

平成 26 年 6 月 13 日現在

機関番号：17104

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23560459

研究課題名(和文)高齢者音声を明瞭化する音声処理インタフェースに関する研究

研究課題名(英文) Research on speech processing interface to improve speech intelligibility for elderly people

研究代表者

中藤 良久 (NAKATOH, YOSHIHISA)

九州工業大学・工学(系)研究科(研究院)・教授

研究者番号：10599955

交付決定額(研究期間全体)：(直接経費) 4,100,000円、(間接経費) 1,230,000円

研究成果の概要(和文)：本研究の目的は、高齢者音声の音響的特徴について解析を行い、その音響的特徴を用いて明瞭性を改善することである。まず成人と高齢者音声の特徴として遷移量差を算出し、母音ごとに比較した。その結果、口や舌を大きく動かす/a/や/i/において、遷移量差の絶対値が大きくなることがわかった。次に、F1-F2平面上における5母音の分布を高齢者と成人で比較した。その結果、高齢者音声では分布が小さくなることがわかった。そこで、フォルマント周波数を補正する方法を提案し、明瞭性の改善効果を検証した。聴取実験の結果、フォルマント補正によって口や舌を大きく動かす必要のある単語音声で明瞭性が改善できた。

研究成果の概要(英文)：The purpose of this study is to analyze for acoustic features of elderly speech, to improve the intelligibility using the acoustic features. First, we were compared with the transition difference for each vowel in adult and elderly. As a result, we found that the absolute value of the transition difference increases on /a/ and /i/ that need to move mouth and tongue largely. Secondly, we were compared with the distribution of the five vowels in the F1-F2 plane in adult and elderly. As a result, we found that the distribution is reduced in the elderly speech. Finally, we propose a method to improve the speech intelligibility by the formant frequency shift. We tested the effect of improving the intelligibility. A result of listening experiments, the intelligibility could be improved on word that need to move mouth and tongue largely.

研究分野：工学

科研費の分科・細目：電気電子工学 通信・ネットワーク工学

キーワード：高齢者支援 音声明瞭化 音声データベース インタフェース

1. 研究開始当初の背景

世界的に高齢化が進んでおり、日本においても高齢化は大きな社会問題の1つである。日本の総人口に占める65歳以上の割合は現在23.1%であり、2055年には40.5%に達する。2.5人に1人が65歳以上となる計算である。その一方で、普段の生活を豊かにしたい、あるいは介護福祉現場や接客業など働きたいと考える高齢者は確実に増えている。そのため、コミュニケーションをとる機会がこれまで以上に増えると考えられる。しかし、高齢者は加齢により声に変質するため、そのことが会話の障害になることは少なくない。

2. 研究の目的

先行研究で、高齢者音声の聴感的印象の一つとして、「めりはりのなさ」があることがわかっている。また、高齢者の声に「めりはり」がなくなると、隣接する音素間のスペクトルの変化量である「遷移量」が低下することがわかっている。しかし、高齢者音声の明瞭性の改善を行うためには、より詳細な高齢者音声の音響的特徴の解析が必要となる。そこで、本研究では、遷移量に対応する音響的特徴を検討し、その音響的特徴を用いて高齢者音声の明瞭性の改善を行う。

3. 研究の方法

先行研究で、高齢者の声にめりはりがなくなると、遷移量が低下することが分かっているが、本研究では、その遷移量の低下の度合いが母音ごとに異なることに着目した。そこで、聴取実験によって選定された「めりはりのある高齢者」、「めりはりのない高齢者」、「成人」の各話者グループ間で遷移量の差分を算出し、遷移量の差分と発声時の調音動作との間にどのような関係があるかを解析した。その結果を元に、あらたに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フォルマントに着目して高齢者音声の解析を行う。最後に、その音響的特徴を用いて高齢者音声の明瞭性の改善方法を検討し、聴取実験によって改善効果を検証する。

