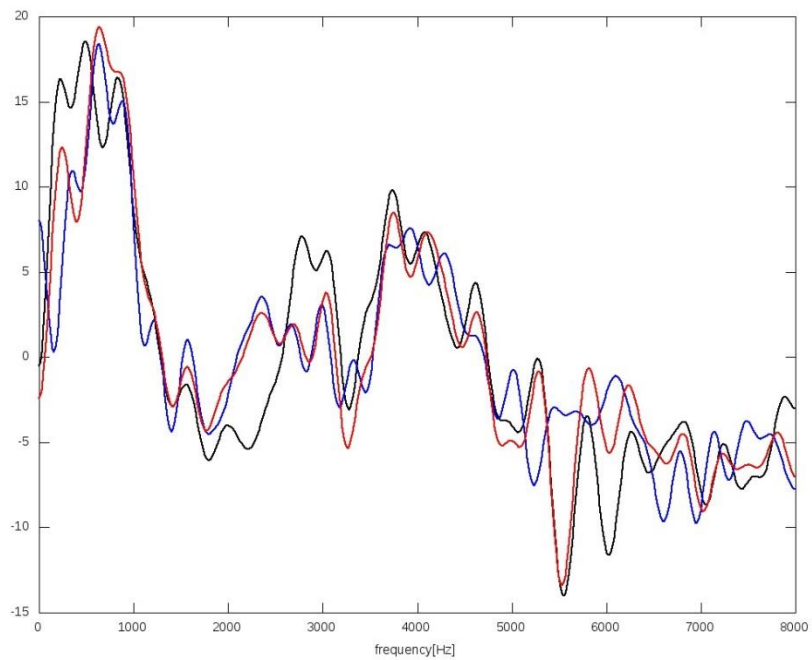


図 4.9 /o/のスペクトル変換結果(黒：通常音声 青：変換前 赤：変換後)

第5章 種々の音源による電気式人工喉頭の音質 改善

喉頭摘出者は、音源である声帯音源は無いが、声道は健聴者と同様にあるため、調音機能は有している。そのため、外部から与える音源として、人間の声に近くなるような音源を選定する必要があると考えられる。音源の選定により、提案する発声支援システムに適した擬似声帯音源について検討する。

5.1 電気式人工喉頭と擬似声帯音源による音声の比較実験

5.1.1 擬似声帯音源を用いた音声の生成

我々は、システムに用いる音源の予備検討として、声帯音源として近いと言われている「パルス波音源」、「のこぎり波音源」、「Rosenberg 波音源」の3種類の比較を行った[45]。その結果、「Rosenberg 波」が肉声に近いことがわかった。この結果を踏まえ、新たに、音声から声帯情報を抽出した「LPC 残差波形音源」を用いた音質改善方法を提案し、比較検討を行った。以下に、本実験で用いた「Rosenberg 波音源」、「LPC 残差波音源」の2種類の生成方法について述べる。

咽喉内へ伝える音源として、種々の音源として以下の2つの擬似声帯音源を生成した。

● LPC 残差波音源

LPC 残差波とは、発声話者の音声を LPC 分析し、声道特性であるスペクトル包絡を原音声へ逆フィルタリングすることで生成する音源波形のことである。LPC 残差波は、通常音声の/a/, /i/, /u/, /e/, /o/から声帯振動を抽出した。図 5.1 の左が通常音声の/a/, 右が LPC 残差波音声の/a/である。図 5.1 より、包絡情報が除去され、ピッチ（音源情報）が表現できていることがわかる。

● Rosenberg 波音源

声帯振動は、非対称な三角波に近似すると言われており[46]、Rosenberg 波は音声合成時に声帯振動として用いられている。Rosenberg 波とは、Rosenberg が提唱したモデルを Klatt らが改良した波形のことである[48]。本研究では、式(5.1)を用いて波形を作成した。式中では、 $\tau_1 = 0.95$ $\tau_2 = 0.90$ とし、サンプリング周波数は 16kHz, 16bit, 波形の周期は 150Hz になるように作成した。また、図 5.2 は Rosenberg 波形である。

$$f(t) = \begin{cases} 3\left(\frac{t}{\tau_1}\right)^2 - 2\left(\frac{t}{\tau_1}\right)^3 & (0 \leq t \leq \tau_1) \\ 1 - \left(\frac{t - \tau_1}{\tau_2}\right)^2 & (\tau_1 \leq t \leq \tau_2) \end{cases} \quad (5.1)$$

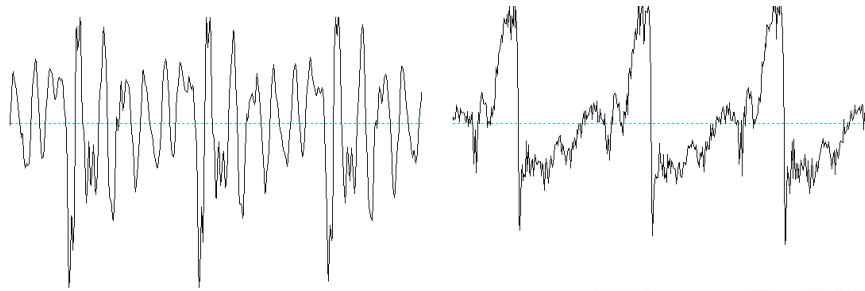


図 5.1 /a/波形 (左：通常音声 右：LPC 残差波)

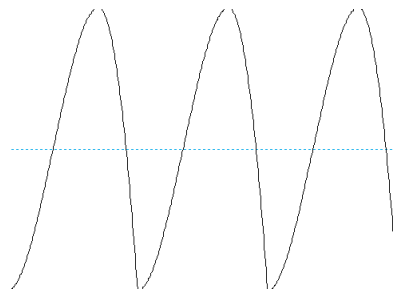


図 5.2 Rosenberg 波

本研究では音声分析ソフト praat を用いて LPC 残差波を抽出した. 分析する際の設定は以下の表 5.1 のようにした.

表 5.1 LPC 残差波作成時の praat 設定

Prediction order	16
Window length(s)	0.025
Time step(s)	0.005
Pre-emphasis frequency(Hz)	50.0

図 5.3 の左図は/a/の LPC 残差波を作成する際に用いた原音声のスペクトルで, 横軸が周波数で縦軸がパワーである. ピークが複数見られ, 周波数の低い方から順に第 1 ホルマント, 第 2 ホルマント…と呼ばれる. ホルマントは音声の特徴を表しており, 音韻性に大きく関わる. 右図は/a/を LPC 分析し逆フィルタの処理をした疑似声帯音源のスペクトラムである. ホルマントがほとんどわからず, 音韻性が無くなっていることがわかる.

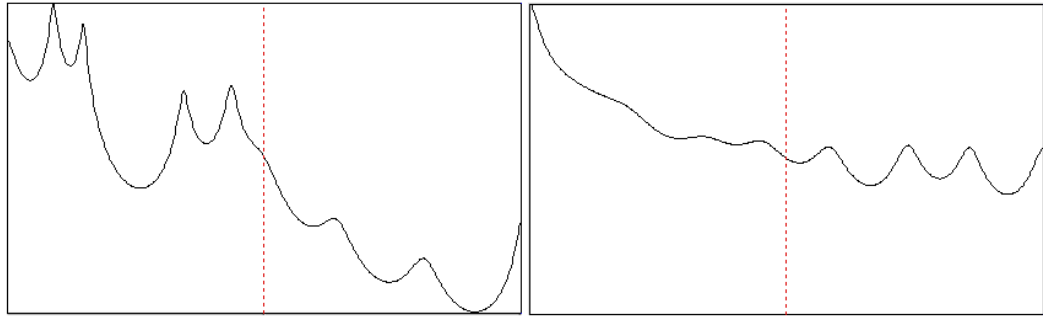


図 5.3 スペクトラム(左 : /a/ 右 : /a/からの疑似声帯音源)

5.1.2 音声収録

本研究では、EL を用いて発声した音声と 2 種類の疑似声帯音源による音声を比較するために、喉頭を失っていない成人男性 1 名分の音声を収録した。

EL は DENCOM 社 DEL-Y200「ユアトーン II ゆらぎ」を使用した。また、疑似声帯音源を咽喉内に伝えるために、電気喉頭に替えて振動スピーカを使用した。本実験では、コバテル社 VR3000 を用いた。

収録場所は防音室で行い、収録音声は、EL を用いた発声法でも容易に発声が可能であり、日本語の音声の単語の殆どを構成する母音/a/, /i/, /u/, /e/, /o/の 5 つを収録した。また、LPC 残差波音源は/a/を発声するときには通常発声の/a/から抽出した音源を、/i/を発声するときには通常音声の/i/から抽出した音源のように、それぞれの母音から抽出した音源を使用した。

使用した EL は音量と声の高さをそれぞれ「1」から「9」まで調整する機能がついており、声の高さは「4」に設定し、音量は振動スピーカとほぼ同じ大きさになるよう「7」に設定し収録を行った。収録した音声ファイルは 48kHz, 16 ビットでデジタル録音である。この収録した音声ファイルを 16kHz にダウンサンプリングし、分析、聴取実験を行った。収録マイクは、口元から 15cm のところに設置した無指向性マイクである。表 5.2 に収録条件をまとめる。

表 5.2 収録条件

話者	EL の扱いに慣れた健常者 1 名
収録マイク	口元に設置した無指向性マイク
収録場所	防音室（周囲雑音なし）
収録音声	/a/, /i/, /u/, /e/, /o/
音源の種類	EL（声の高さ 4, 音量 7） 振動スピーカ（音韻毎に異なる LPC 残差波形） 振動スピーカ（Rosenberg 波）

5.1.3 聴取実験条件

疑似声帯音源を用いて発声した音声は EL に比べてどれだけ肉声に近いかを聴取実験により評価した。比較音声は「EL」「振動スピーカ（LPC 残差波）」「振動スピーカ（Rosenberg 波）」の 3 種類の音源でそれぞれ 5 母音発声した。被験者は EL 音声を日常的に聞く環境にない大学生 10 名に協力してもらった。場所は、静かな環境で、音声分析ツール「Praat」の聴取実験システムを用いて実験を行った。被験者には各音声の中からランダムに選出した 2 つの音声を A 音声、B 音声として提示し、肉声らしいものについて「2 : A 音声の方がかなりいい」、「1 : A 音声の方が少しいい」、「0 : 同程度」、「-1 : B 音声の方が少しいい」、「-2 : B 音声の方がかなりいい」の 5 段階で評価してもらった。被験者 1 名に対して、1 母音につき 3 種類の組み合わせについて評価してもらい、それを 5 母音分実施した。被験者の判断誤差を低減するために、5 回の評価を行ってもらった。この聴取実験により相対的な肉声らしさの度合いがわかる。表 5.3 に聴取実験条件について示す。

表 5.3 聴取実験の条件

評価方法		2つの音声の肉声らしさについて 5段階評価
評価話者	年齢	20代(健常者)
	人数	男性1名
聴取実験の被験者	年齢	20代10名
	人数	男性8名 女性2名
評価回数		5回
評価音声		/a/, /i/, /u/, /e/, /o/の5母音 母音毎に、音源の異なる3種類の音声 (EL, LPC 残差波, Rosenberg 波)

5.1.4 シェッフエの一対比較による主観評価

シェッフエの一対比較の中屋変法によって算出された各音声間の平均嗜好度の差 ($\alpha_i - \alpha_j$) とその99%信頼度区間を、図5.4に示す。

まず、「EL音源」と「LPC残差波音源」の比較を行う。図5.4より、青のグラフがEL音源とLPC残差波音源の得点分布である。また、青のグラフ上にある黒のバーが99%信頼区間であり、全母音に対して一側に位置しているため、平均嗜好度の差が有意であることがわかる。このことより、EL音源とLPC残差波音源を比較すると、LPC残差波音源の方が肉声に近いことがわかった。

次に、「EL音源」と「Rosenberg波音源」の比較を行う。図5.4より、赤のグラフがEL音源とRosenberg波音源の得点分布である。99%信頼区間を確認すると、/a/, /i/, /u/では区間が+側と一側にまたがっており平均嗜好度の差が有意であるとは言えない。しかし、/e/と/o/に関しては区間が+側に位置しており平均嗜好度の差が有意であることがわかる。このことより、EL音源とRosenberg波音源を比較すると/e/と/o/のみでEL音源の方が肉声に近く、それ以外の母音では変わらないという結果だった。

最後に、「LPC残差波音源」と「Rosenberg波音源」の比較を行う。図5.4より、緑のグラフがLPC残差波音源とRosenberg波音源の得点分布である。99%信頼区間を確認すると、すべての母音に対して区間が+側に位置しており、平均嗜好度の差が有意である。

このことより、LPC 残差波音源と Rosenberg 波音源を比較すると、LPC 残差波音源の方が肉声に近いことがわかった。

また、図 5.5 は集計結果に基づく一次元尺度上の布置である。図 5.5 より LPC 残差波音源が EL 音源、Rosenberg 波音源よりも肉声に近いことがわかる。以上のことより今回調査した音源の中で LPC 残差波音源が最も肉声に近いことがわかった。

最後に、EL 音源と Rosenberg 波音源を比較した際に、Rosenberg 波音源の方が EL 音源よりも肉声らしくないという結果がでたことは使用するデバイスによる影響が大きかったのではないかと考えられる。振動スピーカは元々机などの上において振動を伝え、音にするデバイスである。そのため、喉に押し当て、声道内に振動を伝えることが EL に比べて難しかったためではないかと思われる。そのようなデバイスによる不利がありながらも LPC 残差波音源の評価が良かったことは音源振動の違いが音声に与える影響が非常に大きいことを示す。

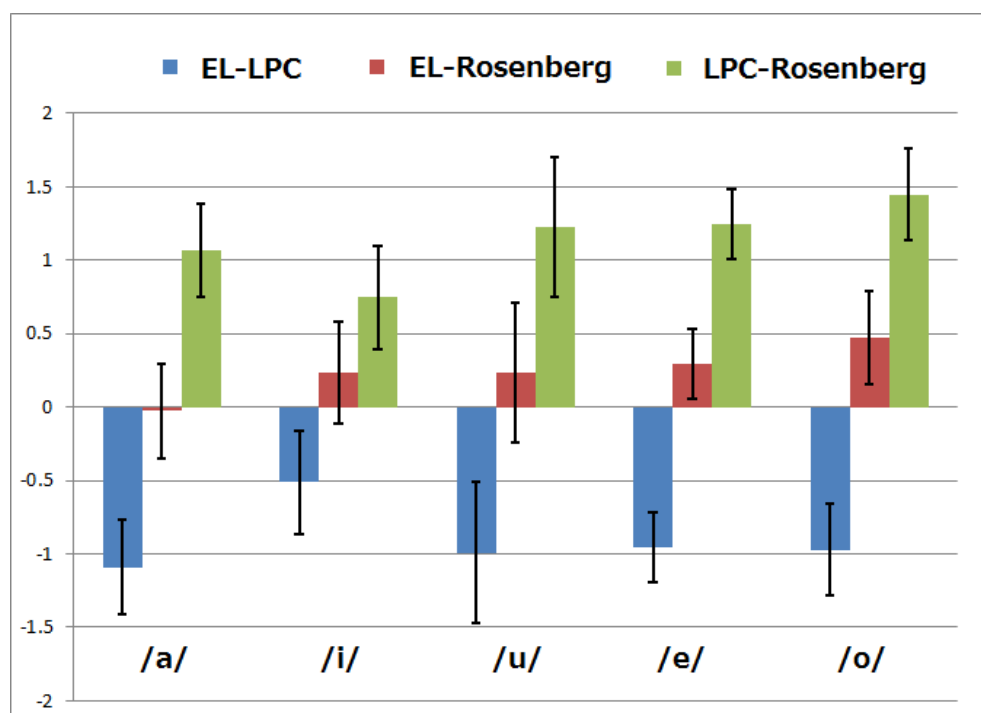


図 5.4 平均嗜好度の差と 99%信頼区間

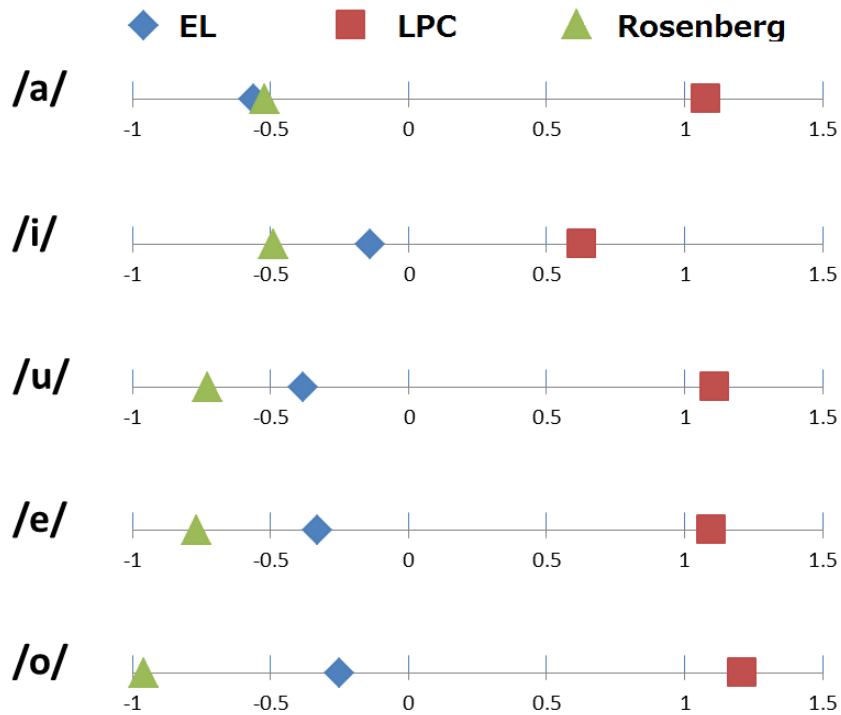


図 5.5 集計結果に基づく一次元尺度

5.1.5 ケプストラム分析による客観評価

本項ではケプストラム分析を用いて母音毎に通常音声，EL 音声，振動スピーカによる LPC 残差波音声，振動スピーカによる Rosenberg 波音声の四種類を比較する．図 5.6 から図 5.20 までが通常音声と代替発声音声の対数振幅スペクトルの比較である．分析する音声は聴取実験に用いた音声を使用し，声の安定している音声の中心部分の 100ms を切り出し，分析を行った．フーリエ変換の次数は 1024 とし，リフタリングの次数は 55 とした．また比較する対数振幅スペクトルはそれぞれのパワーの平均が 0 になるように編集してある．

図 5.6 の/a/の通常音声と EL 音声に着目すると，全体的に形状が似てはいるが，1000Hz 付近のホルマンントの位置がずれており，3000Hz から 4000Hz 付近ではパワーが出ていないことがわかる．次に，図 5.7 の/a/の通常音声と LPC 残差波音声に着目すると，全体的に形状が似ており，EL 音声で低下していた 3000Hz から 4000Hz でもパワーが出ていることがわかる．また，ホルマンントのずれも EL 音声よりも小さいことがわかる．最後に，図 5.8 の/a/の通常音声と Rosenberg 波音声に着目すると，ホルマンントのピークが

通常音声に比べて少なく、音韻性が低いことがわかる。他の音韻でもほぼ同じような傾向が見られた。よって、以上の結果より、EL 音声、LPC 残差波音声、Rosenberg 波音声を通常音声と比較すると、通常音声の対数振幅スペクトルの包絡に EL 音声と LPC 残差波音声の包絡が似ていることがわかった。また、LPC 残差波音声の包絡は他の音声よりもホルマントがはっきりしていることがわかった。

さらに、2 音間の音色距離としてケプストラム係数のユークリッド距離を算出した。それぞれ通常音声との距離を算出している。発声区間の中心の 0.5[s]を切り出し、窓幅 25[ms]、フレームシフト幅 5[ms]、フーリエ変換の次数 1024、リフタリングの次数は 55 とし、1 単語につき 95 フレームの平均を算出した。結果を表 5.4 にまとめる。表中では電気喉頭音声を EL、振動スピーカによる LPC 残差波音声を LPC、Rosenberg 波音声を Rosen と記す。全体の傾向として LPC 残差波音声最も通常音声に近く、次が電気喉頭音声、最後が Rosenberg 波音声と言う結果になり、聴取実験の結果と一致する。しかし、音韻毎に比較すると、/a/では EL 音声の評価が悪いが、聴取実験の結果では Rosenberg 波音声との差異はなかった。

以上の結果より、ケプストラム分析により音声を比較すると LPC 残差波音声は肉声らしく、音韻性が強いことがわかった。

表 5.4 ケプストラム係数の通常音声とのユークリッド距離

	/a/	/i/	/u/	/e/	/o/	sum total
EL	1.5307	1.063766	1.610358	0.948568	1.454734	6.608126
LPC	0.863566	0.786679	0.816069	1.159214	1.328449	4.953977
Rosen	2.053559	2.488611	2.397599	1.975046	2.734057	11.648872