4. 研究成果

(1) 音声データベース及び評価話者の選定

(1-1) 音声データベースの概要

高齢者音声の特徴を詳細に解析するためには、各音韻での特徴を分析する必要がある。そこで、本研究では電子技術総合研究所で選定した543語単語の音声を収録した音声データベースを使用する。音声データベースは音声データとラベル情報で構成されている。音声データに収録されている543単語は、音素出現頻度のバランスを考慮して選定されている。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1.1に音声データベースの各年代別人数を示す。

表 1.1 音声データベースの年代別人数

年齢	20代	30代	40代	50代	60代	70代	80代
男性	19名	4名	10名	13名	16名	18名	2名
女性	21名	6名	4名	4名	18名	17名	0名

また、音声データベースには音声情報の一部がラベルとしてそれぞれ付加されている。ラベルデータには各音素の始端終端や音素中心、エポックなどが含まれており、音素単位で分析が必要な本研究には欠かせないものである。ここでエポックとは、それぞれの音素の特徴を最もよく表現している時間軸上の特徴点のことを指し、例えば、母音は定常区間の中心点が、破裂音は閉鎖後の破裂部がエポックに相当する。

(1-2) 評価話者の選定

予備実験により、めりはり度合いを基準として、上位6名の高齢者を「めりはりのない高齢者（以降、めりはりなし）」グループ、下位6名を「めりはりありのある高齢者（以降、めりはりあり）」グループとして抽出した。高齢者女性も同様な傾向が伺えたが、本研究では、評価話者として、高齢者男性における「めりはりのない高齢者」6名、「めりはりのある高齢者」6名を用いて高齢者音声の解析を行う。また、比較のため6名の成人男性を「成人」グループとして抽出した。

(2) 遷移量と調音位置との関係

(2-1) 遷移量の定義

先行研究で、高齢者の声に「めりはり」がなくなると、隣接する音素間のスペクトル差である「遷移量」が低下することが述べられている。また、他の研究では、隣接する音素間における遷移量よりも、単語の先頭音素とその前に存在する無音区間との間における遷移量の方が、高齢化の影響がより顕著に現れることが述べられている。

そこで、本研究では、単語の先頭音素とその前に存在する無音区間との間の「遷移量」（以下、「語頭遷移量」と定義する）について解析することにする。先行研究と同様に、語頭遷移量は、(2.1)式によって求める。

$$DIS = \sqrt{\sum_{n=1}^{30} (C_n^{(t1)} - C_n^{(t2)})^2} \quad (式 2.1)$$

ここで、Cはケプストラム係数、t1とt2はスペクトル変化を求める2点の時間フレーム、nはケプストラム次数を表している。ケプストラム係数はFFTケプストラムを使用しており、サンプリング周波数24kHz、1

フレーム10 [ms]、FFT サイズ512、次数は30次まで用いた。t1, t2 は、語頭遷移量を算出する範囲(フレーム番号)であり、t1 は先頭音素の前に存在する無音区間の任意の点であり、t2 は単語の先頭音素のエポック点とする。

(2-2) 単語の語頭における遷移量

各単語の先頭音素ごとに算出した遷移量(語頭遷移量)を、「成人」、「めりはりあり」、「めりはりなし」の3つの話者グループで平均した結果を図2.1に示す。その結果、「めりはりあり」と「めりはりなし」を比較すると、全ての母音で「めりはりあり」よりも「めりはりなし」の遷移量が小さくなっていることがわかる。すなわち、「めりはりなし」の話者の場合、筋力が低下しているため、口や舌をほとんど動かしていない状態から発声動作を行うことが困難になり、結果的に遷移量が小さくなると考えられる。特に/i/や/e/が大きく低下していることがわかる。

一方「成人」と「めりはりあり」を比較すると、全ての母音で「成人」よりも「めりはりあり」の遷移量の方が小さくなっていることがわかる。この傾向は、「めりはりあり」と「めりはりなし」の比較の場合と同様であるが、その差は小さい。その中でも、/a/や/e/の遷移量の差は大きい、/u/や/i/の遷移量の差は小さいことがわかる。

(2-3) 単語の語頭における遷移量差の算出

各単語の先頭音素ごとの各母音の遷移量の低下の度合いを見るために、「成人」と「めりはりあり」と「めりはりなし」の話者グループ間の差分を「遷移量差」として、それぞれ算出した。例えば、「成人」と「めりはりあり」の間の遷移量差は、成人を基準として、「めりはりあり」の遷移量から「成人」の遷移量を引いた差として算出する。一方、「めりはりあり」と「めりはりなし」の遷移量差は、「めりはりあり」を基準として、「めりはりなし」の遷移量から「めりはりあり」の遷移量を引いた差として算出する。その結果を図2.2~図2.3にそれぞれ示す。図から、「めりはりあり」と「めりはりなし」との間の遷移量差において、/i/と/e/の遷移量差が小さ

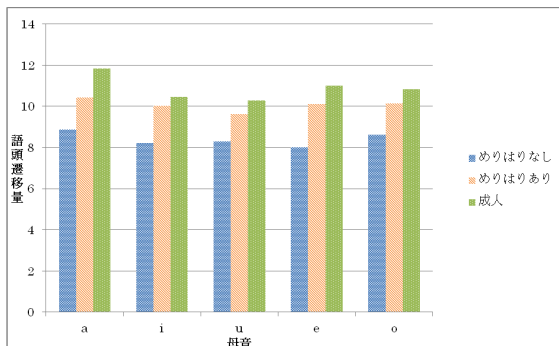


図2.1 各母音における語頭遷移量

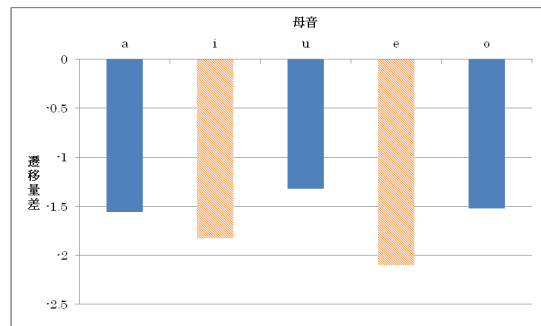


図2.2 語頭における遷移量差 (めりはりなし - めりはりあり)

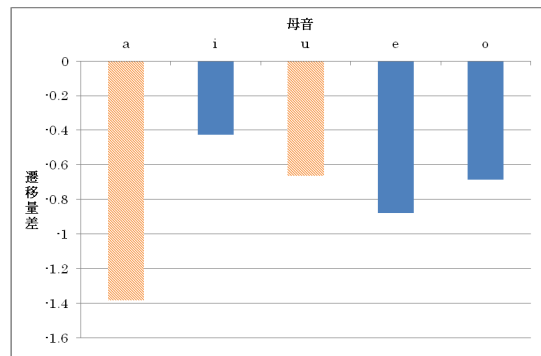


図2.3 語頭における遷移量差 (めりはりあり - 成人)

くなることわかる。反対に、/u/は遷移量差が負の方向に比較的小さいことがわかる。「成人」と「めりはりあり」との間の遷移量差において、/a/と/e/が特に遷移量差が負の方向に大きいことがわかる。反対に、/i/、/u/は遷移量差が負の方向に比較的小さいことがわかる。

(2-4) 単語の語頭における遷移量差の絶対値と調音位置との関係

前節の結果より、「めりはりあり」と「めりはりなし」との間の遷移量差において、口や舌を大きく動かす必要のある母音において、遷移量差の絶対値が大きくなっていることにわかった。そこで、「遷移量差」の絶対値と発声時の口の大きさや舌の位置(調音位置)との対応関係を整理した。図2.4は、「めりはりあり」と「めりはりなし」との間の遷移量差の絶対値と調音位置との関係を示す。同様に、図2.5は、「成人」と「めりはりあり」との間の関係を示している。縦軸が発声の際の口の大きさ、横軸が発声の際の舌の位置を表している。各母音の円の面積の大きさは、前節で算出した各母音の遷移量差の絶対値を半径として算出したものである。