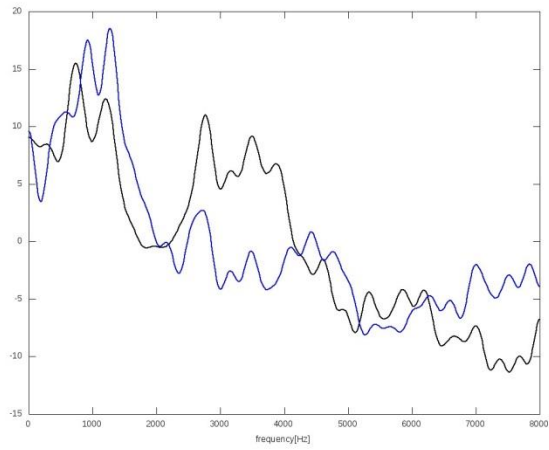


図 5.6 /a/の通常発声(黒)と EL 音声(青)の比較

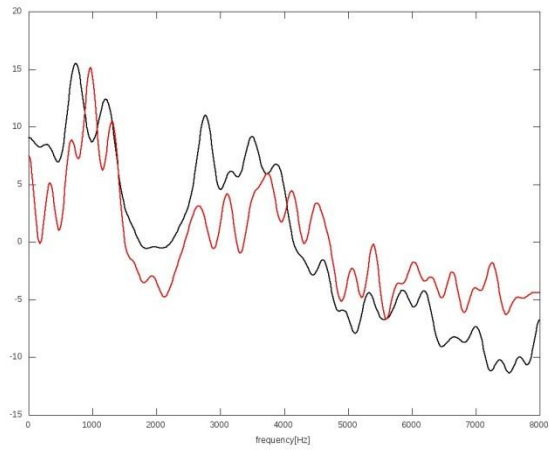


図 5.7 /a/の通常発声(黒)と LPC 残差波音声(赤)の比較

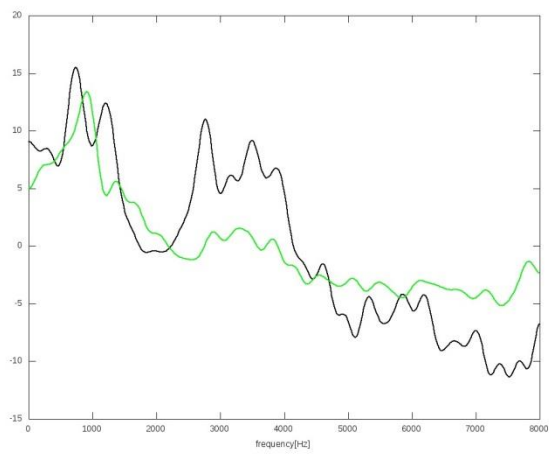


図 5.8 /a/の通常発声(黒)と Rosenberg 波音声(緑)の比較

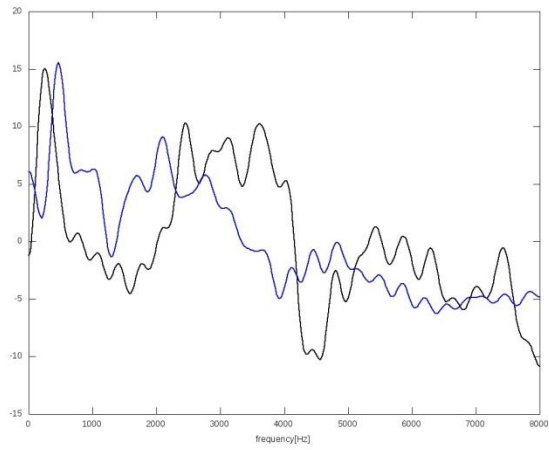


図 5.9 /i/の通常発声(黒)と EL 音声(青)の比較

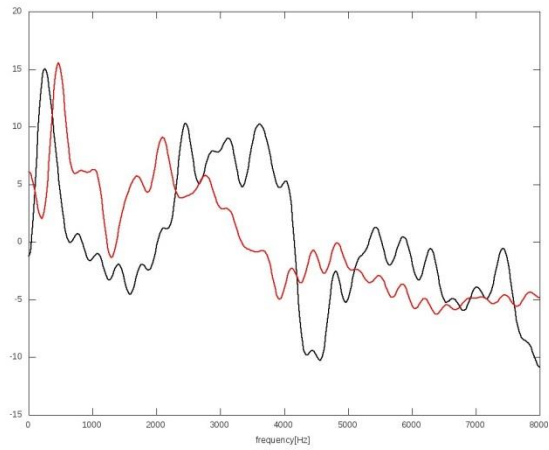


図 5.10 /i/の通常発声(黒)と LPC 残差波音声(赤)の比較

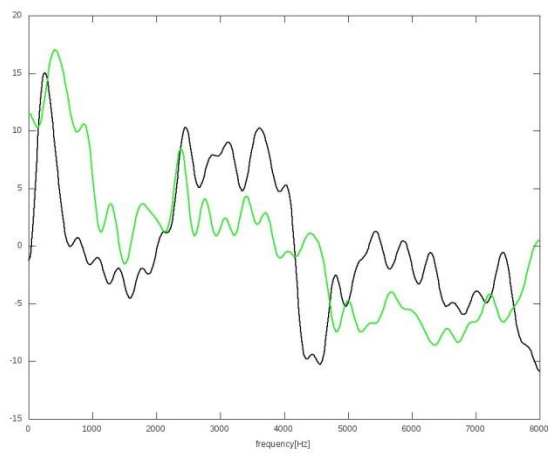


図 5.11 /i/の通常発声(黒)と Rosenberg 波音声(緑)の比較

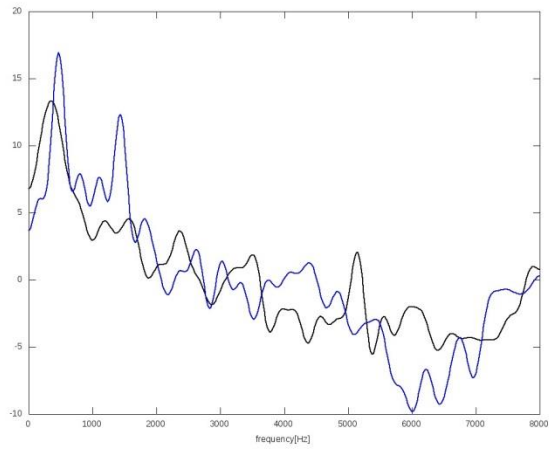


図 5.12 /u/の通常発声(黒)と EL 音声(青)の比較

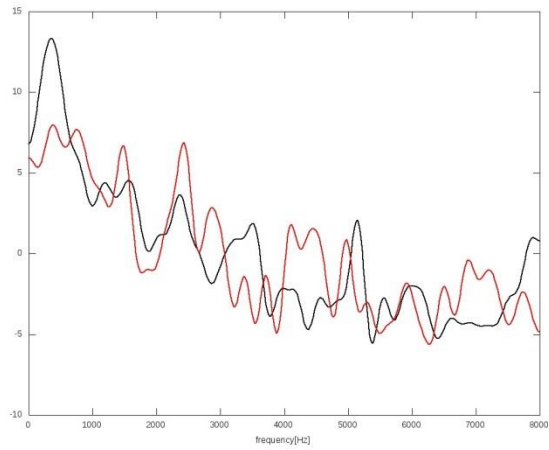


図 5.13 /u/の通常発声(黒)と LPC 残差波音声(赤)の比較

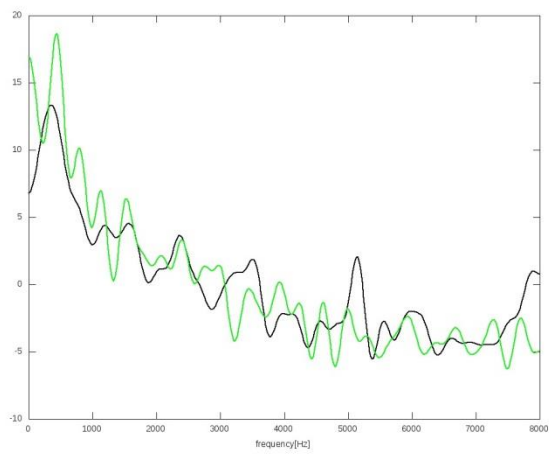


図 5.14 /u/の通常発声(黒)と Rosenberg 波音声(緑)の比較

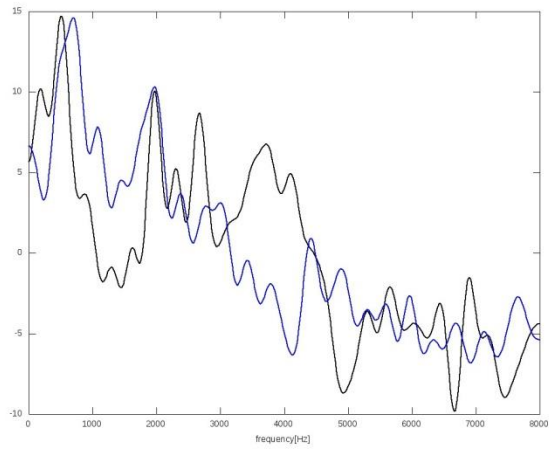


図 5.15 /e/の通常発声(黒)と EL 音声(青)の比較

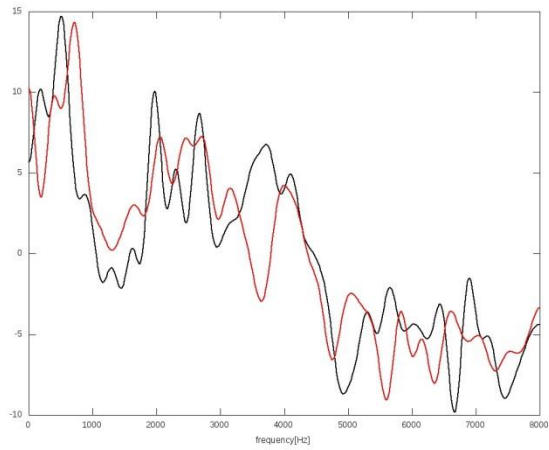


図 5.16 /e/の通常発声(黒)と LPC 残差波音声(赤)の比較

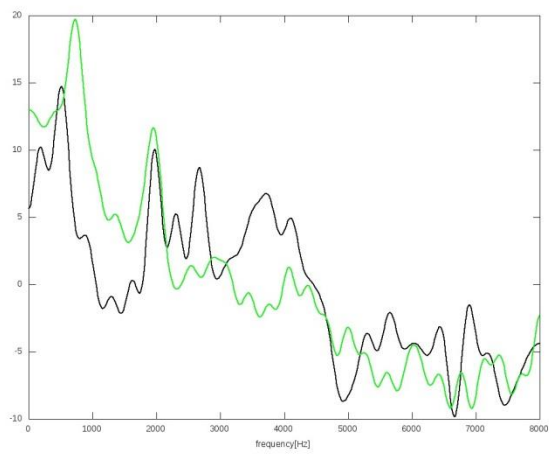


図 5.17 /e/の通常発声(黒)と Rosenberg 波音声(緑)の比較

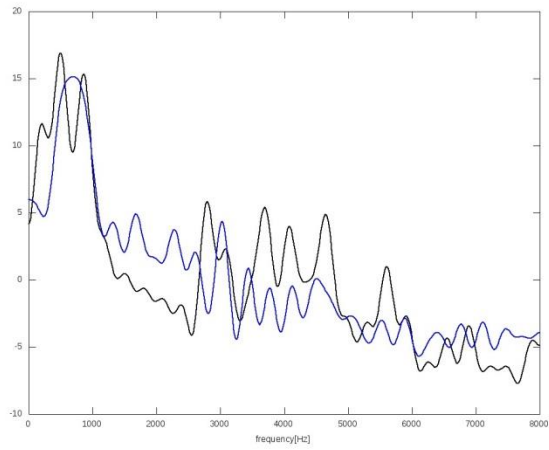


図 5.18 /o/の通常発声(黒)と EL 音声(青)の比較

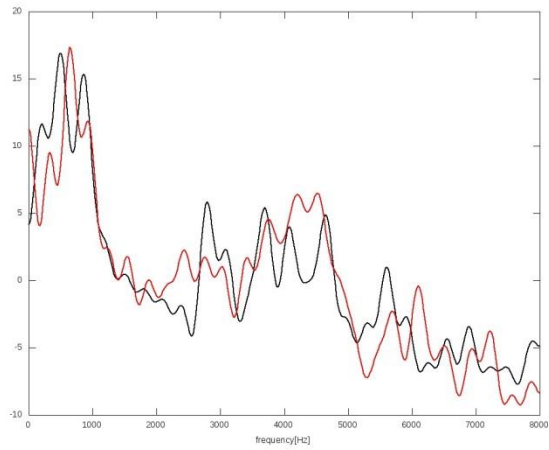


図 5.19 /o/の通常発声(黒)と LPC 残差波音声(赤)の比較

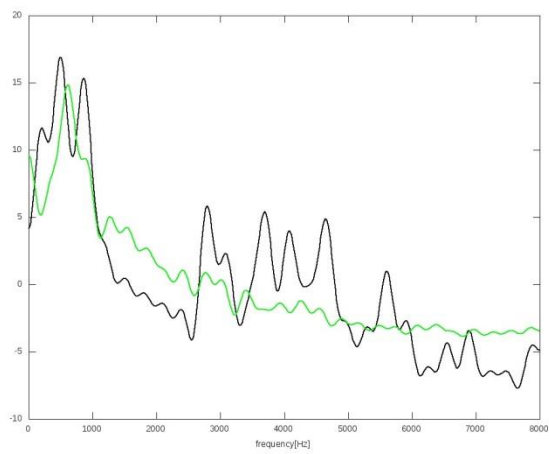


図 5.20 /o/の通常発声(黒)と Rosenberg 波音声(緑)の比較

5.2 LPC 残差波における抽出音源の違いによる音韻性への影響

5.2.1 音声収録

5.1 の実験で生成した 5 母音による LPC 残差波音源(LPC/a/, LPC/i/, LPC/u/, LPC/e/, LPC/o/)を, 振動スピーカを用いてそれぞれ 5 母音(/a/, /i/, /u/, /e/, /o/)発声し, 収録した. 話者は 1 名で, LPC 残差波を抽出した本人が行い, EL の扱いに慣れた健常者である. 収録マイクは口元に設置した無指向性マイク, 収録場所は防音室で, 収録した音声ファイルは 48kHz, 16 ビットでデジタル録音である. この収録した音声ファイルを 16kHz にダウンサンプリングし, 聴取実験を行った. 収録条件を表 5.5 にまとめる.

表 5.5 収録条件

話者	EL の扱いに慣れた健常者 1 名
収録マイク	口元に設置した無指向性マイク
収録場所	防音室 (周囲雑音なし)
収録音声	/a/, /i/, /u/, /e/, /o/
音源の種類	LPC/a/, LPC/i/, LPC/u/, LPC/e/, LPC/o/

5.2.2 聴取実験条件

5 母音から抽出した疑似声帯音源を用いてそれぞれ 5 母音の音声を発声し一対比較により評価する. 前章での実験と同じように, 被験者は EL 音声を日常的に聞く環境にならない大学生 10 名に協力してもらった. 場所は, 静かな環境で, 音声分析ツール「Praat」の聴取実験システムを用いて実験を行った. 被験者には各音声の中からランダムに選出した 2 つの音声を A 音声, B 音声として提示し, 肉声らしいものについて「2 : A 音声の方がかなりいい」, 「1 : A 音声の方が少しいい」, 「0 : 同程度」, 「-1 : B 音声の方が少しいい」, 「-2 : B 音声の方がかなりいい」の 5 段階で評価してもらった. 聞き直しについては制限しないものとした. 1 母音につき 5 種類, 10 組の組み合わせについて評価してもらい, それを 5 母音分実施した. 被験者の判断誤差を低減するために, 5 回の評価を行ってもらった. この聴取実験により相対的な肉声らしさの度合いがわかる. 表 5.6 に聴取実験の条件について示す.

表 5.6 聴取実験の条件

評価方法		2つの音声の肉声らしさについて 5段階評価
評価話者	年齢	20代(健常者)
	人数	男性1名
聴取実験の被験者	年齢	20代10名
	人数	男性8名 女性2名
評価回数		5回
評価音声		/a/, /i/, /u/, /e/, /o/の5母音 母音毎に, 5種類のLPC残差音源 (LPC/a/, LPC/i/, LPC/u/, LPC/e/, LPC/o/)

5.2.3 シェッフエの一対比較による主観評価

本実験もシェッフエの一対比較の中屋変法によって分析した。算出された各音源間の平均嗜好度の差 ($\alpha_i - \alpha_j$) とその95%信頼度区間を、図5.21から図5.30に示す。また、図5.31から図5.35にそれぞれの母音の一次元尺度を示し、図5.36で音源の一次元尺度の平均値を示す。

その結果、発声するLPC残差波音源の違いにより、音声の評価値に有意な差があることがわかった。つまり、LPC残差波音源の音韻性の違いが音声へ影響を与えている。また、母音毎に嗜好度を一次元尺度にすると音韻の一致する音声の嗜好度が高い傾向にあることがわかり、その中でも/o/より抽出されたLPC残差波音源が最も肉声らしく95%信頼区間を見てみると、65%の音声に対して優位であるという有意差があり、/o/より抽出したLPC残差波が劣位であるという有意差のある音声は無かった。また、残りの35%の音声には有意な差があるとは言えなかった。

よって、LPC残差波形から更に音韻生を除去することで、更なる明瞭化が図れると考えられる。

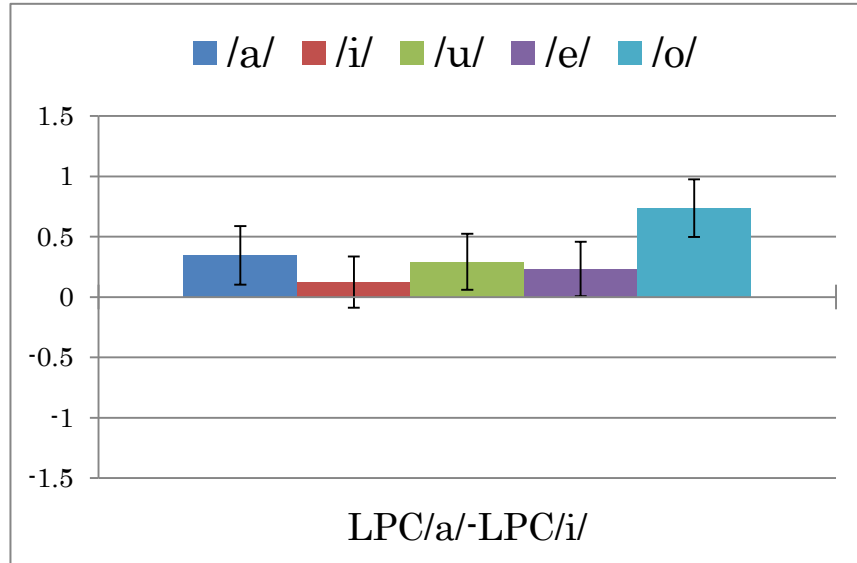


図 5.21 LPC/a/と LPC/i/の平均嗜好度の差と 95%信頼区間

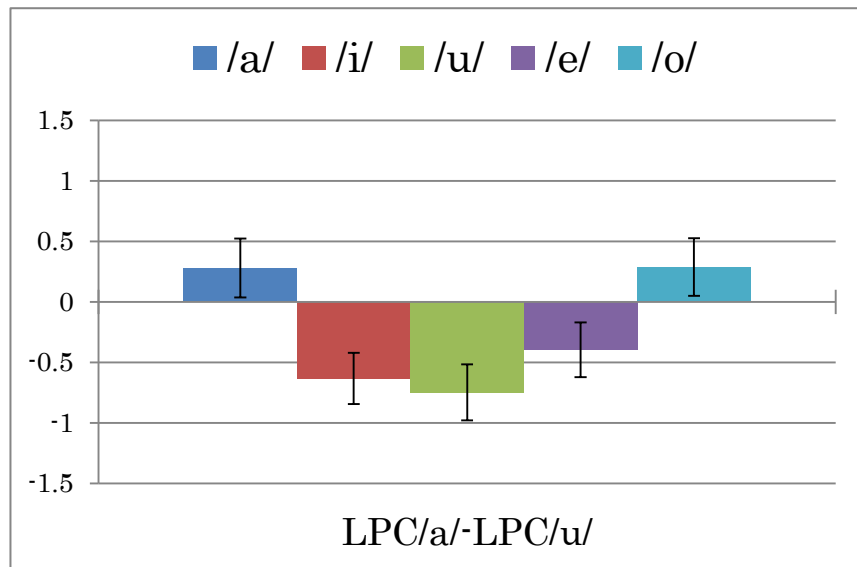


図 5.22 LPC/a/と LPC/u/の平均嗜好度の差と 95%信頼区間

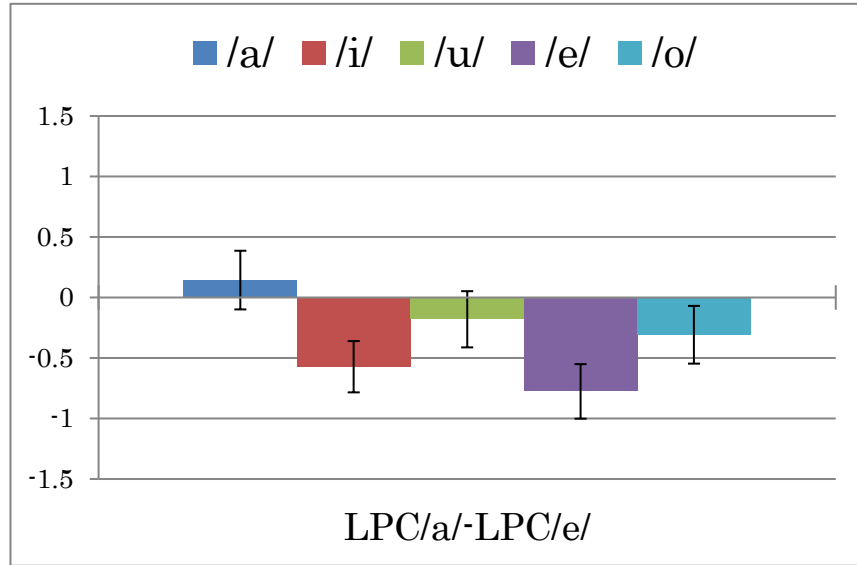


図 5.23 LPC/a/と LPC/e/の平均嗜好度の差と 95%信頼区間

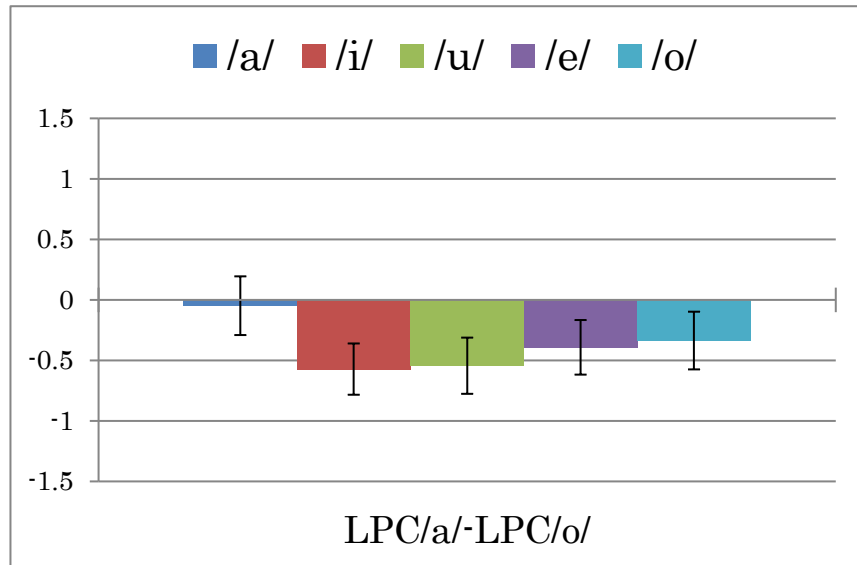


図 5.24 LPC/a/と LPC/o/の平均嗜好度の差と 95%信頼区間

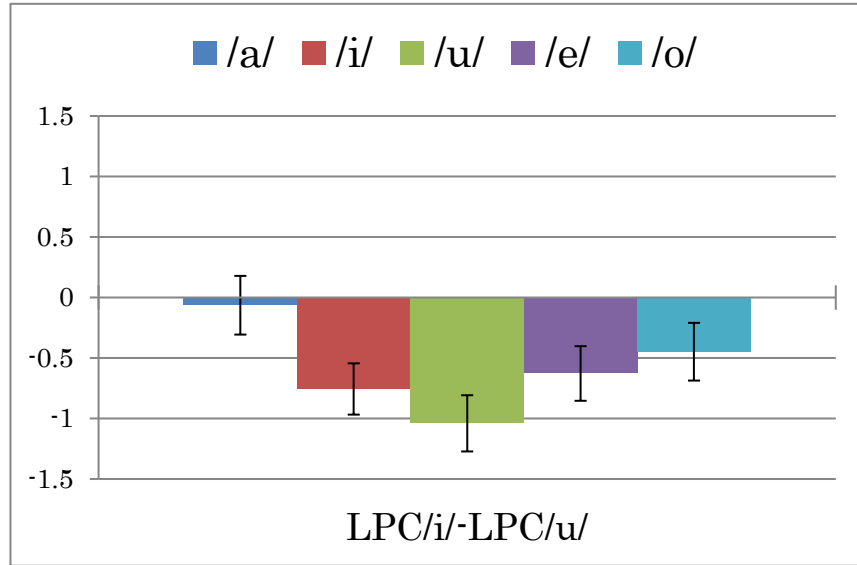


図 5.25 LPC/i/と LPC/u/の平均嗜好度の差と 95%信頼区間

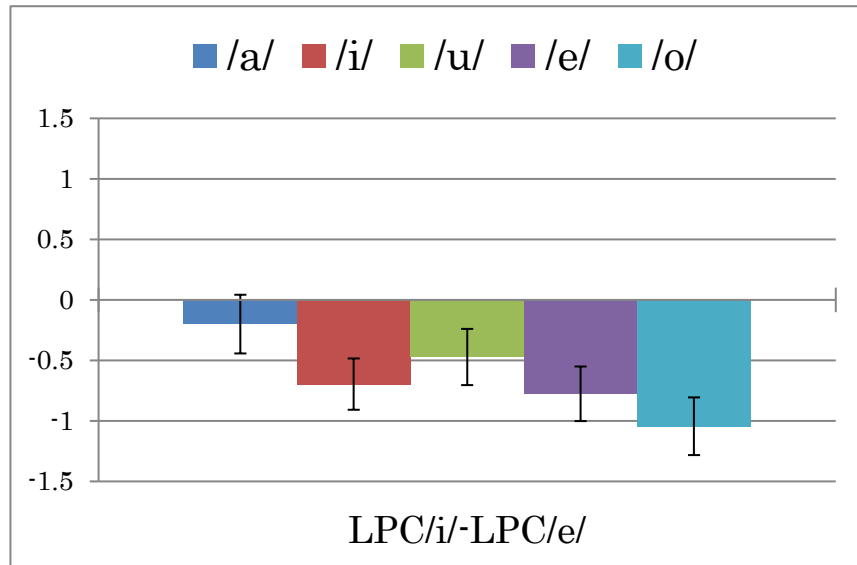


図 5.26 LPC/i/と LPC/e/の平均嗜好度の差と 95%信頼区間

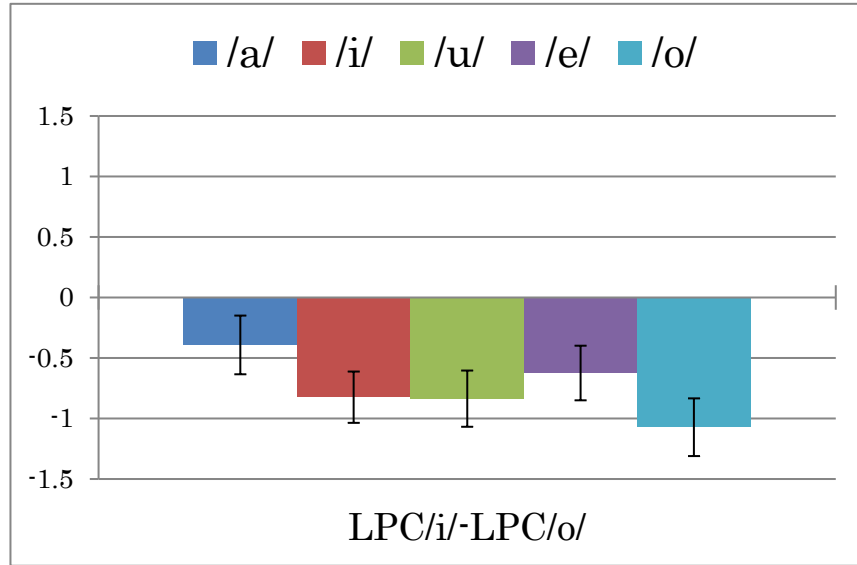


図 5.27 LPC/i/と LPC/o/の平均嗜好度の差と 95%信頼区間

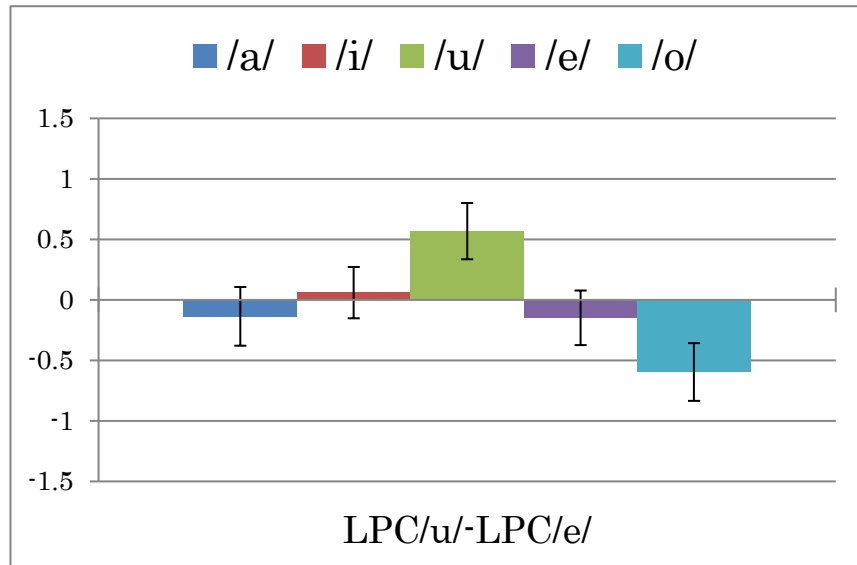


図 5.28 LPC/u/と LPC/e/の平均嗜好度の差と 95%信頼区間

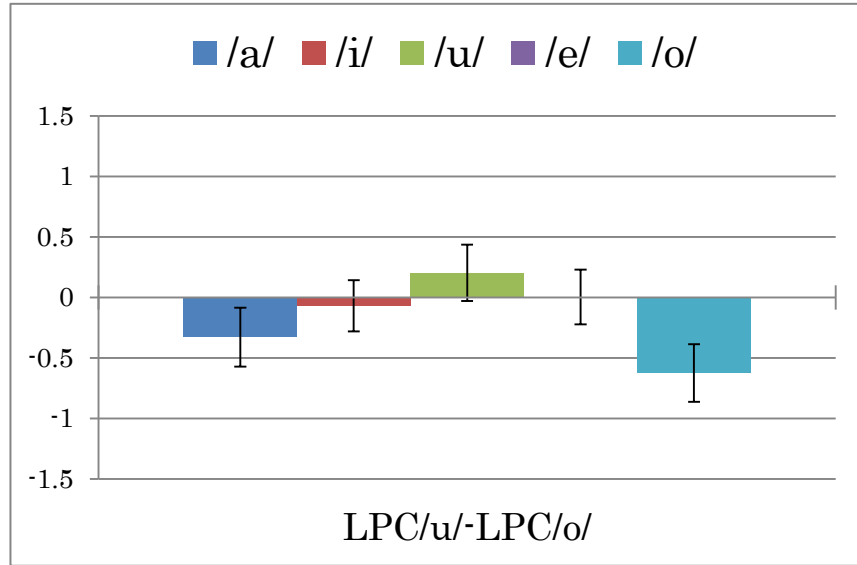


図 5.29 LPC/u/と LPC/o/の平均嗜好度の差と 95%信頼区間

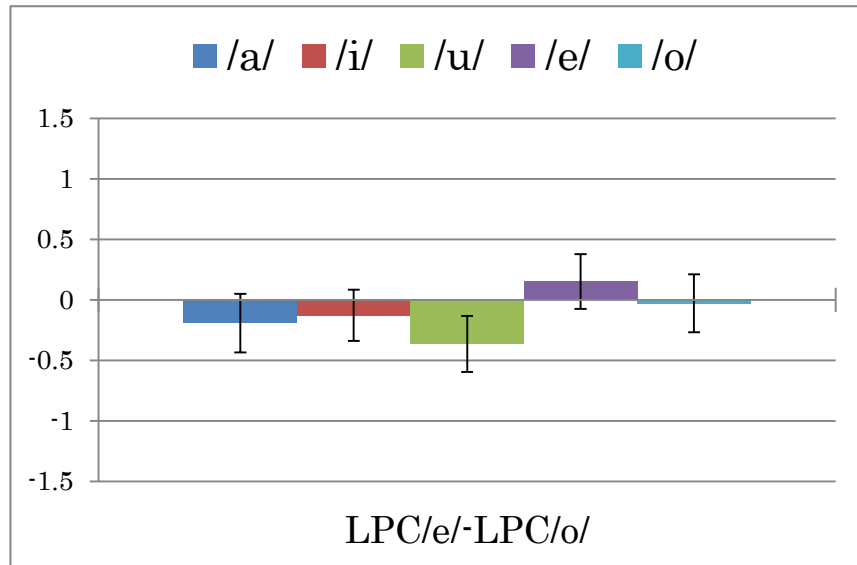


図 5.30 LPC/e/と LPC/o/の平均嗜好度の差と 95%信頼区間

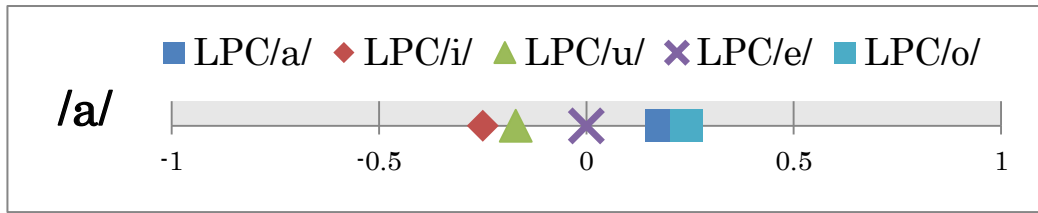


図 5.31 /a/の一次元尺度

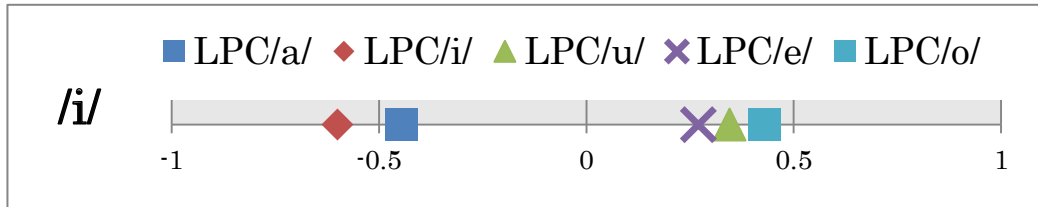


図 5.32 /i/の一次元尺度

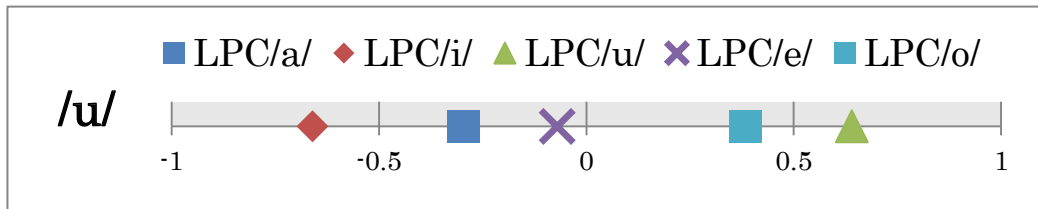


図 5.33 /u/の一次元尺度

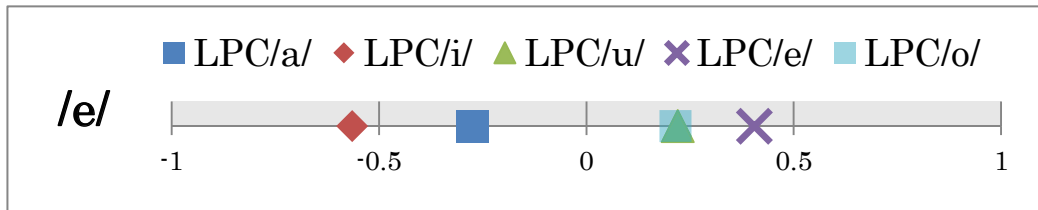


図 5.34 /e/の一次元尺度

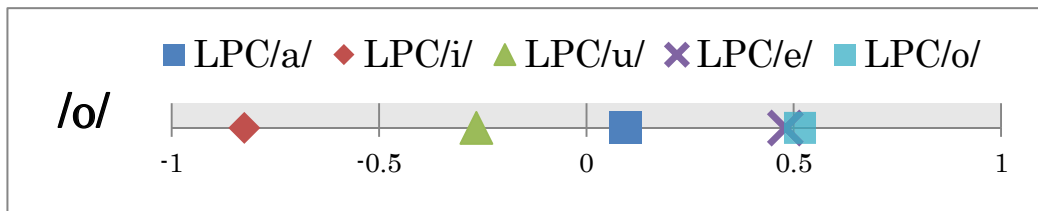


図 5.35 /o/の一次元尺度

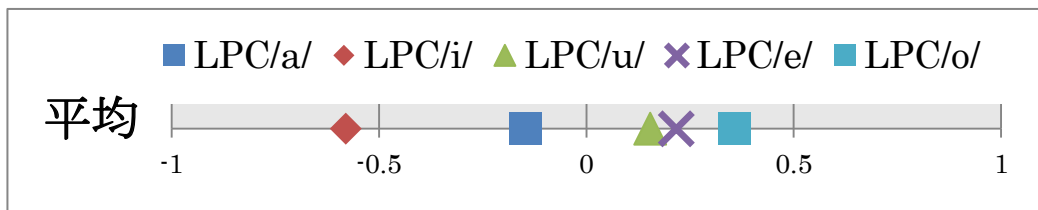


図 5.36 一次元尺度の平均値

第6章 スペクトル包絡のマッピングに基づく 音質改善方法の検討

本章では、提案する発声支援システムに用いる発声音声補正に用いるために、通常音声と EL 音声の LPC 特徴量を用いてスペクトル補正を行う。先行研究[44]により、EL 音声を補正するにあたって、音源情報のみではなく、包絡情報も並行して補正を行う必要があるのではないかと検討されている。そこで、スペクトル補正による EL 音声の音質の改善効果について調査した。

6.1 スペクトル包絡のマッピング概要

スペクトル包絡のマッピングに基づく音質改善システムの基本構成を図 6.1 に示す。まず、入力音声(改善したい音声)を LPC 分析し、音源情報(残差信号)と声道情報(スペクトル包絡)に分離する。スペクトル包絡は入力音声の音響モデルによりクラスタリングされ、分布間距離表により最も似ている目的音声の音響モデルのクラスに変換される。変換されたクラスより、目的音声のスペクトル包絡を求め、最初に分離した残差信号と LPC 合成することにより、補正された音声を作成する。

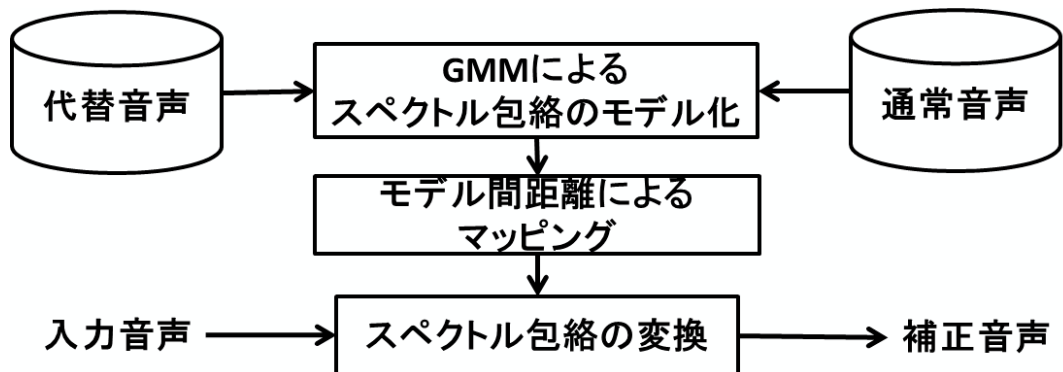


図 6.1 音質改善システムの基本構成

6.2 スペクトル補正音声作成

6.2.1 GMMによる音声モデルの作成

GMM[48]による音声モデルを作成する際の特徴量には, praat の LPC 分析による LPC 係数を用いて, HTK[49]内の HInit, HRest を使用した. HInit は k-means アルゴリズムを用いて, 状態ごとに特徴パラメータをクラスタリングし, 初期パラメータ(ガウス分布の平均ベクトルと分散共分散表列)を生成する. HRest は, HInit で生成された初期モデルに対して EM 学習によりパラメータの再推定を行う. 表 6.1 にモデルを作成するのに用いた設定パラメータを記す.

表 6.1 モデル作成時の設定パラメータ

Praat LPC(burg)	Prediction order	20
	Window length(s)	0.025
	Time step(s)	0.005
	Pre-emphasis frequency(Hz)	50.0
HTK (HInit&HRest)	SOURCEKIND	LPC
	TARGETKIND	LPC
	SAVEGLOBOPTS	TRUE
	KEEPDISTINCT	F
	prototype の出力状態	1
	Mixture	1024

6.2.2 GMM モデル間距離を用いたマッピング

疑似声帯音声のスペクトル包絡を、目的とする通常音声のスペクトル包絡へ近づける変換を行うために、元の通常音声と疑似声帯音声のモデル間距離を用いて対応関係の学習を行う。距離の測定にはユークリッド距離を用いてそれぞれの一番近いクラスをマッピングした。疑似声帯音声と通常音声のスペクトル包絡の特徴量は GMM によりモデル化され、クラスタリングされている。ここで、疑似声帯音声の音声モデルのクラスと通常音声の音声モデルのクラス間の距離を計算し、最もクラス間距離の近いクラス同士をマッピングすることで変換表を作成する。

図 6.2 はマッピングの概念図である。図 6.2 中の変換表の横軸が通常音声のクラスであり、縦軸が疑似声帯音声のクラスである。もし、疑似声帯音声のクラス 2 に対応した疑似声帯音声のスペクトル包絡の特徴量が入力された場合、変換表により通常音声のクラス 1 に変換される。

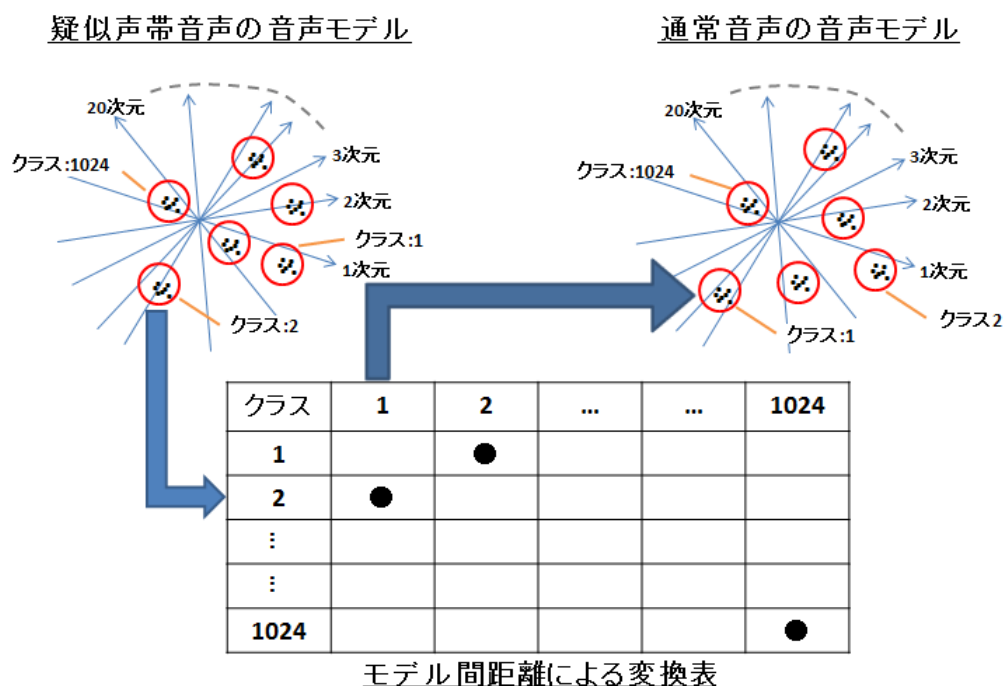


図 6.2 GMM モデル間距離を用いたマッピングの概念図

6.2.3 LPC 分析合成による補正音声の生成

ここでは、LPC 分析合成による補正音声の生成方法について説明する。6.2.2 の GMM モデル間距離を用いたマッピングにより、疑似声帯音声のスペクトル包絡の特徴量は通常音声のスペクトル包絡の特徴量へと変換することができる。よって、未知の疑似声帯音声が入力された場合、LPC 分析により音源情報と包絡情報に分離された後、包絡情報は変換表に従って通常音声のスペクトルへと変換される。変換されたスペクトルを声道フィルタとして、最初に分離した音源情報へ畳み込むことで補正音声を生成する。図 6.3 は LPC 分析合成による補正音声の生成過程のフローチャートである。