その結果、「めりはりあり」と「めりはりのなし」との間の遷移量差の絶対値と調音位置との関係を見てみると、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。このことから、「めりは

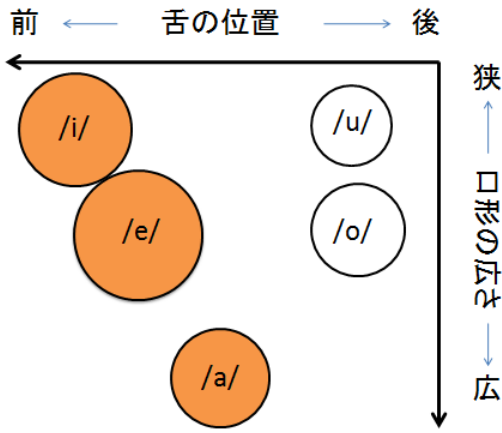


図 2.4 語頭における遷移量差の絶対値と調音位置との関係 (めりはりなし - めりはりあり)

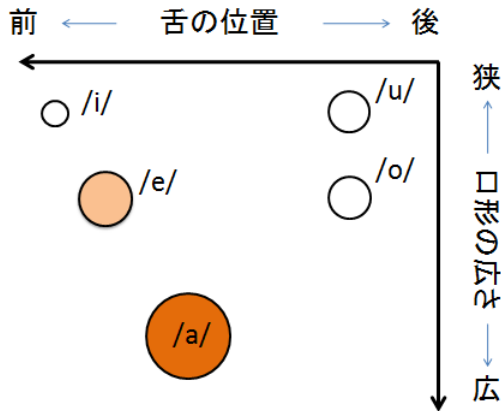


図 2.5 語頭における遷移量差の絶対値と調音位置との関係 (めりはりあり - 成人)

りなし」は、口や舌をあまり大きく動かすことができているために、正しく母音を発声できていないと考えられる。「成人」と「めりはりあり」との間の遷移量差の絶対値と調音位置との関係を見てみると、語頭では口や舌を大きく動かす/a/や/e/が、遷移量差の絶対値が大きいことがわかる。ただし、「めりはりあり」と「めりはりなし」の差に比べて差は大きくない。一方で、口をあまり動かす必要のない/i/、/u/では遷移量差の絶対値が小さくなった。このことから、高齢者は口をあまり大きく動かすことができているために、正しく母音を発声できていないと考えられる。

(3) 高齢者音声の F1-F2 平面上的の母音分布

高齢者は加齢により筋力が衰えており、口や舌を大きく動かす音素の発声が困難になっていると考えた。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フォルマントに着目して高齢者音声の解析を行った。

(3-1) 話者グループ別の F1-F2 平面上における 5 母音の分布

ここでは、「めりはりあり」、「めりはりなし」、「成人」の間におけるフォルマントの違いの全体的な傾向について述べる。図 3.1 は、単語の語頭における各母音の第 1、2 フォルマントをそれぞれ算出し、各話者グループで平均したものである。図から、声にめりはりがなくなると、/u/以外の母音の第 1、2 フォルマントが/u/に収束していることがわかる。これは筋力の低下から、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、「成人」と「めりはりあり」には各母音の分布が近いため、F1-F2 平面上の 5 母音の分布において高齢化の影響はほとんど見られない。

(4) フォルマント面積の拡張による明瞭性改善の検討

ここでは、前章で述べためりはりのない高齢者におけるフォルマントの特徴を用いて、実際に高齢者音声に対して補正処理を行う。そして、そのフォルマント補正による高齢者音声の明瞭性改善効果について、聴取実験を行って考察する。