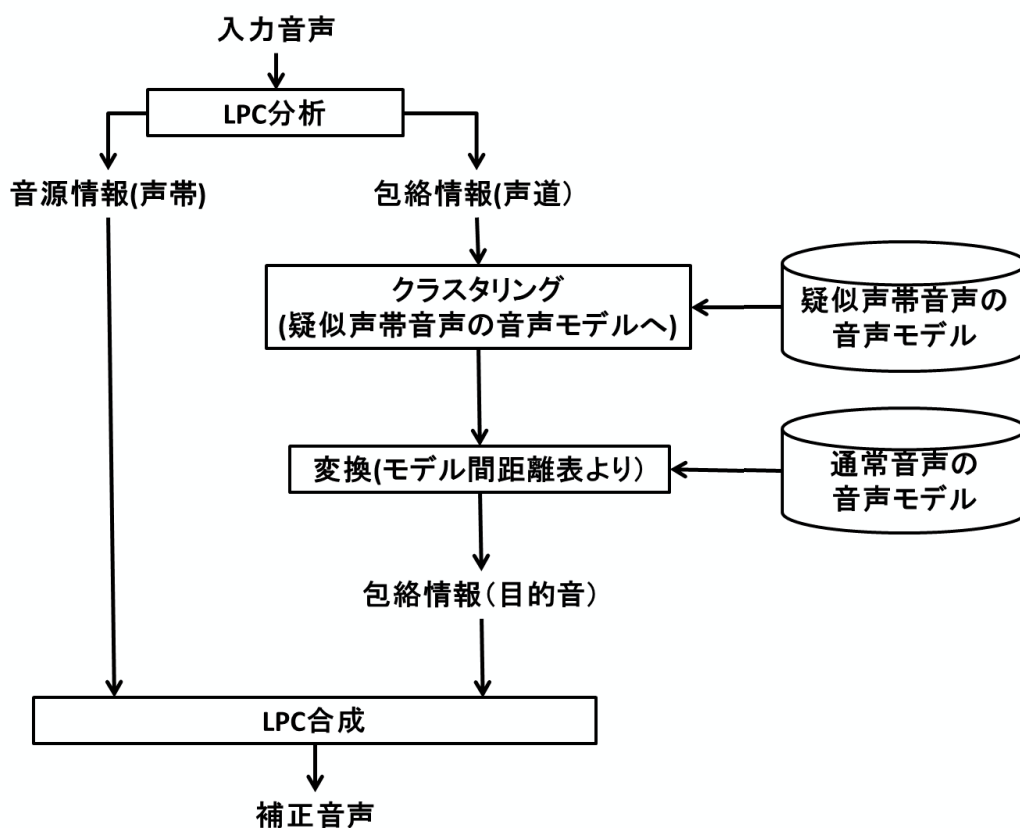


図 6.3 LPC 分析合成による補正音声の生成過程

6.3 音声収録

GMM による音声モデルを作成し、スペクトル補正を行うために、SP 音声、EL 音声、振動スピーカによる LPC 残差波音声の三種類の音声を収録した。

前章と同じように EL は DENCOR 社 DEL-Y200 「ユアトーンⅡ ゆらぎ」を使用し、振動スピーカはコバテル社 VR3000 を用いた。LPC 残差波は第 3 章の主観評価で最も評価のよかった母音/o/より抽出した音源を使用した。収録場所は防音室で行い、収録音声は音韻バランスのとれた 543 単語と、未知の音声に対して補正を行うために「あいうえお」、「おはよう」、「こんにちは」、「ありがとう」、「さようなら」の 5 単語を収録した。音韻バランスのとれた 543 単語については付録 1 に示す。

使用した EL は音量と声の高さをそれぞれ「1」から「9」まで調整する機能がついており、声の高さは「4」に設定し、音量は振動スピーカとほぼ同じ大きさになるよう「7」に設定し収録を行った。

収録した音声ファイルは 48kHz, 16 ビットでデジタル録音である。この収録した音声ファイルを 16kHz にダウンサンプリングし、分析、聴取実験を行った。

収録マイクは、口元から 15cm のところに設置した無指向性マイクである。収録条件を表 6.2 にまとめる。

表 6.2 収録条件

話者	EL の扱いに慣れた健常者 1 名
収録マイク	口元に設置した無指向性マイク
収録場所	防音室（周囲雑音なし）
収録音声	音韻バランスのとれた 543 単語(付録 1) 「あいうえお」 「おはよう」 「こんにちは」 「ありがとう」 「さようなら」
音源の種類	通常発声音声 EL（声の高さ 4, 音量 7） 振動スピーカ (/o/ の LPC 残差波形)

6.4 客観評価による比較

スペクトル包絡のマッピングに基づく音質改善方法により生成した音声の客観的評価を行う。

6.4.1 5母音のスペクトル包絡のマッピング

スペクトル包絡のマッピングに基づく音質改善が行われているのか評価するために、EL と/o/より抽出したLPC残差波音声の5母音による発声を用いてスペクトルの変換を行った。図6.4から図6.8に、EL音声の5母音に対してスペクトル包絡の補正を行った結果を示し、図6.9から図6.13にLPC残差波音声の5母音に対してスペクトル包絡の補正を行った結果を示す。スペクトルはoctaveを用いて作成し、作成時の設定は表6.3に示す。また、通常音声とのケプストラム係数のユークリッド距離を算出し、表6.4にまとめた。

スペクトル包絡のマッピングを行うことにより、1000[Hz]付近の低域のパワーが補正され、3000[Hz]付近の帯域での特性が通常音声に近くなっていることがわかる。また、補正によりホルマントが強調されている箇所があるが、位置を通常音声に近づけるような補正が行われている箇所は少ないことがわかった。

ケプストラム係数のユークリッド距離を比較すると、EL音声を補正すると差が大きくなったが、LPC残差波音声では補正により差が小さくなった。LPC残差波音声に対してはスペクトル包絡の変換による補正が有効であることがわかった。

表 6.3 スペクトル作成時の設定

分析区間	Center 1.0[s]
窓幅	25[ms]
フレームシフト幅	5[ms]
フーリエ変換の次数	1024
フレーム数	195
リフタリング次数	55

表 6.4 ケプストラム係数の通常音声とのユークリッド距離

	/a/	/i/	/u/	/e/	/o/	Sum total
EL	1.570127	1.047754	1.607602	0.910538	1.537923	6.673944
変換 EL	1.571969	1.224881	1.372029	1.035174	1.730226	6.934279
LPC	0.97686	0.776961	1.347197	1.402736	1.270773	5.774527
変換 LPC	0.934634	0.637464	1.214916	0.740868	0.958773	4.486655

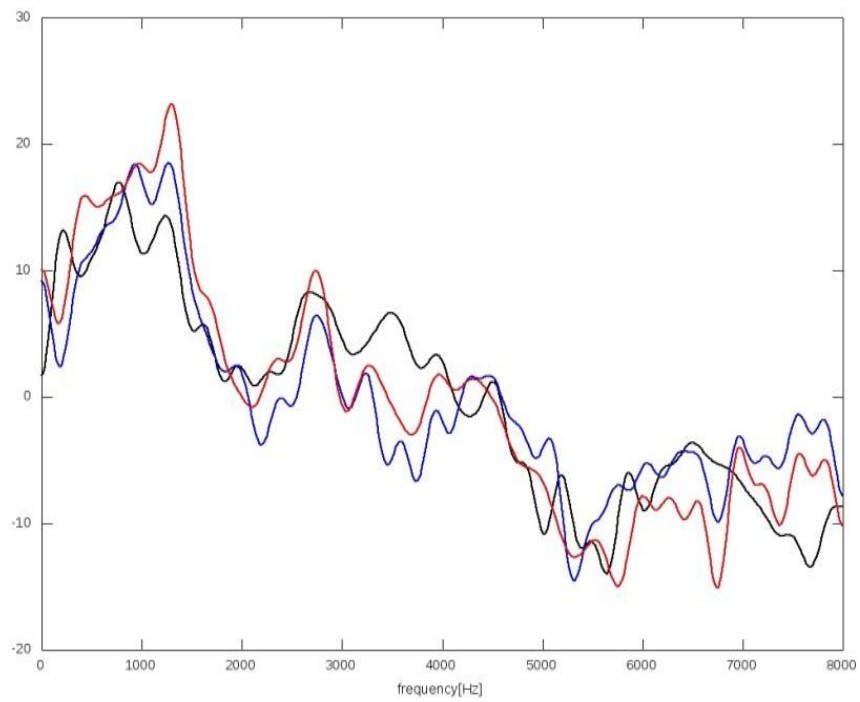


図 6.4 /a/のスペクトル(黒：SP 青：EL 変換前 赤：EL 変換後)

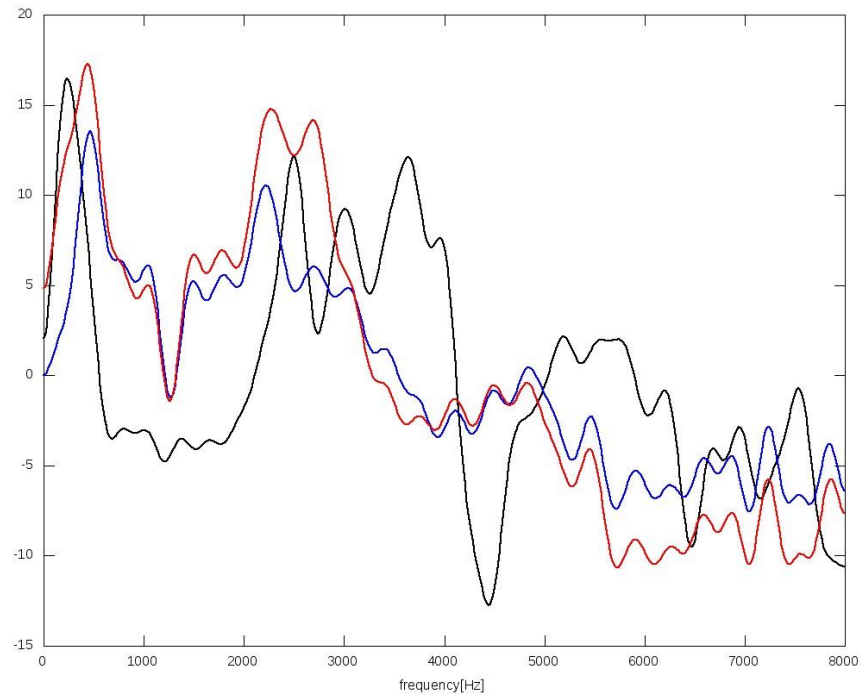


図 6.5 /i/のスペクトル(黒：SP 青：EL 変換前 赤：EL 変換後)

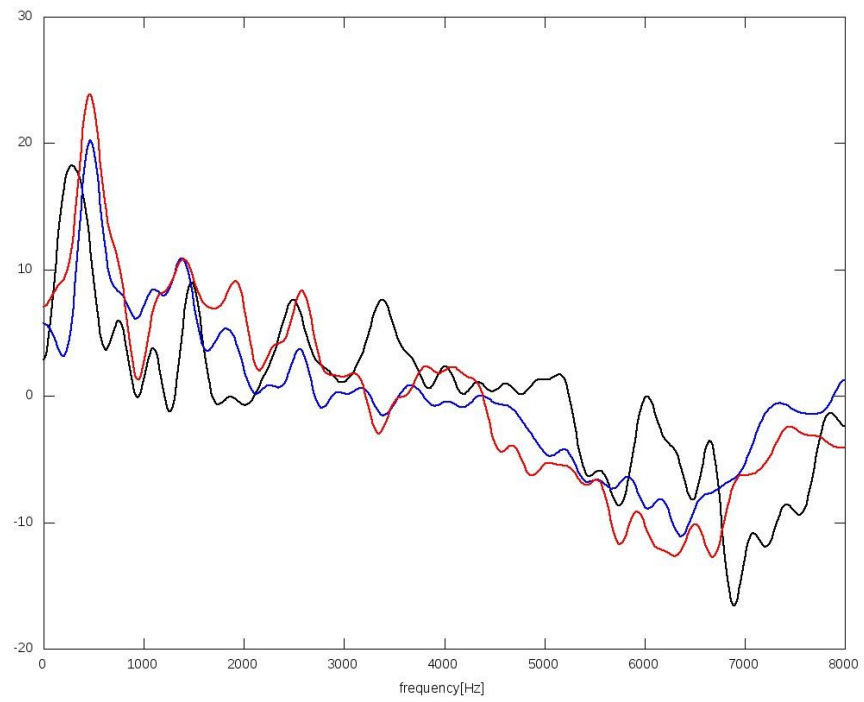


図 6.6 /u/のスペクトル(黒：SP 青：EL 変換前 赤：EL 変換後)

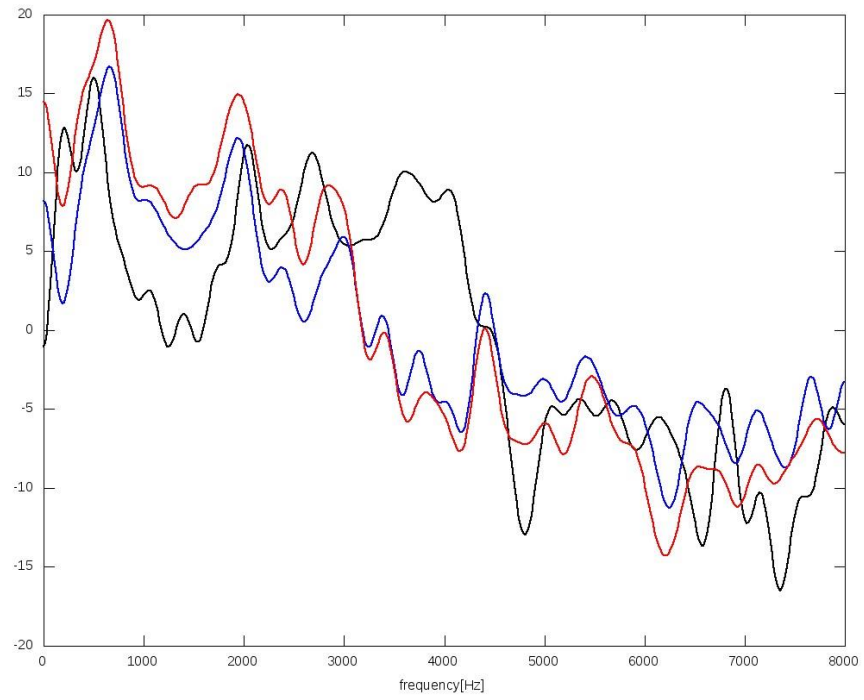


図 6.7 /e/のスペクトル(黒：SP 青：EL 変換前 赤：EL 変換後)

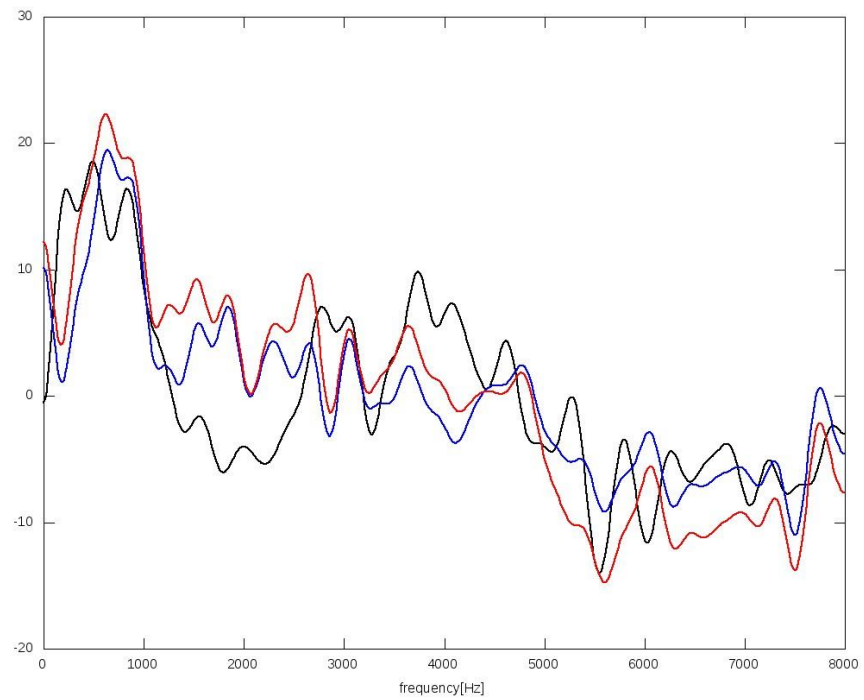


図 6.8 /o/のスペクトル(黒：SP 青：EL 変換前 赤：EL 変換後)

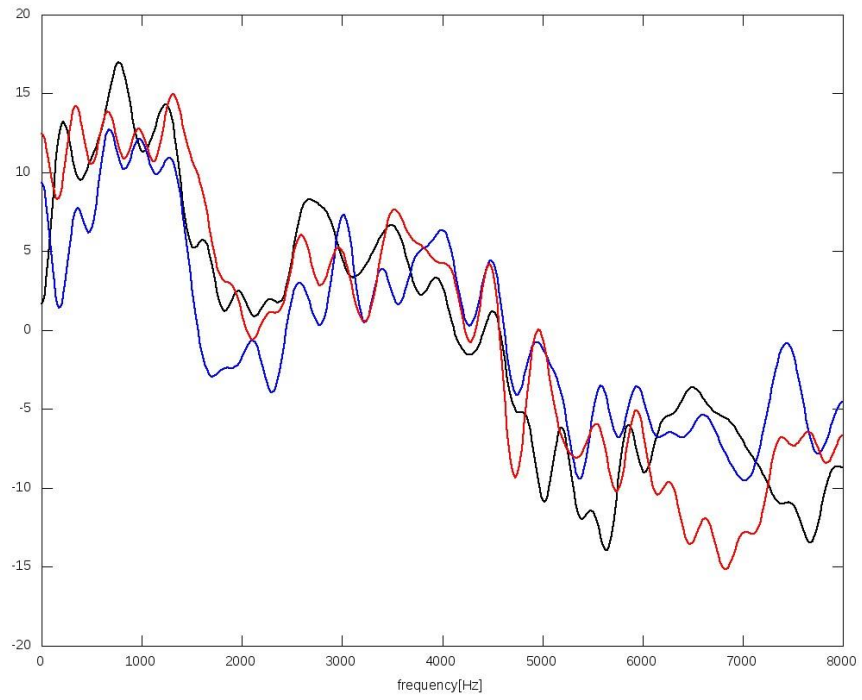


図 6.9 /a/のスペクトル(黒：SP 青：LPC 変換前 赤：LPC 変換後)

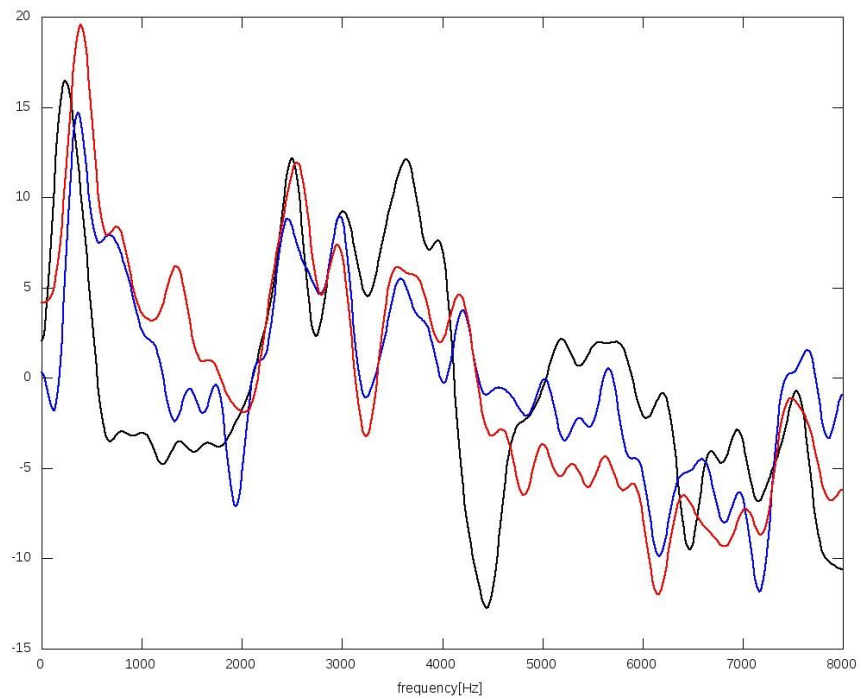


図 6.10 /i/のスペクトル(黒：SP 青：LPC 変換前 赤：LPC 変換後)

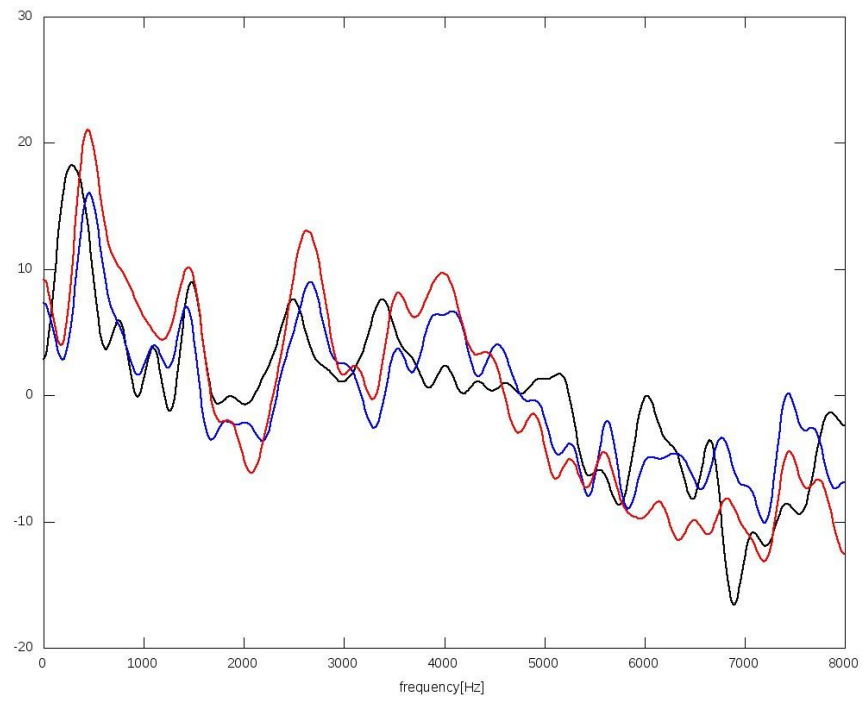


図 6.11 /u/のスペクトル(黒：SP 青：LPC 変換前 赤：LPC 変換後)