(4-1) フォルマント補正方法の概要

前節で、めりはりのない高齢者はめりはりのある高齢者に比べて、F1-F2 平面上における 5 母音を頂点とした五角形の面積（以下、「フォルマント面積」と定義する）が縮小していることがわかった。そこで、ここではこのフォルマント面積の拡大を行い、フォルマント補正音声を作成する。

フォルマント補正処理の流れを以下説明する。まず、原音声から、母音区間を検出する。今回は、音を耳で聞いて知覚できる区間を母音区間とした。そして、その母音区間における第 1 フォルマント周波数と第 2 フォルマント周波数を時間フレームごとに算出する。時間フレームごとに算出された第 1、2 フォルマント周波数に対して、それぞれ異なるフォルマント補正值 f_1 、 f_2 を足し合わせる。最後に、フォルマントの遷移が滑らか

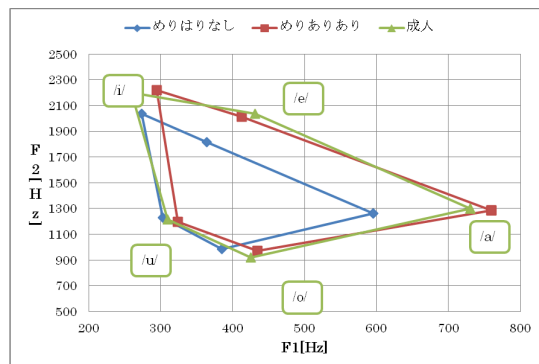


図 3.1 F1-F2 平面上における 5 母音の分布 (語頭)

になるように補正して、補正音声を作成する。フォルマント補正値の算出方法の概要を図 4.1 に示す。めりはりのある高齢者全体の平均フォルマント周波数から、被評価話者であるめりはりのない高齢者 1 名の平均フォルマント周波数を引いた差分をフォルマント補正値 f_1 、 f_2 として、母音ごとに算出する。また、語中、語頭では特徴が異なることから、語中、語頭それぞれでフォルマント補正値を算出する。

(4-2) 補正音声を生成する明瞭度変換システムのプロトタイプ構築

補正音声は、音声分析フリーソフトである Praat を用いた明瞭度変換システムにより作成する。明瞭度変換システムではフォルマント分析合成法を用いて、フォルマント補正音声を作成することができ、このプロトタイプを構築して実験を行った。フォルマント分析合成法とは、LPC 分析により、原音声のスペクトルの微細構造（音源情報）とスペクトル包絡パラメータに分離し、抽出したスペクトル包絡パラメータからフォルマントを抽出し補正した後に、再度音源情報とスペクトル包絡パラメータを再合成することで補正音声を作成する手法である。ここで「原音声」と作成した「補正音声」の時間波形とスペクトログラムを図 4.2、4.3 に示す。スペクトログラム上の赤色の線がフォルマント周波数を表している。下から順に第 1 フォルマント、第 2 フォルマントとなっている。原音声に比べて、フォルマント周波数の遷移が急峻になっていることがわかる。

(4-3) 評価システムの構築および聴取実験

前節で述べた補正方法による明瞭性改善効果について検証するために、Praat を用いた評価システムを構築し、実験を行った。パソコンのディスプレイに表示される画面のボタンを押すと、それぞれ異なる音声の流れ、- 2 ~ 2 までの 5 段階評価に対応した 5