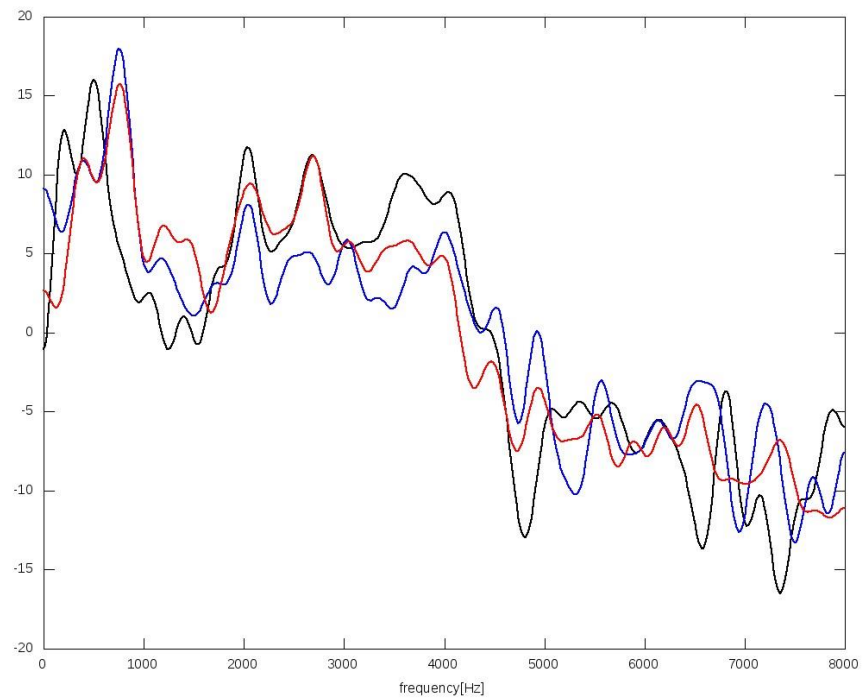


図 6.12 /e/のスペクトル(黒：SP 青：LPC 変換前 赤：LPC 変換後)

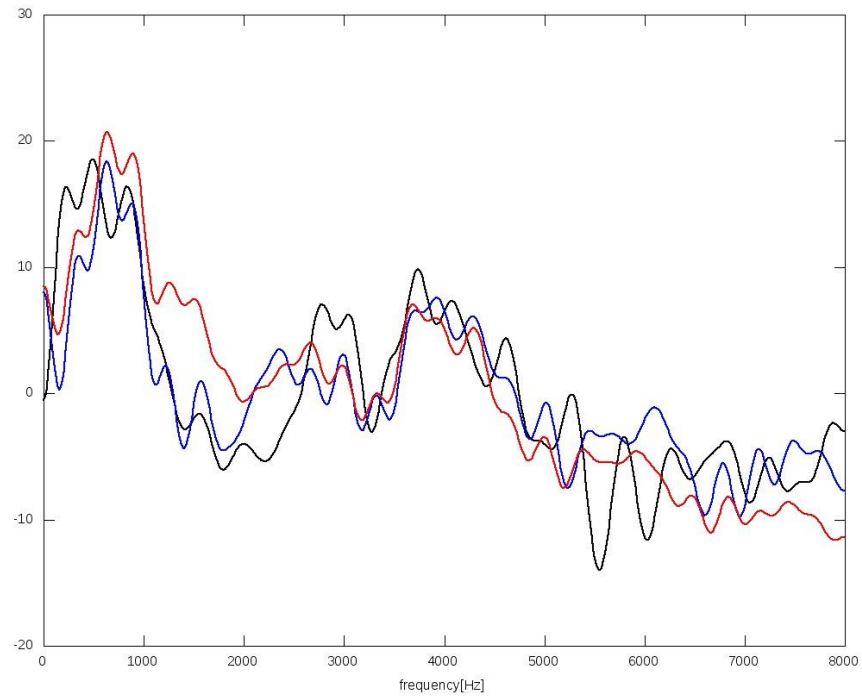


図 6.13 /o/のスペクトル(黒：SP 青：LPC 変換前 赤：LPC 変換後)

6.4.2 未知単語へのスペクトル包絡のマッピング

未知単語である「あいうえお」、「おはよう」、「こんにちは」、「さようなら」、「ありがとう」のスペクトル変換の評価を行う。スペクトル変換の効果を見るため、スペクトル変換前後と目的音声のスペクトル包絡の比較例を示す。図 6.14 から図 6.18 は、単語「あいうえお」、図 6.19 から図 6.23 が単語「おはよう」、図 6.24 から図 6.28 が単語「こんにちは」、図 6.29 から図 6.33 が単語「さようなら」、図 6.34 から図 6.38 が単語「ありがとう」のそれぞれ通常音声、補正前の EL 音声、補正後の EL 音声、補正前の LPC 残差波音声、補正後の LPC 残差波音声である。また、赤の点がホルマントの位置であり、青の線が基本周波数である。

簡単な音質評価により低域での改善が見られたが、音韻の変化する部分などで音質の劣化が激しかった。そこでスペクトルグラムにより通常音声との比較を行った。その結果、スペクトル包絡のマッピングによる補正を行うことで、ホルマントの不連続性や、通常音声のスペクトルへの変換が正確に行われておらず、逆に音質が劣化してしまった。また、振動スピーカを用いた発声は、/a/などの単音の発声は EL と同様にできるが、単

語の発声は難しく，音声の特徴が出にくいことがわかった.

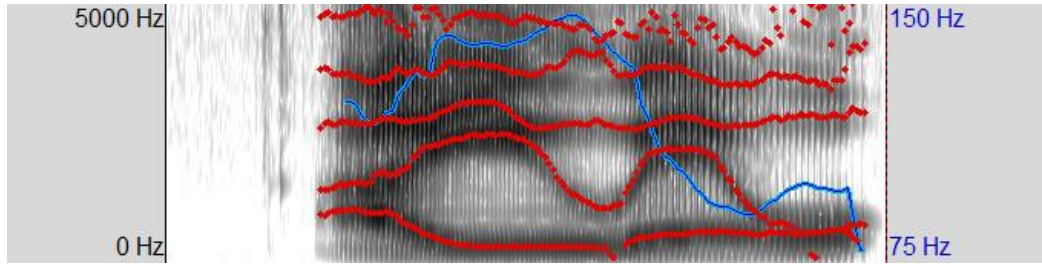


図 6.14 「あいうえお」の通常音声のスペクトル

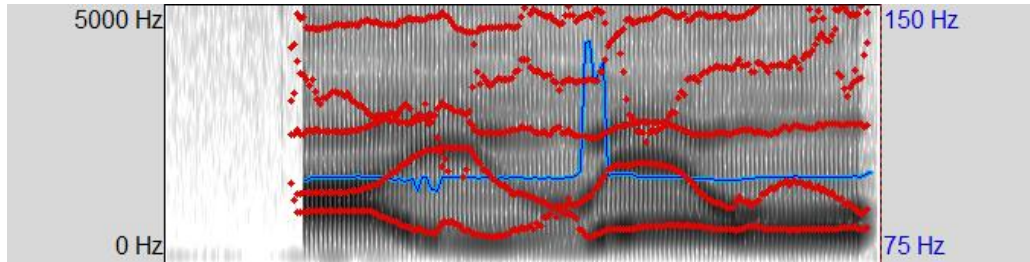


図 6.15 「あいうえお」の補正前の EL 音声のスペクトル

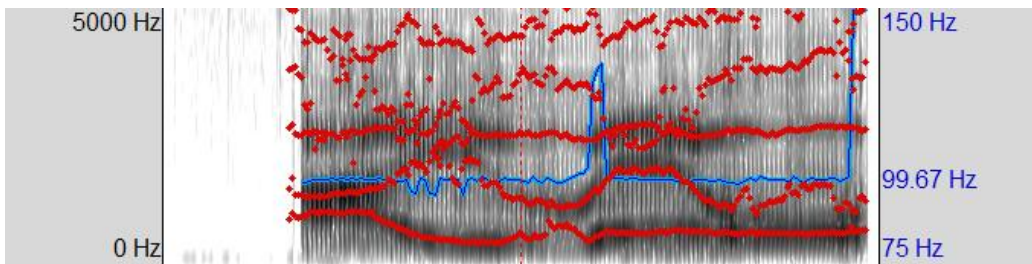


図 6.16 「あいうえお」の補正後の EL 音声のスペクトル

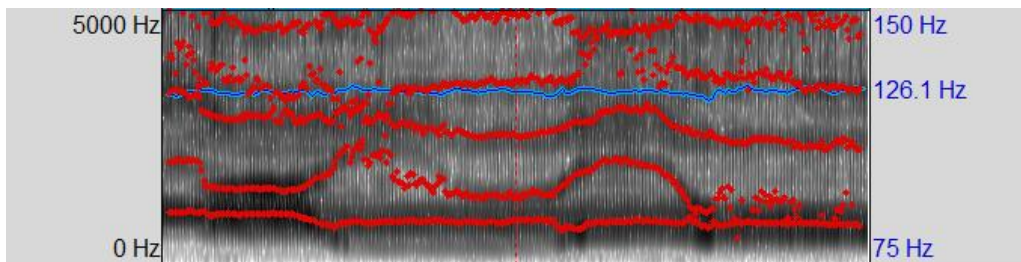


図 6.17 「あいうえお」の補正前の LPC 残差波音声のスペクトル

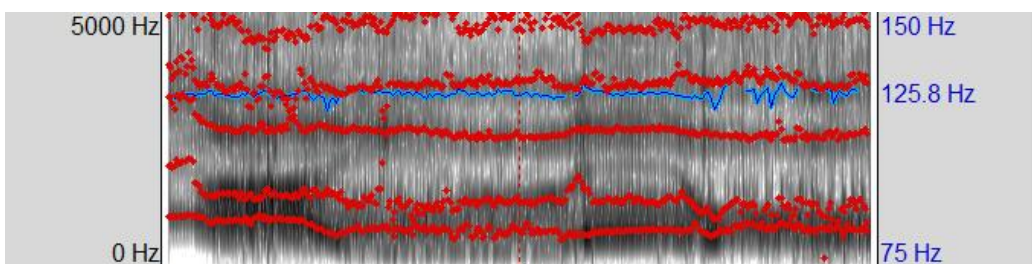


図 6.18 「あいうえお」の補正後の LPC 残差波音声のスペクトル

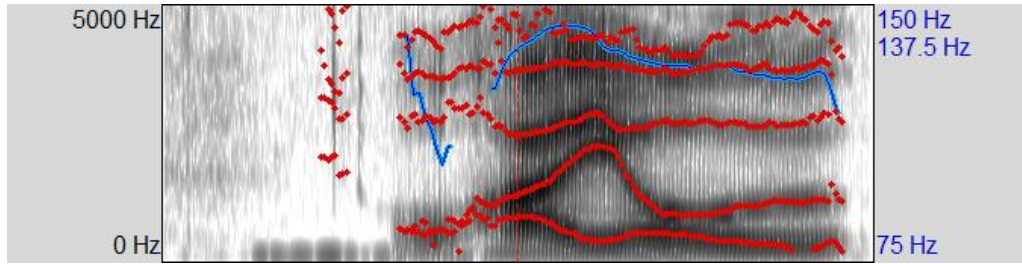


図 6.19 「おはよう」の通常音声のスペクトル

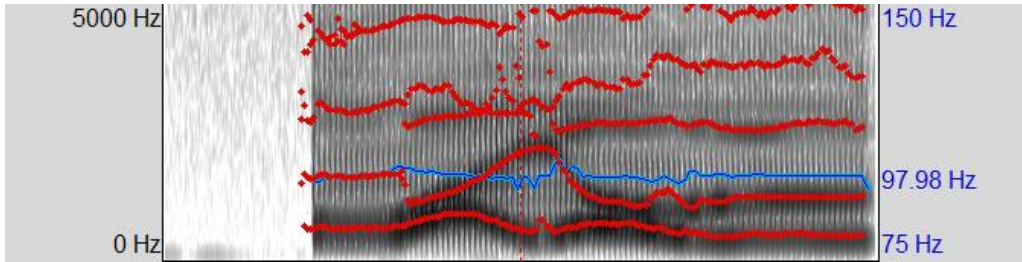


図 6.20 「おはよう」の補正前の EL 音声のスペクトル

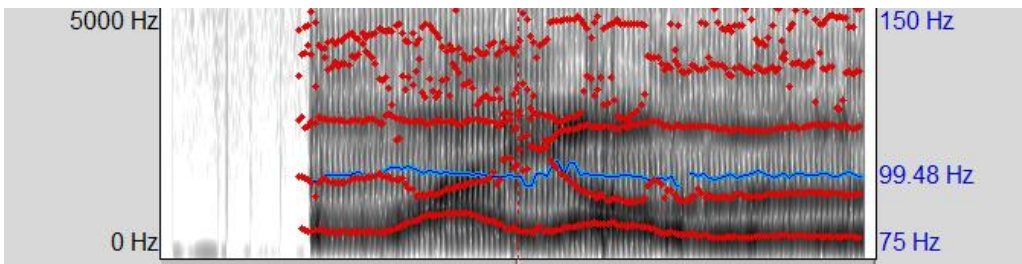


図 6.21 「おはよう」の補正後の EL 音声のスペクトル

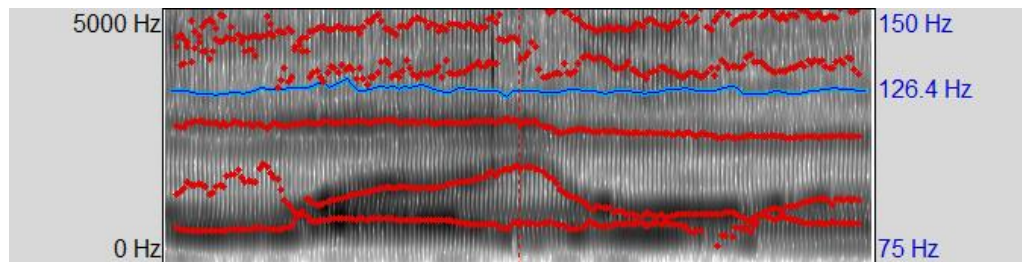


図 6.22 「おはよう」の補正前の LPC 残差波音声のスペクトル

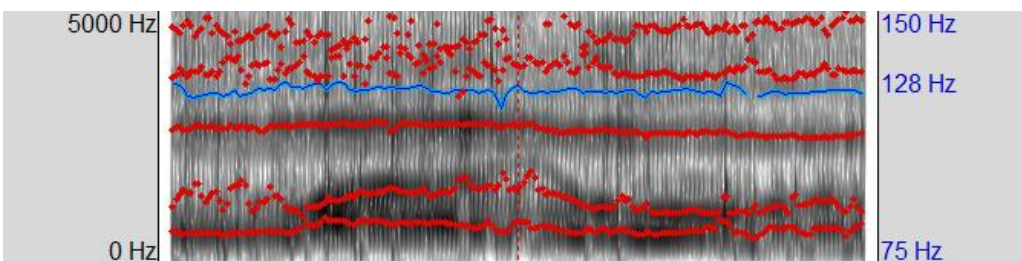


図 6.23 「おはよう」の補正後の LPC 残差波音声のスペクトル

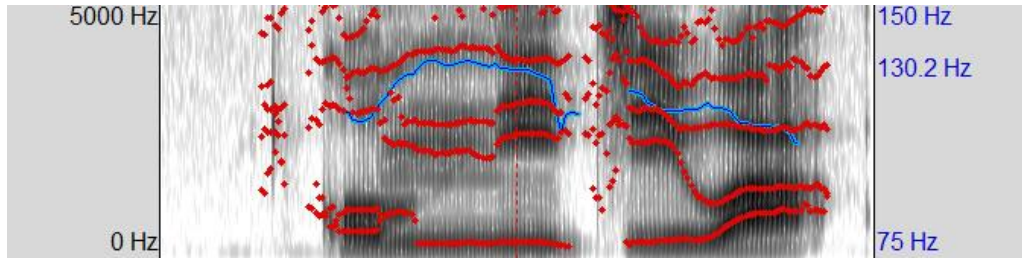


図 6.24 「こんにちは」の通常音声のスペクトル

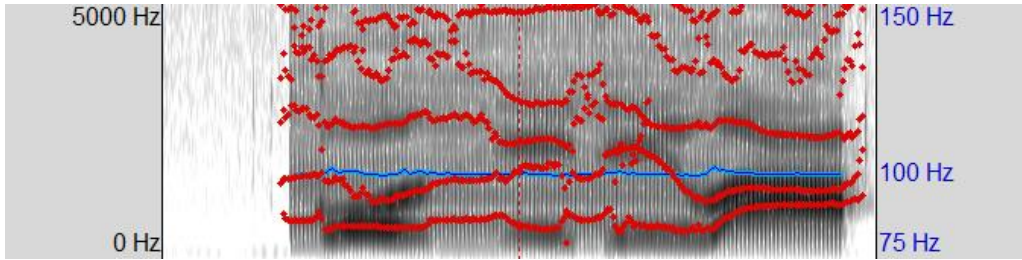


図 6.25 「こんにちは」の補正前の EL 音声のスペクトル

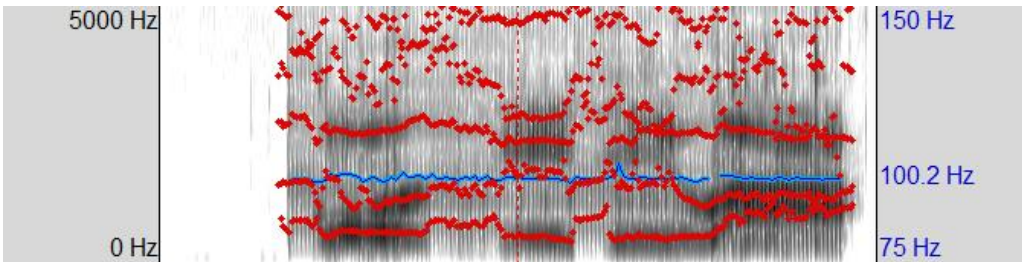


図 6.26 「こんにちは」の補正後の EL 音声のスペクトル

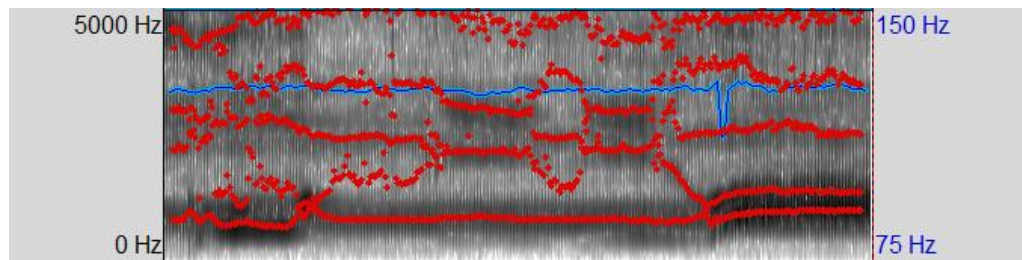


図 6.27 「こんにちは」の補正前の LPC 残差波音声のスペクトル

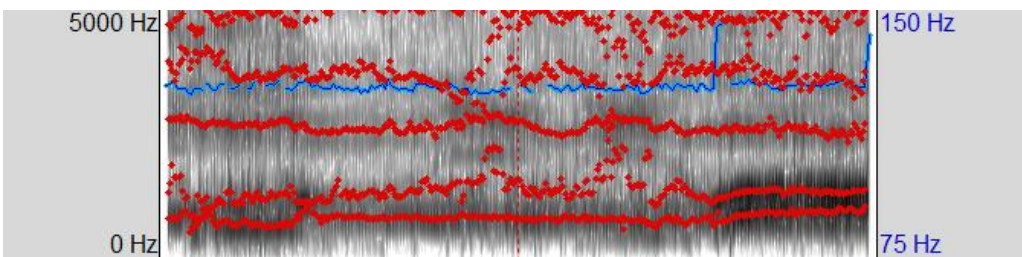


図 6.28 「こんにちは」の補正後の LPC 残差波音声のスペクトル

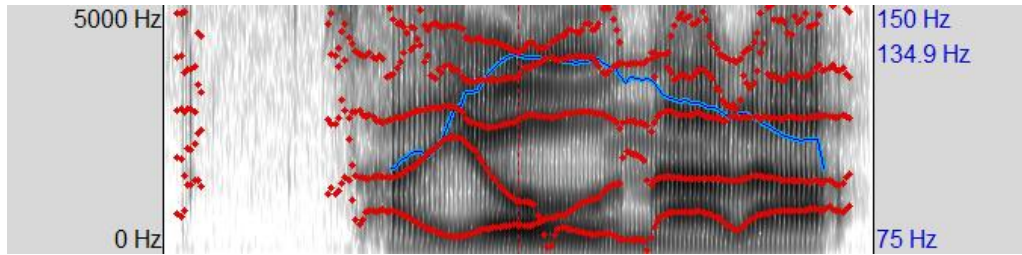


図 6.29 「さようなら」の通常音声のスペクトル

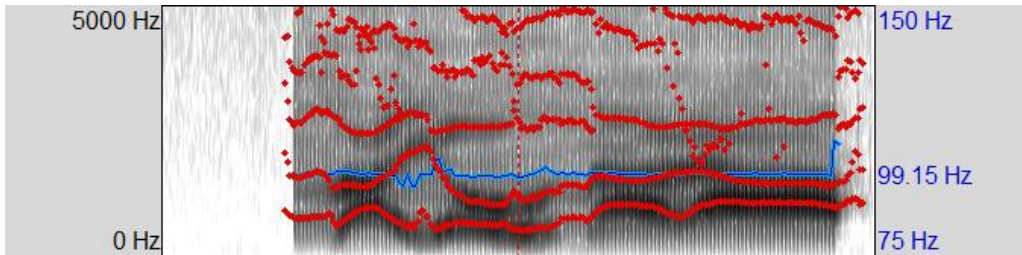


図 6.30 「さようなら」の補正前の EL 音声のスペクトル

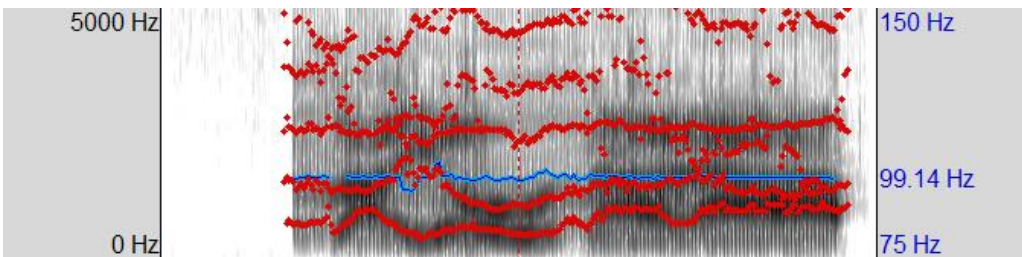


図 6.31 「さようなら」の補正後の EL 音声のスペクトル

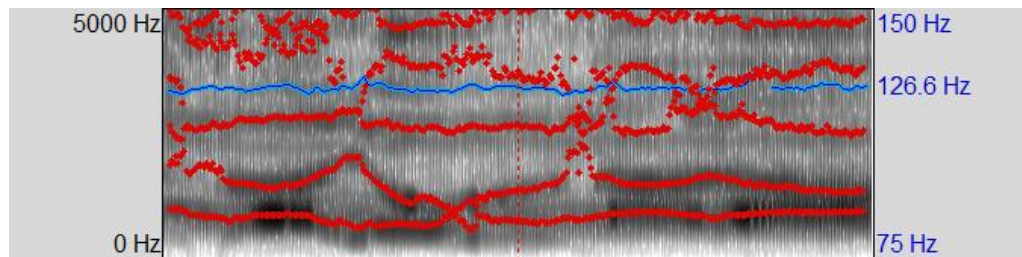


図 6.32 「さようなら」の補正前の LPC 残差波音声のスペクトル

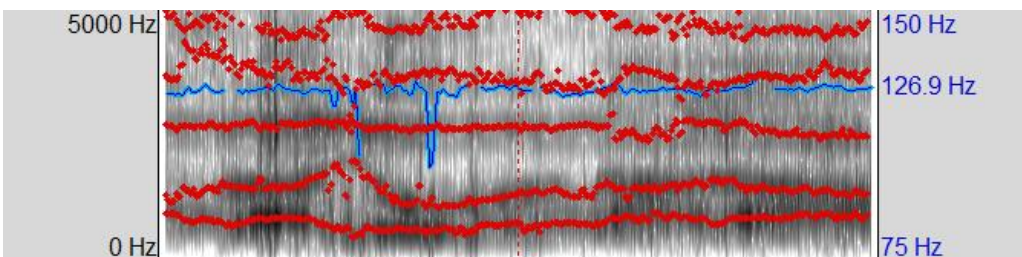


図 6.33 「さようなら」の補正後の LPC 残差波音声のスペクトル

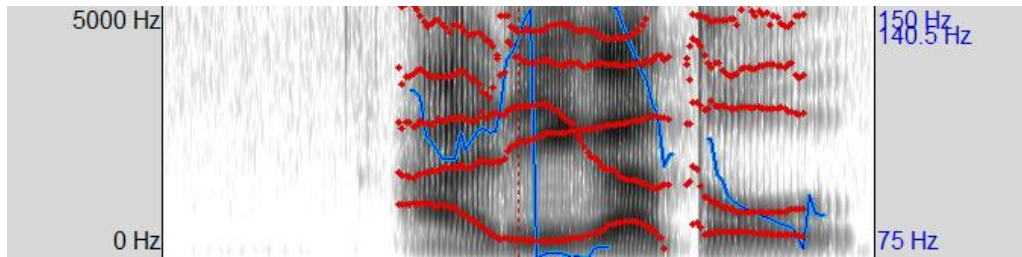


図 6.34 「ありがとう」の通常音声のスペクトル

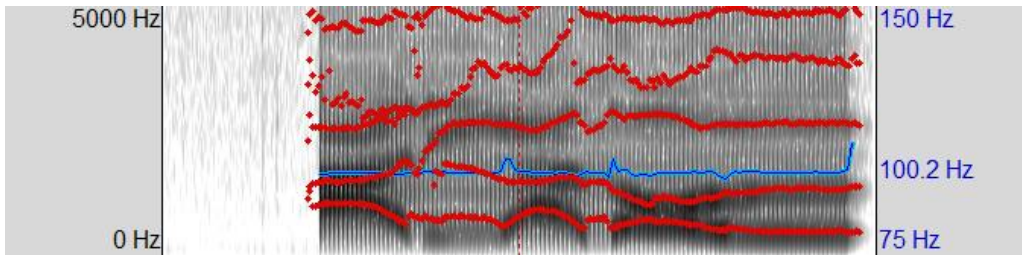


図 6.35 「ありがとう」の補正前の EL 音声のスペクトル

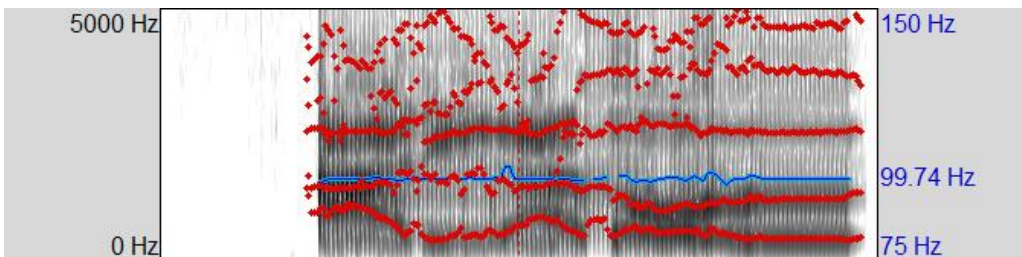


図 6.36 「ありがとう」の補正後の EL 音声のスペクトル

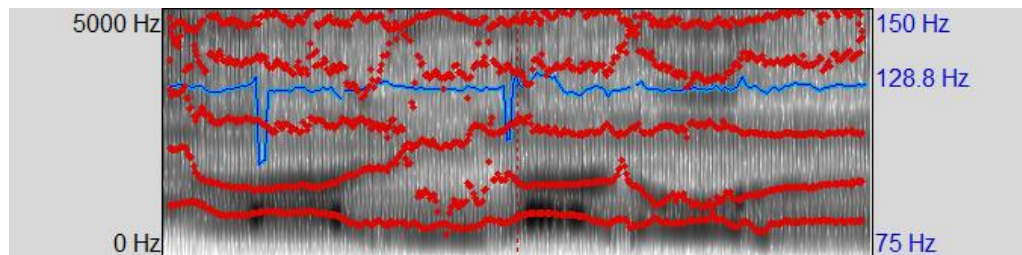


図 6.37 「ありがとう」の補正前の LPC 残差波音声のスペクトル

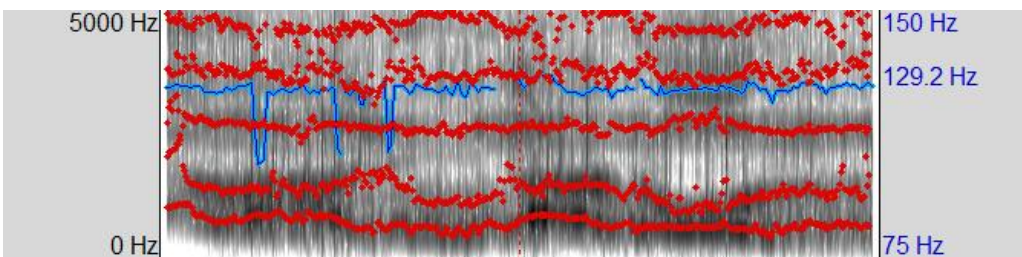


図 6.38 「ありがとう」の補正後の LPC 残差波音声のスペクトル

第 7 章 結論

7.1 本論文の要約

我々人間は、非常に高度なコミュニケーションを行っており、中でも、音声によるコミュニケーションは、言語情報のみならず、発話意図、感情と言った情報も含まれており、これを音声信号にのせてやり取りを行っている。高度なコミュニケーションを行うことで、社会的適応能力の向上や脳の発達に期待ができるという報告もある。円滑な音声によるコミュニケーションをサポートすることは、我々のQoL向上に対して、非常に有効である。

しかし、年々進行する高齢化により、高齢者との音声言語によるコミュニケーションをすることにおいて、加齢による発生期間の衰えによる発声音声の劣化が生じてしまう。また、発声器官の喪失による代替発声を余儀なくされる高齢者もいる。代替発声法の中でも近年電気式人工喉頭を用いる発声法は扱いが容易という率腕需要が伸びてきているが、非常に不明瞭な音声である。このような劣化音声の明瞭性の低下の原因は何か、あるいはどのような音響的特徴があるか、といった基礎的な研究は非常に少ない。

そこで、本研究では、劣化音声の基礎的な分析から行い、この音響的特徴の対応関係を適切に表現できる変換パラメータを選定し、変換パラメータに基づいて劣化音声から明瞭な音声へと変換する音声処理アルゴリズムの検討を行う。

第1章では、音声コミュニケーションの必要性和高齢者、喉頭摘出者の音声コミュニケーションの現状と課題について述べ、本研究の位置づけと目的を明確化した。

第2章では、高齢者音声の明瞭度に関する音響的特徴について分析した結果について述べた。まず、使用した高齢者音声データベースについて説明し、そのデータベースを用いて高齢者音声の「明瞭度の低い音声」と「明瞭度の高い音声」の2つの比較を、遷移量(ケプストラム距離)と、ホルマント周波数について分析した。その結果、遷移量においては、明瞭度の低い音声は、高い音声と比較すると、遷移量が低下していることがわかった。ホルマント周波数についてF1-F2平面上で比較すると、母音/u/は明瞭度の高低による変化はあまり見られなかったが、母音/a/, /i/, /e/, /o/については、音声の明瞭度が低くなると、母音/u/に収束していることがわかった。

第3章では、第2章のホルマントの結果を踏まえ、高齢者音声の明瞭度改善方法について述べた。まず、明瞭度を改善するためにホルマント周波数シフトによる明瞭度改善アルゴリズムを構築した。このアルゴリズムは、第2章で分析したホルマント周波数の差を補償するというものである。この手法の有効性について、補償するホルマント周波

数差分値が既知である場合と、未知である場合について、シェッフェの一对比較法に基づいた聴取実験により検証を行った。差分値が既知である場合には、明瞭度が 10 単語中 10 単語とも明瞭度が改善した。差分値が未知である場合では、F1-F2 平面上に 5 母音が構成する五角形の面積が広がるように母音/u/を中心に補償を行うこととした。その結果、補償効果に個人差が存在することがわかった。これは、個人のホルマント周波数の傾向にばらつきが存在するためであると考えられる。未知の場合においては、ホルマント周波数の傾向が、一律に母音/u/に収束しているような傾向の場合においては有効であると考えられる。

第 4 章では、電気式人工喉頭音声の問題である「音声不明瞭である」、「電気式人工喉頭の発する振動音自体がノイズになってしまう」、という問題を解決するための電気式人工喉頭の発声方法を基に、新しい発声支援システムの提案を行った。

第 5 章では、電気式人工喉頭音声の問題である「音声不明瞭である」という問題に着目し、音源である電気式人工喉頭の発する信号を改善することによる音質改善を図り、効果の検証を行った。その結果、音声から抽出した LPC 残差波を音源として使用することで、音声の肉声感が増すことがわかった。

第 6 章では、電気式人工喉頭音声の問題である「音声不明瞭である」という問題に着目し、音声の包絡情報を、GMM によるマッピングに基づく音質改善方法を提案し、効果の検証を行った。その結果、物理量であるケプストラム距離上での改善は見られた。聴感印象では、マッピングがうまく行えている定量的な部分においては、改善がみられたが、マッピングがうまく行えなかった変化部分においては、滑らかな変化が表現できていないことがわかった。

7.2 今後の課題

本研究で得られた知見の発展及び汎用性を高めるために、今後、以下に挙げる課題を解決することが望ましい。

子音に関する音響的特徴量の分析及び明瞭性改善に関する調査

本研究においては、日本語において重要である5母音について解析を行い、明瞭化を行った。更なる明瞭化を図るために、子音についての明瞭性について、聴感印象とそれに対応する物理量について調査することが必要であると考えられる。

更なるマッピング対応付けの向上

本研究においてはGMMによる学習結果を、ユークリッド距離を用いて対応付けを行い、一部分においては改善傾向にあることがわかった。しかし、電気式人工喉頭音声やLPC残差波形を用いて発声した音声の渡りの部分（遷移部分）に関しては対応付けがうまくできていないことがわかった。そこで、更なる音質改善のために、急峻な変化を抑制するアルゴリズムの導入や、深層学習を用いた学習への転換などが必要であると考えられる。

リアルタイムで動作する装置の開発

本研究成果は、電話や補聴器などに応用できると考えられる。そのために、本論文で提案したアルゴリズムをリアルタイム処理に適した形でマイコン等を実装し、効果の検証を行うことが必要である。加齢による明瞭度の低下を改善するアルゴリズムは、リアルタイム処理に適していると考えられる。そのために、入力音声が低明瞭かどうかを判断するアルゴリズムや、第4章で述べたような低明瞭な音声となる要因であるF1-F2平面でのホルマント面積低下の傾向に外れた場合における補正方法の検討を行う必要がある。

謝辞

本研究を進め学位論文をまとめるに当たり、多くの支援とご指導を承りました。

指導教員である中藤良久教授には深く感謝しています。研究活動を通して、大変多忙な中、数々のご指導、ご鞭撻を賜りましたこと、ここに厚く御礼申し上げます。有難うございました。

本論文に対して貴重なご意見を頂きました、九州工業大学 芹川 聖一 教授、金 亨 燮 教授、水町 光徳 准教授に感謝申し上げます。芹川先生には、学会等でも貴重な御助言をして頂きました。金先生には、暖かい言葉を掛けて頂き、本論文執筆の大きな励みとなりました。水町先生には、本研究を遂行する上で御指導、貴重な御助言を頂きました。有難うございました。

折りに触れて御指導、貴重なご助言を頂きました 張 力峰 准教授に心より感謝致します。国際会議、国際交流へ向けた御指導、大変お世話になりました。有難うございました。

本研究を円滑に行う上で研究に関連する書籍、器具の調達等で本当にお世話になりました、原田 勝也 技術職員に心より感謝致します。有難うございました。

中藤研究室、水町研究室の皆様、二矢田研究室の先輩方には、本研究に関する数多くの御討論や御意見、そして聴取実験にご協力を頂きました。ここに深く感謝致します。

父、母には、博士後期課程への進学を快く認めていただき、博士取得のために必要な経済的、精神的援助をして頂きました。心より感謝致します。有難うございました。

最後に、大学・大学院在学中に渡って暖かく見守っていただいた親族、友人の皆様我心から感謝致します。

参考文献

- [1] Hall E.T. ,The Silent Language ,New York, 1959.
- [2] 岩田誠,脳のコミュニケーション,朝倉書店,1987
- [3] 上床弘幸,小林豊,新美康永,“音声の感情表現の分析とモデル化”,電子情報通信学会技術研究報告,92 卷,pp. 65-72,1993
- [4] 中央教育審議会,青少年の意欲を高め心と体の相伴った成長を促す方策について(中間まとめ)”,
http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/06112713.htm,
2006年9月28日
- [5] 内閣府,平成28年版高齢社会白書,
http://www8.cao.go.jp/kourei/whitepaper/w-2016/gaiyou/28pdf_indexg.html
- [6] 宮崎健ら,“高齢者音声を印象づける聴覚的特徴に関する検討”,音講論(秋),pp. 283-286,2008
- [7] 藤原敬記,広重真人,荒木健治,栃内香次,“F0・パワー・ホルマントを用いた合成音声の明瞭制御に関する検討”,通学技法,Vol.103, No.517, pp.217-222,2003
- [8] 粕谷秀樹ら,“高齢者音声の分析と合成”,信学技法,SP87-134, pp.33-39,1987
- [9] 原田大輔,水町光徳,二矢田勝行,“高齢者のめりはりのない声に関する音響的解析”,信学技法,pp.13-18,2010
- [10] 濱崎健太ら,“高齢者音声の「めりはりの無い声」を表す物理量に関する考察”,音講論(春),2010
- [11] Mueller PB, Sweeney RJ and Barbeau LJ, “Acoustic and morphologic study of the senescent voice”, Ear Nose Throat Journal 63, pp. 71-75, 1985.
- [12] 佐藤公則,坂口伸治,栗田茂二郎,平野実,“高齢者喉頭の形態学的研究”,日本喉頭科学会 喉頭 4(2), pp. 84-94, 1992
- [13] 春原政浩,寺田清昭,奥野貴俊,岩倉行志,“聴覚情景分析に基づく難聴者のための音声強調手法”,信学技法,応用音響 110(131), pp. 25-30, 2010
- [14] 小林敬,安啓一,程島奈緒,荒井隆行,進藤美津子,“母音のエネルギー一定常部抑圧による高齢者に対する音節強調の検討”,日本音響学会誌 64(5), pp. 278-289,

2008

- [15] 大島賢一, “重度難聴でも高い明瞭度で聞こえる補聴器 (アナログ方式)”, 電子情報通信学会技術研究報告. WIT, 福祉情報工学 106(57), pp.11-16, 2006
- [16] 安啓一, 菱谷誠人, 荒井隆行, 村原雄二, “デジタル補聴器のための臨界帯域圧縮アルゴリズム”, 電子情報通信学会技術研究報告. WIT, 福祉情報工学 102(419), pp. 41-45, 2002
- [17] 三国一郎, 舘野誠, 巻の康一, 山田一郎, “内部スペクトル上で補正を行なう補聴方式とその評価”, 電子情報通信学会技術研究報告. SP, 音声 93(521), pp.71-78, 1994
- [18] 目片強司, 山田義則, 鈴木良二, 田中豊, 河野淳, 松坂宗太郎, “補聴器への応用を目的としたホルマント強調の一検討”, 電子情報通信学会技術研究報告. SP, 音声 93(155), pp.1-8, 1993
- [19] 関本荘太郎, 齋藤收三, “補聴器のための周波数圧縮音声の明瞭度特性”, 日本音響学会誌 41(12), pp. 841-849, 1985
- [20] 清山信正, 今井篤, 三島剛, 都木徹, 宮坂栄一, “高品質リアルタイム話速変換システムの開発”, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理 J84-D-II(6), pp. 918-926, 2001
- [21] 並木寿枝, 古本友紀, 吉村宏紀, 清水忠昭, 井須尚紀, 菅田一博, “砂時計型ニューラルネットワークの補聴器用雑音除去フィルタへの応用”, 電子情報通信学会技術研究報告. SP, 音声 101(271), pp. 17-23, 2001
- [22] 岩崎紀子, 白石君男, “デジタル補聴器の適応型指向性マイクロホンによる雑音抑制効果”, *Audiology Japan* 48(6), pp. 623-632, 2005
- [23] 井手隆宏, 水町光徳, 中藤良久, “高齢者の音響的特徴と音声認識性能との関係性の検討”, 電子情報通信学会総合大会講演論文集 2012 年情報・システム(1), pp. 195, 2012
- [24] 宮崎健, 毛利年成, 水町光徳, 二矢田勝行, “高齢者音声標準パターンによる高齢者音声認識性能の評価”, 電気関係学会九州支部連合大会講演論文集 2005(0), pp. 414, 2005
- [25] 馬場朗, 芳澤伸一, 山田 実一, 李晃伸, 鹿野 清宏, “高齢者向け音響モデルによる大語彙連続音声認識の評価”, 情報処理学会研究報告音声言語情報処理 (SLP)

- 2001(11(2000-SLP-035)), pp.13-18, 2001
- [26] 長崎ひとみ, 中村美知子, “喉頭全摘出術患者のコミュニケーション状態認識の特徴”, *Yamanashi Nursing Journal* Vol.6 No.2, pp.11-16, 2008
- [27] 伊福部達, “音の福祉工学”, pp.19-24, 165-189, コロナ社, 1997
- [28] 中村圭吾, 戸田智基, 猿渡洋, 鹿野清宏, “肉伝導人工音声の変換に基づく喉頭全摘出者のための音声コミュニケーション支援システム”, *電子情報通信学会論文誌 D* Vol.J90-D No.3 pp.780-787, 2007
- [29] 板倉淳, “人工喉頭音声”, *音声言語医学* 39, pp.443-449, 1998
- [30] 中村圭吾, 戸田智基, 猿渡洋, 鹿野清宏, “外部音源を用いた発声補助システムにおける統計的声質変換の評価”, *電子情報通信学会技術研究報告. WIT, 福祉情報工学* 109(260), pp.49-54, 2009
- [31] 中村圭吾, 戸田智基, 猿渡洋, 鹿野清宏, “喉頭摘出者の会話支援システムにおける微弱な音源信号に関する検討”, *電子情報通信学会技術研究報告. SP, 音声* 107(165), pp.91-96, 2007
- [32] 中村圭吾, 戸田智基, 猿渡洋, 鹿野清宏, “喉頭摘出者を想定して微弱な音源信号を用いて収録された肉伝導音声の声質変換”, *電子情報通信学会技術研究報告. WIT, 福祉情報工学* 106(57), pp.65-70, 2006
- [33] 田中宏, 戸田智基, グラム・ニュービッグ, サクリアニ・サクティ, 中村哲, “雑音抑圧及び統計的音源予測に基づくハイブリッドな電気音声強調法の評価”, *研究報告音楽情報科学 (MUS)* 2014-MUS-103(64), pp.1-6, 2014
- [34] 田中宏, 戸田智基, グラム・ニュービッグ, サクリアニ・サクティ, 中村哲, “統計的音源予測に基づく電気式人工喉頭制御法のシミュレーションによる評価”, *電子情報通信学会技術研究報告. SP, 音声* 114(91), pp.33-38, 2014
- [35] 上三憲弘, 伊福部達, 高橋誠, 松島純一, “ピッチ周波数制御型人工喉頭の提案とその評価”, *電子情報通信学会論文誌. D-II, 情報・システム, II-情報処理* J78-D-2(3), pp.571-578, 1995
- [36] 上見憲弘, 橋場参生, 須貝保徳, 山口悦範, 伊福部達, “抑揚を制御できる電気式人工喉頭の製品化と喉頭摘出者による評価”, *電子情報通信学会技術研究報告. SP, 音声* 98(638), pp.47-52, 1999
- [37] 菊池義信, 粕谷英樹, “電気式人工喉頭の F0 制御に関する検討”, *電子情報通信*