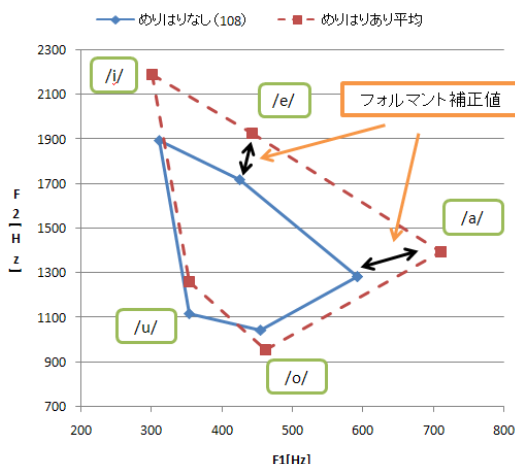


図 4.1 フォルマント補正値の算出方法

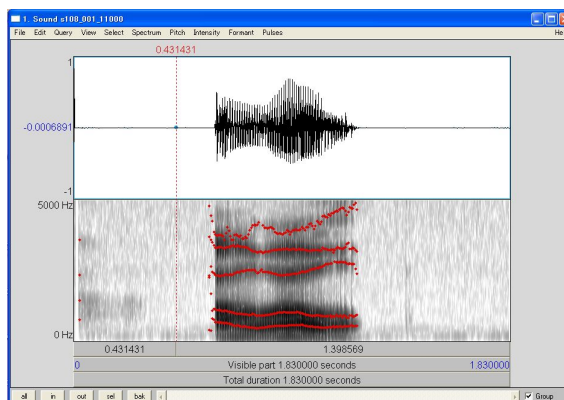


図 4.2 原音声の時間波形（上）とスペクトログラム（下）

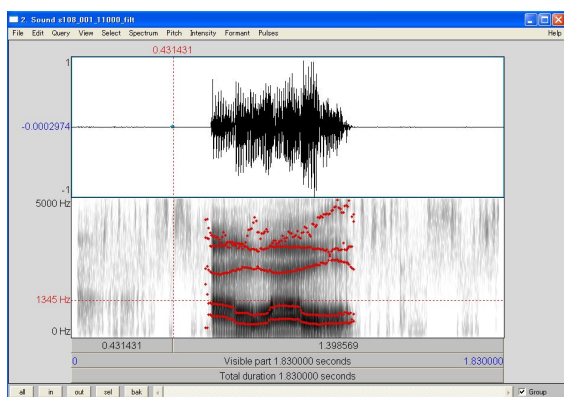


図 4.3 補正音声の時間波形（上）とスペクトログラム（下）

つのボタンのどれかを選択することで評価を行うことができるシステムである。主観評価には、シェッフェの対比較法（中屋変法）を用いて、実験結果の解析を行う。評価話者として、めりはりのない高齢者 6 名から、最も声にめりはりのないと思われる話者を任意で 1 名選出した。被験者としては、20 代の学生 10 名に協力してもらった。評価単語としては、音素バランスのとれた 5 4 3 単語から 2 ~ 3 種類の母音を含む 10 種類の単語音声を選択した。各々の単語について、「原音声」、フォ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フォ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。評価実験では、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 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。

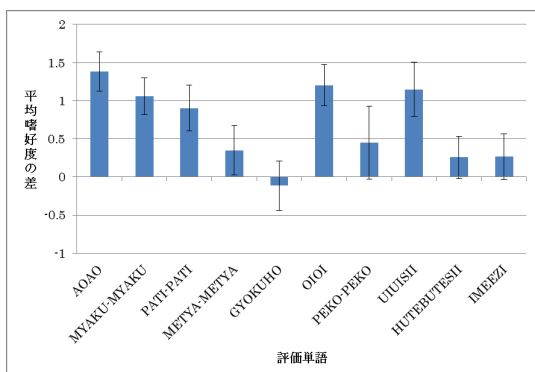


図 4.4 無補正音声と補正音声との間の平均嗜好度の差と 99%信頼区間

(4-4) 実験結果

図 4.4 に結果 (各音声間の平均嗜好度の差とその 99%信頼度区間) を図示する。まず、「無補正音声」と「補正音声」との比較を行う。図から、/AOAO/、/MYAKU-MYAKU/、/METYA-METYA/、/PATI-PATI/、/OIOI/、/UIUISI/の 5 つの単語の 99%信頼区間全体が、+側に位置しており、平均嗜好度の差が有意であると言える。このことから、今回の補正方法により明瞭性が向上したと考えられる。これらの単語は、発声する際に口や舌を大きく動かす必要があり、高齢者にとって発話が難しい単語のため、フォルマント補正による明瞭性の改善効果が顕著に現れたと考えられる。一方で、/GYOKUHO/や/IMEEZI/などは、99%信頼区間が+側と-側にまたがっているため、差が有意であるとは言えない。フォルマント補正による明瞭性の向上が有意とは言えない。これらの単語は、発話の際に、口や舌をあまり動かす必要がなく、高齢者にとって発話しやすい単語のため、フォルマント補正による明瞭性の改善が見られなかったと考えられる。これらの口や舌の動きと「めりはり」の関係は、高齢者音声の音響的特徴の分析結果とも一致する。

(5) まとめ

本研究では、高齢者音声の明瞭性を改善するために、高齢者音声の音響的特徴について解析を行い、その音響的特徴を用いた音声補正の明瞭性改善効果について検証した。まず、「成人」、「めりはりのある高齢者」、「めりはりのない高齢者」の各話者グループ間の遷移量の差分を「遷移量差」を母音ごとに算出した。そして、「遷移量差」の絶対値と調音位置との対応関係を解析した。その結果、どの話者グループ間においても、口や舌を大きく動かす/a/や/i/において、遷移量差の絶対値が特に大きいことがわかった。この結果から、遷移量に対応する音響的特徴として、発声の際の口の大きさや舌の位置に大きく関係するフォルマントに着目することにした。遷移量とフォルマントとの対応関係を解析するために、F1-F2 平面上における 5 母音の分布

を比較した。その結果、高齢者の声にめりはりがなくなると、/u/以外の母音が、/u/を中心として収束しており、5 母音を頂点とした五角形の面積 (フォルマント面積) が小さくなっていることがわかった。この結果から、めりはりのない高齢者の発話音声から補正音声を作成し、フォルマント補正による明瞭性の改善効果を検証した。聴取実験の結果、フォルマント補正によって口や舌を大きく動かす必要のある単語音声において、明瞭性が改善していることがわかった。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

- (1) Yoshiaki Yuda, Ryo Fukuhara, Mitsunori Mizumachi, Yoshihisa Nakatoh, "Investigation on variation of the initial phonemes of Japanese words due to aging", Proceedings of the 1st International Conference on Industrial Application, 2013, pp.49-52, (査読有)

[学会発表](計 6 件)

- (1) 福原諒, 水町光徳, 中藤良久, "F1-F2 平面上の母音間距離に着目した高齢者音声の解析, 日本音響学会秋季発表会, 2013 年 9 月 27 日
- (2) 中藤良久, "高齢者や障害者の支援に向けた音声技術の応用, 産業応用工学会全国大会 (招待講演), 2013 年 9 月 27 日
- (3) 福原諒, 湯田吉亮, 水町光徳, 中藤良久, "高齢者音声における音響的特徴と調音位置との関係", 日本音響学会春季研究発表会, 2013 年 3 月 13 日
- (4) 湯田吉亮, 水町光徳, 中藤良久, "遅い声に着目した語頭における高齢化の影響の解析, 電子情報通信学会総合大会, 2012 年 3 月 20 日
- (5) 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, 2012 年 3 月 20 日
- (6) 湯田吉亮, 福原諒, 井手隆宏, 水町光徳, 中藤良久, "語頭音素における高齢化の影響の検討", 日本音響学会九州支部学生のための研究発表会, 2011 年 11 月 26 日

6. 研究組織

(1) 研究代表者

中藤 良久 (Yoshihisa Nakatoh)
九州工業大学・大学院工学研究院・教授
研究者番号: 10599955

(2) 研究分担者

水町 光徳 (Mitsunori Mizumachi,)
九州工業大学・大学院工学研究院・教授
研究者番号: 90380740