- 学会技術研究報告. SP, 音声 102(417), pp. 65-68, 2002
- [38] 村上浩司, 荒木健治, 広重真人, 枋内香次, “電気発声音声の健常者音声への音声変換手法の性能評価”, 信学論 J87-D-I, pp. 1030-40, 2004
- [39] 森田龍彌, 松村 雅史, “頸部電気インピーダンスを用いた音声ピッチの推定”, ヒューマンインタフェースシンポジウム 2003 論文集, pp.273-280, 2003
- [40] 古井貞熙, 音声情報処理, 森北出版, 1998
- [41] 板橋秀一, 赤羽真, 石川泰, 大河内正明, 粕谷英樹, 桑原尚夫, 田中和世, 知恒雄, 矢頭隆, 渡辺隆夫, 音声工学, pp. 6-12, 森北出版, 2008
- [42] 中田和男, 音声の合成と認識, 総合電子出版, 1980
- [43] Julius O Smith III, “Bark and ERB Bilinear Transforms”, Final draft accepted for publication in the IEEE Transactions on Speech and Audio Processing, pp. 1-32, November, 1999
- [44] 佐藤信, 統計的官能検査法, 日科技連, 1985
- [45] 屏貴文, 田中裕人, 水町光徳, 中藤良久, 松井謙二, “音源波形の違いが電気喉頭音声に与える影響の検討”, 産業応用工学会全国大会 2013, p24-25, 2013
- [46] 古井 貞熙, 音響・音声工学, 近代科学社, 2001
- [47] Dennis H. Klatt and Laura C. Klatt, “Analysis, synthesis, and perception of voice quality variations among female and male talkers”, J. Acoust. Soc. Am., 87, pp. 820-857, 1990
- [48] 上田修功, “統計的学習の基礎”, 電子情報通信学会誌, vol. 85, no. 4, pp. 265-271, 2002.
- [49] S. Young, J. Jansen, J. Odell, D. Ollason, and P. Woodland, The HTK BOOK, 1995

業績一覧

<学術雑誌等に発表した論文>

- [1] Yuto Tanaka, Hiroaki Igaue, Mitsunori Mizumachi, Yoshihisa Nakatoh, “Study of Improvement of Intelligibility for the Elderly Speech Based on Formant Frequency Shift”, International Journal of Computer, Consumer and Control 2015 Vol. 3, No. 3, pp. 57-64, August 2014[査読有]
- [2] Takashi Himeno, Yuto Tanaka, Mitsunori Mizumachi, Yoshihisa Nakatoh, Kenji Matsui, “Singing System by Electric-Larynx Using LPC Residual Wave” Journal of the Institute of Industrial Applications Engineers Vol. 3, No. 4, pp. 183-186, December 25, 2015, DOI: 10.12792/JIIAE.3.183 [査読有]
- [3] 姫野剛至, 田中裕人, 水町光徳, 松井謙二, 中藤良久, “LPC 残差駆動歌唱システムにおける音韻性除去による音質改善の検討”, 産業応用工学会論文誌 Vol. 4, No. 1, pp. 19-25, September 2016[査読有]
- [4] Yuto Tanaka, Mitsunori Mizumachi, Yoshihisa Nakatoh, “Method for improving intelligibility of elderly speech based on formant frequency shift”, Journal of the Institute of Industrial Applications Engineers, (公開予定)

<国際会議における発表>

- [1] Yuto Tanaka, Hiroaki Igaue, Mitsunori Mizumachi, Yoshihisa Nakatoh, “Study of Improvement of Intelligibility for Elderly Speech Based on Formant Frequency Shift”, ICISIP2014, GS4-4, pp. 163-166, NISHINIPPON INSTITUTE OF TECHNOLOGY, September 2014[査読有]
- [2] Yuto Tanaka, Hiroaki Igaue, Mitsunori Mizumachi, Yoshihisa Nakatoh, “STUDY OF IMPROVEMENT OF INTELLIGIBILITY BASED ON FORMANT FREQUENCY SHIFT FOR VOWELS OF ELDERLY SPEECH” Joint conference on Acoustics 2015, pp. 7-10, NAGASAKI UNIVERSITY, January 2015
- [3] Yuto Tanaka, Mizumachi Mitshunori, Yoshihisa Nakatoh, “Improvement of Speech Intelligibility for Elderly People Based on Formant Frequency Shift”, NCSP2016, pp. 343-346, March 2016

- [4] Takumi Misawa, Yuto Tanaka, Yoshihisa Nakatoh, “Study of Text Reading Method Using Important Words on Web Sites for visual impaired” ICISIP2015, pp. 317-320, September 4, 2015
- [5] Yuto Tanaka, Mitsunori Mizumachi, Yoshihisa Nakatoh, “ Study on the Improvement of Intelligibility for Elderly Speech Using Formant Frequency Shift Method” , Speech and Computer 18th International Conference, SPECOM2016, LNAI9811, pp. 683-687
- [6] Takumi Misawa, Yuto Tanaka, Yoshihisa Nakatoh, “Text-reading Support System for the Visually Impaired on Web Sites” NCSP2016, pp. 331-334, march2016
- [7] Ryota Fujiwara, Yuto Tanaka, Yoshihisa Nakatoh, “Study of Word Intelligibility Based on Control of ITD or ILD on Binaural Hearing” , Proceedings of the 3rd IIAE International Conference on Intelligent Systems and Image Processing 2015, pp. 325-328, September 2015
- [8] Ryota Fujiwara, Yuto Tanaka, Yoshihisa Nakatoh, “ Improvement of Word Intelligibility Based on ITD and ILD Control on Binaural Hearing Aids” , NSCP2016, pp. 335-338, march2016
- [9] Teruhisa Higashi, Yuto Tanaka, Yoshihisa Nakatoh, “Study of the Volume Limit for Hearing Protection in the Music Listening” , ICISIP2015, pp. 321-324, September 4, 2015
- [10] Teruhisa Higashi, Yuto Tanaka, Yoshihisa Nakatoh, “Study of Volume Limit for Headphone Hearing Loss Prevention on Portable Music Player” , 2016 IEEE International Conference on Consumer Electronics, pp. 433-434, January 10, 2016
- [11] Takashi Himeno, Yuto Tanaka, Mitsunori Mizumachi, Yoshihisa Nakatoh, Kenji Matsui, “Singing System by Electric-Larynx Using LPC Residual Wave” The 3rd IIAE International Conference on Intelligent Systems and Image Processing 2015, Proceedings of the 3rd IIAE International Conference on Intelligent Systems and Image Processing 2015, pp. 313-316, September 4, 2015 [査読有]
- [12] Takashi Himeno, Yuto Tanaka, Mitsunori Mizumachi, Yoshihisa Nakatoh, Kenji Matsui, “Singing Support System Using LPC Residual Wave for Laryngectomy”

- 2016 IEEE International Conference on Consumer Electronics (ICCE), pp.431-432, January 10, 2016 [査読有]
- [13] Takafumi Hei, Yuto Tanaka, Mitsunori Mizumachi Yoshihisa Nakatoh, Kenji Matsui, “Study of Natural-Voice-Like Vibration Sound for Electrolarynx” , Proc. ICISIP 2014, pp.159-162, Sep, 2014 [査読有]
- [14] Toshiaki Ebihara, Yuto Tanaka, Yoshihisa Nakatoh, “Study on Improvement of Sound Localization for Dichotic-Listening” ,Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.501-504, Septemer 2016
- [15] Shinichiro Asayama, Yuto Tanaka, Yoshihisa Nakatoh, “Study of Listening Training to Improve Auditory Ability for Elderly People” ,Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.505-508, Septemer 2016
- [16] Kazuki Kodama, Yuto Tanaka, Yoshihisa Nakatoh, “Study of Onomatopoeia Expression Caption on TV Contents” ,Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.541-544, Septemer 2016
- [17] Hiroyuki Takamatsu, Yuto Tanaka, Yoshihisa Nakatoh, “Improvement of MAEQ Method for Speech Enhancement in Noisy Environments” , Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.545-548, Septemer 2016
- [18] Kosuke Kuwayama, Yuto Tanaka, Yoshihisa Nakatoh, “Improvement of the Approach Vehicle Detection by the Supersonic Wave” , Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.554-557, Septemer 2016
- [19] Kazuyoshi Sawada, Yuto Tanaka, Yoshihisa Nakatoh, “Study on Evaluation about Unpleasantness of Sounds in Everyday Life by Elderly People” , Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016, pp.558-561, Septemer 2016

<国内会議における発表>

- [1] 田中裕人, 伊賀上祐彰, 水町光徳, 中藤良久, 松井謙二, “ホルマント周波数の補正に基づく高齢者音声の明瞭性改善”, 電子情報通信学会 2014 年ソサイエティ大会 2014, A-4-16, p. 79, 徳島大学, 2014 年 9 月
- [2] 田中裕人, 屏貴文, 水町光徳, 中藤良久, “電気喉頭音声の音質改善に関する予備検討”, 第 66 回電気関係学会九州支部連合大会講演論文集, 01-1A-09, p. 14, 熊本大学, 2013 年 9 月
- [3] 田中裕人, 水町光徳, 中藤良久, “高齢者音声の明瞭性改善に向けたホルマント補正法の検討”, 第 11 回学生のための研究発表会講演論文集, pp. 23-24, 九州工業大学, November 28, 2015
- [4] 田中裕人, 水町光徳, 中藤良久, “ホルマント周波数シフトによる高齢者音声の明瞭度改善”, 音講論 (秋), pp. 1139-1140, 2016
- [5] 澤田知良, 田中裕人, 中藤良久, “認知症高齢者に向けた立体音響によるヒーリング効果の予備検討”, 産業応用工学会全国大会 2015 講演論文集, pp. 30-31, September 4, 2015
- [6] 小玉和貴, 田中裕人, 中藤良久, “TV コンテンツにおける音響情報の字幕化に関する検討”, 産業応用工学会全国大会 2015 講演論文集, pp. 26-27, September 4, 2015
- [7] 小玉和貴, 田中裕人, 中藤良久, “TV 映像におけるオノマトペ字幕付与に関する検討” 2015 年 ソサイエティ大会講演論文集, p. 114, September 9, 2015
- [8] 小玉和貴, 田中裕人, 中藤良久, “TV 映像におけるオノマトペ字幕の有用性の検討” 第 11 回 学生のための研究発表会 講演論文集, p. 41, November 28, 2015
- [9] 桑山晃介, 田中裕人, 中藤良久, “マスクングを用いた周囲騒音による不快感の軽減” 産業応用工学会全国大会 2015 講演論文集, pp. 28-29, September 4, 2015
- [10] 海老原和晃, 田中裕人, 中藤良久, “両耳分離補聴における音声到来方向感の検討” 日本音響学会九州支部第 11 回学生のための研究発表会講演論文集, pp. 39-40, 2015.
- [11] 海老原和晃, 田中裕人, 中藤良久 “両耳分離補聴の分離条件が音像定位に与える影響の検討” 日本音響学会 2016 年 春季研究発表会 講演論文集, pp. 543-544, 2016
- [12] 三澤匠, 田中裕人, 中藤良久, “視覚障害者のための Web 利用における重要単語読み上げ方法の検討”, 2015 年ソサイエティ大会講演論文集, p. 113, September 9, 2015

- [13] 三澤匠, 田中裕人, 中藤良久, “視覚障害者のための重要単語を用いた Web 読み上げ方法の検討”, 第 11 回 学生のための研究発表会講演論文集, pp. 3-4, November 28, 2015
- [14] 高松宏行, 田中裕人, 中藤良久, “雑音環境下での音声補正による明瞭性改善方法の検討” 日本音響学会九州支部 第 11 回学生のための研究発表会 講演論文集, pp. 47-48, November 28, 2015
- [15] 高松宏行, 田中裕人, 中藤良久, “MAEQ 法の改良による拡声音の明瞭性改善効果の検討” 日本音響学会 2016 年春季研究発表会 講演論文集, pp. 837-838, 2016
- [16] 藤原良太, 田中裕人, 中藤良久, “両耳受聴における時間差・音圧差が単語理解度に与える影響の検討” 日本音響学会 2015 年 秋季研究発表会 講演論文集, pp. 445-446, September 18, 2015
- [17] 藤原良太, 田中裕人, 中藤良久, “両耳受聴における時間差・音圧差の違いが単語理解度に与える影響の検討” 第 11 回 学生のための研究発表会 講演論文集, pp. 37-38, November 28, 2015
- [18] 姫野剛至, 田中裕人, 水町光徳, 中藤良久, 松井謙二, “パワー補正した LPC 残差を用いる代替発声歌唱法の音質改善” 平成 27 年度(第 68 回) 電気・情報関係学会九州支部連合大会 講演論文集, p. 55, September 26, 2015
- [19] 伊賀上祐彰, 田中裕人, 水町光徳, 中藤良久, “高齢者音声の音響的特徴に基づいた明瞭性改善方法の検討” 産業応用工学会全国大会 2014 講演論文集, pp. 20-21, September 28, 2014
- [20] 屏貴文, 田中裕人, 水町光徳, 中藤良久, 松井謙二, “音源波形の違いが電気喉頭音声に与える影響の検討”, 産業応用工学会全国大会 2013, pp. 24-25, 2013
- [21] 姫野剛至, 屏貴文, 田中裕人, 水町光徳, 中藤良久, 松井謙二, “電気式人工喉頭を用いた歌声の音質改善” 産業応用工学会全国大会 2014 講演論文集, pp. 16-17, September 28, 2014
- [22] 姫野剛至, 田中裕人, 水町光徳, 中藤良久, 松井謙二, “パワー補正した LPC 残差を用いる代替発声歌唱法の音質改良” 第 11 回 学生のための研究発表会 講演論文集 pp. 43-44, November 28, 2015
- [23] 浅山慎一郎, 田中裕人, 中藤良久, “高齢者における聴能訓練による音声聞き取り改善効果の検討” 日本音響学会 2015 年秋季研究発表会講演論文集 pp. 1247-1248,

September 16, 2015

- [24] 浅山慎一郎, 田中裕人, 中藤良久, “聴能訓練が高齢者の語音明瞭度に与える影響の検討” 電気・情報通信学会九州支部第 68 回連合大会, p. 54, September 26, 2015
- [25] 浅山慎一郎, 田中裕人, 中藤良久, “高齢者に対する聴能訓練による語音明瞭度改善効果の検討” 日本音響学会九州支部第 11 回学生のための研究発表会講演論文集 pp. 45-46, November 28, 2015
- [26] 東照久, 田中裕人, 中藤良久, “音楽聴取における聴力保護のための音量制限の検討” 日本音響学会 2015 年秋季研究発表会 講演論文集, pp. 1249-1250, September 16, 2015
- [27] 東照久, 田中裕人, 中藤良久, “音楽聴取における聴力損失防止のための音量制限の検討” 第 11 回学生のための研究発表会 講演論文集, pp. 121-122, November 28, 2015
- [28] 野上博高, 田中裕人, 中藤良久, 松井謙二, “喉頭摘出者に向けた歌唱システムにおける音質改善の検討”, 日本音響学会 2016 年秋季研究発表会 講演論文集, pp. 1133-1134, September 15, 2016
- [29] 高松宏行, 田中裕人, 中藤良久 “騒音下での MAEQ とダイナミックレンジ圧縮を用いた拡声音の明瞭性改善” 日本音響学会 2016 年秋季研究発表会 講演論文集, pp. 1287-1288, September 15, 2016
- [30] 澤田知良, 田中裕人, 中藤良久, “高齢者における日常生活音の不快感についての評価” 日本音響学会 2016 年秋季研究発表会 講演論文集, pp. 1117-1118, September 15, 2016
- [31] 鳥越大貴, 田中裕人, 中藤良久, “視覚障がい者のためのスマートフォン操作支援の検討”, 電気情報通信学会 ソサイティ大会 2016 講演論文集, A-15-12, p. 8 北海道大学, 2016 年 9 月

<国際シンポジウム>

- [1] Yuto Tanaka, Hiroaki Igaue, “Improvement of intelligibility for elderly speech based on formant frequency shift”, Association of Computing Machinery in California State University Monterey Bay, September 4, 2014

<国内シンポジウム>

- [1] 田中裕人, “ホルマント周波数シフトによる音声の明瞭度改善に関する研究”, 第14回聴覚障害とその補償に関する情報交換会, 2016年9月24日

<受賞>

学生優秀賞受賞, 日本音響学会九州支部, March 28, 2015

付録

付録1 音韻バランスのとれた 543 単語

単語番号	ローマ字表記	仮名表記	漢字表記
1	AOAO	アオアオ	青青
2	EIRI	エイリ	営利
3	EGGU	エッグ	
4	IIKI	イイキ	いい気
5	KENNO	ケンオ	嫌悪
6	NANNA	ナンア	南ア
7	SEOU	セオウ	背負う
8	SUAE	スアエ	酢和え
9	UREU	ウレウ	憂う
10	ANNGYA	アンギャ	行脚
11	ADJU-	アデュー	
12	AYAUI	アヤウイ	危うい
13	BUNNUNN	ブンウン	文運
14	INNIONN	インイオン	陰イオン
15	ONNUCHI	オンウチ	御内
16	SOFA-	ソファー	
17	URUOU	ウルオウ	潤う
18	BYUFFE	ビュッフエ	
19	EURUSHI	エウルシ	絵漆
20	GOBYU-	ゴビュウ	誤謬
21	KARIIO	カリイオ	仮庵
22	KIEIRU	キエイル	消え入る
23	KOROOI	コロオイ	頃

24	MAEASHI	マエアシ	前足
25	MAEOKI	マエオキ	前置き
26	MISOAE	ミソアエ	味噌和え
27	MONOUI	モノウイ	物憂い
28	NAMAANN	ナマアン	生餡
29	SAFAIA	サファイア	
30	SEHYO-	セヒョウ	世評
31	SHIOOSHI	シオオシ	塩押し
32	SHIRAAE	シラアエ	白和え
33	SOEUMA	ソエウマ	副馬
34	SYANTSE	シャンツェ	
35	TOGYU-	トギユウ	斗牛
36	WAMYO-	ワミョウ	倭名
37	AOENNDO-	アオエンドウ	青豌豆
38	ATASSHU	アタッシュ	
39	TSETSEBAE	ツェツェバエ	
40	DJUETTO	デュエット	
41	ECHU-DO	エチュード	
42	GERYAKU	ゲリャク	下略
43	PIPETTO	ピペット	
44	GOMYAKU	ゴミャク	語脈
45	KADENNTSA	カデンツァ	
46	KAGYAKU	カギャク	可逆
47	SUMOGGU	スモッグ	
48	KARYAKU	カリャク	下略 or 嘉暦
49	KENNGYU-	ケンギユウ	牽牛
50	KEHHERU	ケッヘル	

51	KOHERU	コッヘル	
52	UIUISHI-	ウイウイシイ	初初しい
53	BACHERA-	バチェラー	
54	BONNPYAKU	ボンピャク	凡百
55	DEFORUME	デフォルメ	
56	INNTABYU-	インタビュー	
57	KAIBYAKU	カイビャク	開白 or 開關
58	KANNTSO-NE	カンツォーネ	
59	KIEUSERU	キエウセル	消え失せる
60	KONNMJU-NN	コンミュン	
61	KONYAKKU	コニャック	
62	MOEAGARU	モエアガル	燃え上がる
63	POPYURA-	ポピュラー	
64	SUKERUTSO	スケルツォ	
65	NIEAGARU	ニエアガル	煮え上がる
66	SUKYABBU	スキャップ	
67	ZINNFI-ZU	ジンフィーズ	
68	KOKUBYAKU	コクビャク	黒白
69	BA-BEKYU-	バーベキュー	
70	MEKYABETSU	メキャベツ	芽キャベツ
71	KONNCHEUTO	コンチェルト	
72	MENNSYEBIKI	メンシェビキ	
73	NYORONYORO	ニョロニョロ	にょろにょろ
74	KONNPYU-TA	コンピュータ	
75	PYOKOPYOKO	ピョコピョコ	ぴよこぴよこ
76	SHATTOAUTO	シャットアウト	
77	SYOGENNHYO-	ショゲンヒョウ	諸元表

78	YANYO-SYO-	ヤニョウショウ	夜尿症
79	BORUSYEBIKI	ボルシェビキ	
80	PURODJU-SA-	プロデューサー	
81	SHIMYURE-TA	シミュレータ	
82	TEREFAKKUSU	テレファックス	
83	NIHYAKUTO-KA	ニハクトオカ	二百十日
84	AMYU-ZUMENNTO	アミューズメント	
85	ICHIZENNMESHIYA	イチゼンメシヤ	一膳飯屋
86	GIHUTOCHEKKU	ギフトチェック	
87	NEMIDAREGAMI	ネミダレガミ	寝乱れ髪
88	OPOCHUNIZUMU	オポチュニズム	
89	AIMOKAWARANU	アイモカワラヌ	相も変らぬ
90	HITOTSUBUDANE	ヒトツブダネ	一粒種
91	KUYASHIMAGIRE	クヤシマギレ	悔し紛れ
92	TSUTAMOMIJI	ツタモミジ	蔦紅葉
93	HATSUHIKAGE	ハツヒカゲ	初日影
94	HUKANASAKE	フカナサケ	深情け
95	NORIKUMIINN	ノリクミイン	乗組員
96	ZEMUKURIPPU	ゼムクリップ	
97	BINNGOOMOTE	ビンゴオモテ	備後表
98	DANNSONNJOHI	ダンソンジョヒ	男尊女卑
99	TEKIHONNSHUGI	テキホンシュギ	敵本主義
100	ORESENGURAHU	オレセングラフ	折れ線グラフ
101	AINARUBEKUWA	アイナルベクワ	

102	WASURENAGUSA	ワスレナグサ	勿忘草
103	HONECHIGAI	ホネチガイ	骨違い
104	TABINOSORA	タビノソラ	旅の空
105	BENISASHIYUBI	ベニサシユビ	紅差指
106	UNNMAKASE	ウンマカセ	運任せ
107	IMIKOTABA	イミコトバ	忌詞・忌言葉
108	OTAZUNEMONO	オタズネモノ	御尋ね者
109	RENNKINNJUTSU	レンキンジュツ	錬金術
110	RANNHANNSHA	ランハンシャ	乱反射
111	MUNAZANNYO-	ムナザンヨウ	胸算用
112	INNHURUENNA	インフルエンザ	
113	RUIGIGO	ルイギゴ	類義語
114	KYAPUTENN	キャプテン	
115	PEDHIKYUA	ペディキュア	
116	DABOHAZE	ダボハゼ	だぼ鯨
117	ENNSADAME	エンサダメ	縁定め
118	TOMESODE	トメソデ	留袖
119	TSUJIBANNSHO	ツジバンショ	辻番所
120	JOSONNDANNPI	ジョソندانピ	女尊男卑
121	YONAYONA	ヨナヨナ	夜な夜な
122	SAIKEDERIKKU	サイケデリック	
123	GURAMUHANNNO-	グラムハンノウ	グラム反応
124	PAGODA	パゴダ	
125	CHANOYU	チャノユ	茶の湯
126	DHIBEROPPA-	ディベロッパ	
127	SENNBONSHIMEJI	センボンシメジ	千本占地
128	NISSHINNGEPPU	ニッシンゲッポ	日進月歩

129	CHINNPANNJI-	チンパンジー	
130	ZEGAHIDEMO	ゼガヒデモ	是が非でも
131	JUNNKANNGOHU	ジュンカンゴフ	准看護婦
132	ISEZAKIORI	イセザキオリ	伊勢崎織
133	UDETATEHUSE	ウデタテフセ	腕立て伏せ
134	SUIAGEPONNPU	スイアゲポンプ	吸上げポンプ
135	DOSANNKO	ドサンコ	道産子
136	HODOHETE	ホドヘテ	
137	CHOBIHIGE	チョビヒゲ	ちょび髭
138	RIPOIDO	リポイド	
139	MUZUGAYUI	ムズガユイ	むず痒い
140	SEMEDAINN	セメダイン	
141	OCHIKUBOMU	オチクボム	
142	RONNKYO	ロンキョ	論拠
143	JESUCHA-	ジェスチャー	
144	GYOKUHO	ギョクホ	玉歩
145	SUSOYOKE	スソヨケ	裾除け
146	NUNOGIRE	ヌノギレ	布切れ
147	PICHAPICHA	ピチャピチャ	ぴちゃぴちゃ
148	HARUTSUGEUO	ハルツゲウオ	春告魚
149	ZOKUJYU	ゾクジュ	俗儒
150	TETORAPOTTO	テトラポット	
151	IENEZUMI	イエネズミ	家鼠
152	HAHAUE	ハハウエ	母上

153	NEGURIJE	ネグリジェ	
154	ADOBENNCHA-	アドベンチャー	
155	OITEKIBORI	オイテキボリ	置いてきぼり
156	ZAI I	ザイイ	在位
157	GINNHONNI	ギンホンイ	銀本位
158	CHENNBARO	チェンバロ	
159	KONNPOJISHONN	コンポジション	
160	UONNBINN	ウオンビン	ウ音便
161	HENNENNKE-	ヘンエンケイ	扁円形
162	FIANNSE	フィアンセ	
163	GYAKUYUNYU-	ギャクユニユウ	逆輸入
164	RYAKUJO	リyakジョ	略叙
165	WAI RUDOPICCHI	ワイルドピッチ	
166	YABUICHIKUANN	ヤブイチクアン	藪井竹庵
167	KUCHIYOGOSHI	クチヨゴシ	口汚し
168	BISHAMONNTENN	ビシャモンテン	毘沙門天
169	NEKONADEGOE	ネコナデゴエ	猫撫で声
170	HOHOEMASHI-	ホホエマシイ	微笑ましい
171	ZUNNBERABO-	ズンベラボウ	ずんべら坊
172	HUSHOZONN	フショゾン	不所存
173	MANNNENNRE-	マンネンレイ	満年齢
174	TEYANNDE-	テヤンデイ	
175	MAHOGANI-	マホガニー	
176	SHIGERIAU	シゲリアウ	繁り合う
177	YAMASOWA	ヤマソワ	山岨
178	NAISHUKKETSU	ナイシュクケツ	内出血
179	KINUKOMACHI	キノコマチ	絹小町

180	GUNNRYO	グンリョ	
181	BUAIKYO-	ブアイキョウ	無愛敬
182	CHU-ONNHUGO-	チュウオンフゴウ	注音符号
183	GORIMUCHU-	ゴリムチュウ	五里霧中
184	PUROYAKYU-	プロヤキュウ	プロ野球
185	RINNGYO	リンギョ	臨御
186	HITZUKIAI	ヒトヅキアイ	人付き合い
187	JIGUZAGU	ジグザグ	
188	OCCHOKOCHOI	オッチョコチョイ	おっちょこち よい
189	YAKEBOKKUI	ヤケボックイ	焼け棒杭
190	MUNYAMUNYA	ムニヤムニヤ	むにやむにや
191	SHOBOSHOGO	ショボショボ	しょぼしょぼ
192	MOCHIE	モチイエ	持ち家
193	UESAMA	ウエサマ	上様
194	TAICHO	タイチョ	大著
195	KOMYUNITHI-	コミュニティー	
196	TENUGUI	テヌグイ	手拭
197	ISSATSUTASHO-	イッサツタショウ	一殺多生
198	ENNRIEDO	エンリエド	厭離穢土
199	KYOKUSHA	キョクシャ	
200	UJAUJA	ウジャウジャ	うじゃうじゃ
201	SUFINKUSU	スフィンクス	
202	MEZOSOPURANO	メゾソプラノ	

203	SHERUPA	シェルパ	
204	DEKASEGI	デカセギ	出稼ぎ
205	SANNWAONN	サンワオン	三和音
206	ANNPEA	アンペア	
207	OAIISO	オアイソ	御愛想
208	NYU-TAUNN	ニュータウン	
209	THISSHUPE-PA-	ティッシュペーパー	
210	RYOGU-	リョグウ	旅寓
211	KOZEWASHI-	コゼワシイ	小忙しい
212	AGEHACHO-	アゲハチョウ	揚羽蝶
213	ZENNNYO	ゼンニョ	善女
214	FO-KUROA	フォークロア	
215	KYU-KANNCHO-	キュウカンチョウ	九官鳥
216	NANNTOIU	ナントイウ	何と言う
217	UINNCHI	ウインチ	
218	NESSYABYO-	ネッシャビョウ	熱射病
219	KAFEINN	カフェイン	
220	HANNBUSSHITSU	ハンブッシツ	反物質
221	GADENNINNSUI	ガデンインスイ	我田引水
222	SHUSSEKENN	シュッセケン	出世間
223	GOSHUKYO-GI	ゴシュキョウギ	五種競技
224	ESHAJO-RI	エシャジョウリ	会者定離
225	MIDHI	ミディ	
226	RYU-KIHE-	リュウキヘイ	竜騎兵
227	EAZO-RU	エアゾール	
228	KUNNONN	クンオン	君恩

229	SUPO-THI	スポーティー	
230	AENKA	アエンカ	亜鉛華
231	GENNANN	ゲンアン	原案
232	OIOI	オイオイ	追い追い
233	RINNU	リンウ	霖雨
234	HYUTTE	ヒュッテ	
235	RAIUNN	ライウン	雷雲
236	PUROFESSA-	プロフェッサー	
237	HYAKUBUNNHI	ヒャクブンヒ	百分比
238	BAIENN	バイエン	煤煙
239	JUNNAI	ジュンアイ	純愛
240	MASSAICHU-	マッサイチュウ	真っ最中
241	ARUFA	アルファ	
242	DAGGUAUTO	ダッグアウト	
243	UYOKYOKUSETSU	ウジョキョクセツ	紆余曲折
244	APACCHI	アパッチ	
245	SENNCHU-HA	センチュウハ	戦中派
246	BIBIDDO	ビビッド	
247	KANZEONN	カンゼオン	観世音
248	SUPIRITTO	スピリット	
249	APENNDHIKKUSU	アペンディックス	
250	SUPPOMUKERU	スップヌケル	すっぽ抜ける
251	MYAKUMYAKU	ミヤクミヤク	脈脈
252	HIRADAIPIANO	ヒラダイピアノ	平台ピアノ

253	YUNNZUE	ユンヅエ	弓杖
254	PARIJENNU	パリジェンヌ	
255	FUTEBUTESHI-	フテブテシイ	ふてぶてしい
256	GESYUTAPO	ゲシュタポ	
257	ROKUNUSUBITO	ロクヌスビト	禄盗人
258	MITSUZOROI	ミツゾロイ	三つ揃い
259	SHISHAGONYU-	シシャゴニュー	四捨五入
260	OERAGATA	オエラガタ	御偉方
261	BOTSUSHUMI	ボツシュミ	没趣味
262	JOJONI	ジョジョニ	徐々に
263	JAINAKYO-	ジャイナキョウ	ジャイナ教
264	KITAHANNKYU-	キタハンキュウ	北半球
265	GANNMODOKI	ガンモドキ	雁擬き
266	FIGYUA	フィギュア	
267	HAIKATSURYO-	ハイカツリョウ	肺活量
268	IUNAREBA	イウナレバ	言うなれば
269	YASUKEAI	ヤスウケアイ	安請合い
270	DOZAEMONN	ドザエモン	土左衛門
271	ZUIJU-	ズイジュウ	随従
272	TOTOKARUCHO	トトカルチョ	
273	MECHAMECHA	メチャメチャ	めちゃめちゃ
274	YUKIJORO-	ユキジョロウ	雪女郎
275	SHOICHINENN	ショイチネン	初一念
276	HESONOO	ヘソノオ	臍の緒
277	TSUIZO	ツイゾ	終ぞ
278	SHUKUU	シュクウ	宿雨
279	DONNTSU-	ドンツウ	鈍痛

280	SOREYUE	ソレユエ	其故
281	JAJAUMA	ジャジャウマ	じゃじゃ馬
282	HUNNBYO-	フンビョウ	分秒
283	YURIAJUSHI	ユリアジュシ	ユリア樹脂
284	MOJAMOJA	モジャモジャ	もじゃもじゃ
285	BUTSUMYO-	ブツミョウ	仏名
286	HYOIHYOI	ヒョイヒョイ	ひょいひょい
287	MINNYU-	ミンユウ	民有
288	WANNRYUU	ワンリュウ	湾流
289	IATONI-	イアトニー	胃アトニー
290	BIDEO	ビデオ	
291	MERODHI-	メロディー	
292	HOKEKYO-	ホケキョウ	法華経
293	NEPUCHU-NN	ネプチューン	
294	KABEHITOE	カベヒトエ	壁一重
295	GEKIRYO	ゲキリョ	逆旅
296	YAMUOEZU	ヤムヲエズ	止むを得ず
297	HINNPYO-	ヒンピョウ	品評
298	AOUNABARA	アオウナバラ	青海原
299	BOCHABOCHA	ボチャボチャ	ぼちゃぼちゃ
300	CHAKACHAKA	チャカチャカ	ちゃかちゃか
301	SOTSUGYO-	ソツギョウ	卒業
302	METASEKOIA	メタセコイア	
303	DHIFENNSU	ディフェンス	

304	SUTEZERIFU	ステゼリフ	捨て台詞
305	KECHONNKECHONN	ケチョンケチョン	けちよんけちよん
306	HIESHO-	ヒエショウ	冷え性
307	CHICHIUE	チチウエ	父上
308	TSUREAU	ツレアウ	連れ合う
309	INNDEANN	インディアン	
310	KYOJU	キョジュ	巨儒 or 巨樹
311	HONNGIMARI	ホンギマリ	本決り
312	JIPUSHI-	ジプシー	
313	YOMEJO	ヨメジョ	嫁女
314	NYOYASHA	ニョヤシャ	女夜叉 or 如夜叉
315	TO-NETSUBYO-	トウネツビョウ	稲熟病
316	MIDORUKYU-	ミドルキュウ	ミドル級
317	SOKOUO	ソコウオ	底魚
318	JUNNJUNNI	ジュンジュンニ	順順に
319	KYADHII	キャディー	
320	GYOGYO-	ギョギョウ	漁業
321	ZO-HIBYO-	ゾウヒビョウ	象皮病
322	KAKEJAYA	カケジャヤ	掛茶屋
323	REKUIEMU	レクイエム	
324	ZACHU-	ザチュウ	座中
325	HAMEE	ハメエ	嵌め絵
326	SHIKASHU-	シカシュウ	
327	KAMIGYO-	カミギョウ	上京
328	PAIONIA	パイオニア	

329	ROHYO-	ロヒョウ	路標
330	DEHO	デホ	出穂
331	CHINMYO-	チンミョウ	珍妙
332	NIKUGYU-	ニクギユウ	肉牛
333	DASHIJAKO	ダシジャコ	出し雑魚
334	REGYURA-	レギュラー	
335	GORYO-NINN	ゴリョウニン	御寮人
336	FANNFA-RE	ファンファーレ	
337	YOBYO-	ヨビョウ	余病
338	HOSHIUO	ホシウオ	乾魚
339	NUREENN	ヌレエン	濡れ縁
340	ZUHYO-	ズヒョウ	図表
341	CHARYO-	チャリョウ	茶寮
342	KAUIRO	カオイロ	顔色
343	HUNYOI	フニョイ	不如意
344	TAIRYU-	タイリュウ	対流
345	DABURYU-	ダブリュー	
346	NYOKINYOKI	ニョキニョキ	によきによき
347	GYORYU-	ギョリュウ	
348	RANNJERI-	ランジェリー	
349	SHAGYO-	シャギョウ	社業
350	JANNRU	ジャンル	
351	DEBYU-	デビュー	
352	RINYU-	リニュウ	離乳

353	KINNRO-	キンロウ	勤労
354	ITOGONNNYAKU	イトゴンニャク	糸蒟蒻
355	OURASAMA	オウラサマ	御裏様
356	PACHIPACHI	パチパチ	ぱちぱち
357	ZAHYO-	ザヒョウ	座標
358	SAIFONN	サイフォン	
359	FURUIOTOSU	フルイオトス	振り落す
360	SYEPA-DO	シェパード	
361	MIHAPPYO-	ミハッピーウ	未発表
362	NOSSHINOSHI	ノッシノッシ	のっしのっし
363	TONNE-	トンエイ	屯営
364	OBUJE	オブジェ	
365	KEBYO-	ケビョウ	仮病
366	WAGYU-	ワギュウ	和牛
367	UOICHIBA	ウオイチバ	魚市場
368	KUIIJI	クイイジ	食い意地
369	SUEOSOROSHI-	スエオソロシイ	未恐ろしい
370	BOKYABURARI-	ボキャブラリー	
371	SEGYO-	セギョウ	
372	POINNTA-	ポインター	
373	HADAAI	ハダアイ	肌合い
374	BIMYO-	ビミョウ	微妙
375	MENYU-	メニュー	
376	SERYO-	セリョウ	
377	FIROSOFI-	フィロソフィー	
378	JU-MANNOKUDO	ジュウマンオクド	十万億土
379	PYU-RITANN	ピューリタン	

380	GUNNI	グンイ	軍医
381	SHIOAJI	シオアジ	塩味
382	UNNE-	ウンエイ	運営
383	BYO-BO	ビョウボ	廟謨
384	CHOKKYU-	チョッキュウ	直球
385	MYO-GA	ミョウガ	茗荷
386	GYAKKYO-	ギャッキョウ	逆境
387	ANENYO-BO-	アネニョウボウ	姉女房
388	BUUNN	ブウン	武運
389	GYU-NABE	ギユウナベ	牛鍋
390	KAEUTA	カエウタ	替え歌
391	NIOU	ニオウ	臭う
392	HANYU-	ハンユウ	
393	ONNANNPO-	オンアンポウ	温罨法
394	REFERI-	レフェリー	
395	CHINNAGE	チンアゲ	賃上げ
396	KOONNNA	コオンナ	小女
397	PITSA	ピッツァ	
398	BEKKYO	ベッキョ	別居
399	IO-BYO-	イオウビョウ	萎黄病
400	SHINNFONI-	シンフォニー	
401	DONNRANN	ドンラン	
402	KINNENSHU	キンエンシュ	近縁種
403	HANNIGO	ハンイゴ	反意語

404	TEUE	テウエ	
405	NAMAUO	ナマウオ	生魚
406	ANNE-	アンエイ	
407	MYU-ZU	ミューズ	
408	MAEIWAI	マエイワイ	前祝
409	DHISUPURE-	ディスプレイ	
410	BYAKUGO-	ビャクゴウ	
411	SUNOBBU	スノッブ	
412	MINOUE	ミノウエ	身の上
413	KUKKYO-	クッキョウ	屈強
414	NEOI	ネオイ	根生い
415	BEPPYO-	ベッピョウ	別表
416	SHUCCHO-	シュッチョウ	出張
417	BUTTSUZUKE	ブツツヅケ	打っ続け
418	TOSSA	トッサ	咄嗟
419	GETTSU-	ゲッター	
420	RIKKYO-	リッキョウ	陸橋
421	SEKIAEZU	セキアエズ	
422	JAJJI	ジャッジ	
423	HOANN	ホアン	保安
424	PURJOKUTA-	プロジェクター	
425	INNO-GO	インオウゴ	印欧語
426	RODDO	ロッド	
427	IPPYO-	イッピョウ	一瓢
428	GUDDOBAI	グッドバイ	
429	RINNGAFONN	リングフォン	
430	TERYU-DANN	シュリュウダン	手榴弾

431	ROJJI	ロッジ	
432	HOPPYO-YO-	ホッピーヨウヨウ	北氷洋
433	BAKYAKU	バキヤク	馬脚
434	HANNDHIKYAPPU	ハンディキャップ	
435	BEDDO	ベッド	
436	GENNUNN	ゲンウン	
437	EJJI	エッジ	
438	OHYARAKASU	オヒヤラカス	おひゃらかす
439	HINNKYAKU	ヒンキヤク	賓客
440	BURIJJI	ブリッジ	
441	GOGYAKU	ゴギヤク	五逆
442	YUMENOYO	ユメノヨ	夢の世
443	DATSURYAKU	ダツリヤク	
444	TSUMEYOSERU	ツメヨセル	詰め寄せる
445	FESUTHIBARU	フェスティバル	
446	KANNMINNZOKU	カンミンゾク	漢民族
447	RANMYAKU	ランミヤク	乱脈
448	WAKEHEDATE	ワケヘダテ	分け隔て
449	PARAFINN	パラフィン	
450	MIZUBASHO-	ミズバショウ	水芭蕉
451	CHIRYAKU	チリヤク	
452	YOIZAME	ヨイザメ	酔い醒め
453	DOREMIFA	ドレミファ	
454	TAIGYAKU	タイギヤク	大逆

455	HENAHENA	ヘナヘナ	へなへな
456	TENNYAWANNYA	テンヤワンヤ	てんやわんや
457	MO-RUSUHUGO-	モールスフゴウ	モールス符号
458	ZENNRYAKU	ゼンリヤク	前略
459	BIRUDHINNGU	ビルディング	
460	CHI JIREGE	チヂレゲ	縮れ毛
461	KYATTSUAI	キャッツアイ	
462	SAGEJU-	サゲジュウ	提げ重
463	ANNNYUI	アンニュイ	
464	YABUHEBI	ヤブヘビ	藪蛇
465	BARAETHI-	バラエティー	
466	JOKYU-	ジョキュウ	女給
467	RU-PE	ルーペ	
468	ENNJA	エンジャ	縁者
469	NEGATHIBU	ネガティブ	
470	GARYU-	ガリュウ	我流
471	TOPA-ZU	トパーズ	
472	CHOSHA	チョシャ	著者
473	MUNNMUNN	ムンムン	むんむん
474	RIKECCHIA	リケッチア	
475	SEPIA	セピア	
476	ANNUTSU	アンウツ	暗鬱
477	MAHHA	マッハ	
478	OTTSUKE	オツツケ	
479	SAJESUTO	サジェスト	
480	SUKUEADANNSU	スクエアダンス	

481	BIGGU	ビッグ	
482	KOBYAKUSHO-	コビャクショウ	小百姓
483	KIMYAKU	キミャク	気脈
484	NOTAREZINI	ノタレジニ	野垂れ死に
485	PEKOPEKO	ペコペコ	ぺこぺこ
486	SYUZE-	シュゼイ	
487	HURONNTHIA	フロンティア	
488	BONNODORI	ボンオドリ	盆踊
489	AKUGYAKU	アクギャク	悪逆
490	AUTOBA-NN	アウトバーン	
491	SORYAKU	ソリャク	粗略
492	EREBE-TA-	エレベーター	
493	KONNFYUUYONN	コンフュージョン	
494	FYUUCHAA	フューチャー	
495	SUTAA^OOZU	スターウォーズ	
496	SUWINNGU	スウィング	
497	KUWEETO	クウェート	
498	HOWEERU	ホエール	
499	MAIWEI	マイウェイ	
500	^OTTI	ウォッチ	
501	WISUKII	ウィスキー	
502	WETTO	ウェット	
503	DHUU	ドゥー	
504	KYADHII	キャディー	
505	THUU	トゥー	

506	VAIORINN	バイオリン	
507	KAVAA	カバー	
508	VORYUUMU	ボリューム	
509	REVORYUUSYONN	レボリューション	
510	SENNSHITHIVU	センシティブ	
511	VIKUTORII	ビクトリー	
512	NAVIGEESHYONN	ナビゲーション	
513	VEERU	ベール	
514	NOVERU	ノベル	
515	MUUVII	ムービー	
516	RANNVANN	ランバン	
517	PYAPYUPYO	ピャピュピョ	ぴゃぴゅぴょ
518	NYAROME	ニャロメ	にゃろめ
519	DODAI	ドダイ	土台
520	DOOMYOOJI	ドウミョウジ	道明寺
521	OONUMA	オオヌマ	大沼
522	RYUUKOOKA	リュウコウカ	流行歌
523	IMEEZI	イメージ	
524	IAINUKI	イアイヌキ	居合抜
525	HONNRYUU	ホンリュウ	本流 or 奔流
526	HYAHYUHYO	ヒャヒュヒョ	ひゃひゅひょ
527	KUKKUDHUU	クックドウ	
528	KARUCHE	カルチェ	
529	SEFIIRO	セフィーロ	
530	ANNFEA	アンフェア	
531	SANNBYAKU	サンビャク	三百

532	WETTHI-	ウェッティ	
533	UDDHI-	ウッディ	
534	YU-THIRITHI-	ユーティリティー	
535	PUTTSURI	プツリ	ぷっつり
536	TSUKUTSUKUBO-SHI	ツクツクボウシ	つくつく法師
537	PIKKUAPPU	ピックアップ	
538	SHUCCHO-	シュッパツ	出発
539	KISHO-KACHI	キシウカチ	希少価値
540	HUHITSUYO-	フヒツヨウ	不必要
541	SHIKKYAKU	シッキャク	失脚
542	TEPPITSUBANN	テップツバン	鉄筆版
543	KUROKKASU	クロッカス	