

平成22年 3月31日現在

研究種目：基盤研究（C）

研究期間：2007～2009

課題番号：19560387

研究課題名（和文） 高齢者の音響的特徴抽出に基づく
高齢者音声認識インタフェースに関する研究研究課題名（英文） Research on speech recognition interface for elderly people
based on acoustic feature extraction of elderly speech研究代表者 二矢田 勝行(NIYADA KATSUYUKI)
九州工業大学・大学院 工学研究院・教授
研究者番号：50363396

研究成果の概要（和文）：高齢者音声の認識率改善を目的として、そのために、高齢者音声の音響的な特徴を調査した。まず、聴取実験によって高齢者音声と非高齢者音声の聴感的な違いを調べ、しゃがれ声、不明瞭な声、遅い発声が高齢者の特徴であることを示した。次に、それぞれの特徴に対応する音響的特徴を調べた。その結果、パワースペクトルの高域部の上昇、声帯発声のゆらぎ、音素間の遷移の緩慢さ、発話速度が遅くリズムが不規則であることが高齢者音声の音響的特徴であることがわかった。

研究成果の概要（英文）：This research aims at investigating acoustic features of elderly speech for improving speech recognition rate of elderly speakers. Listening tests are carried out to know difference of subjective impressions between elderly and non-elderly speech. We summarized that elderly speech was characterized by three adjective phrases: hoarse voice, ill-articulated voice and slow utterance. Then a physical characteristic of elderly speech is investigated corresponding to each adjective phrase. As the result, acoustic characteristics of elderly speech are; energy lift in higher frequency region in average power spectra, fluctuations in spectra during stationary parts of utterance, slow transition speed between adjacent phonemes, slow utterance speed and irregular utterance rhythm.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	1,700,000	510,000	2,210,000
2008年度	1,000,000	300,000	1,300,000
2009年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：音声分析

科研費の分科・細目：情報学 メディア情報学・データベース

キーワード：音声分析、高齢者音声、音声認識、音声データベース

1. 研究開始当初の背景

デジタル化によって、身のまわりのあらゆる電子機器が高度な機能を具え、そして

日々発展している。一方では、ユーザーの高齢化が進んでいる。これからは、高度な機能を高齢者でも使いやすくするためのヒューマ

ン・マシンインタフェース(HMI)が益々重要になってくる。その中で、「音声入力(音声認識)」は高齢化社会における有力なHMIと目されている。

機器の操作は、指先でキーや押しボタンによって行うのが一般的であるが、視力や注意力が衰えた高齢者は、そのために眼鏡をかけたり、細かい動作に集中するのは煩わしく感じる。その点、「声を出す」という動作は気軽に行うことができ、高齢者にとっても負担が小さい。そして、音声入力を用いる方法は、操作を簡易化できる効果が期待される。キーや押しボタンの数は限られているので、複雑な操作には階層化が必要になる。音声入力は多くの言葉から一語を選ぶことができるので、階層を意識することなく、所望の操作ができるというメリットがある。

これまでの音声認識技術は、主に一般成人男女を対象にして開発されたものであり、高齢者(一般に65歳以上)の声に対しては認識性能がかなり低下してしまう。高齢者の音声認識率を一般成人並みに改善することが最重要課題である。

2. 研究の目的

本研究は、高齢者用の音声認識インタフェースを実現するための要素技術研究であり、高齢者音声の音響的特徴を明らかにすることが目的である。本研究の特徴は、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、音声認識に用いる標準パターン(音響モデル)に明示的に反映させるようにすることである。また、聞きにくい高齢者の声を変換して聞きやすくすることも目的である。

3. 研究の方法

以下の手順で研究を遂行した。

(1) 年齢別データベースの整備

20代の成人~80代の高齢者の男女データをバランスよく収録し、データベース化する。

(2) データの試聴とアンケート調査により、高齢者音声の聴覚的特徴を表現語として網羅的に抽出する。

(3) 抽出された聴覚的特徴を発話特徴の要因別に分類する。

(4) 各要因に該当するデータの音響分析を行い、音響的特徴を物理量として抽出する。

(5) 音声認識性能の向上、聞きやすさの改善効果を確かめる。

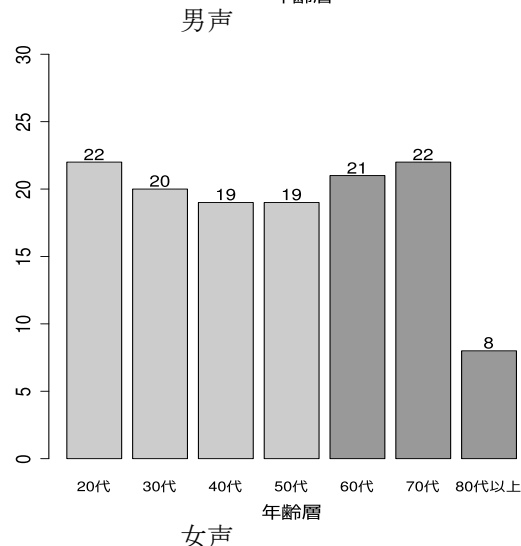
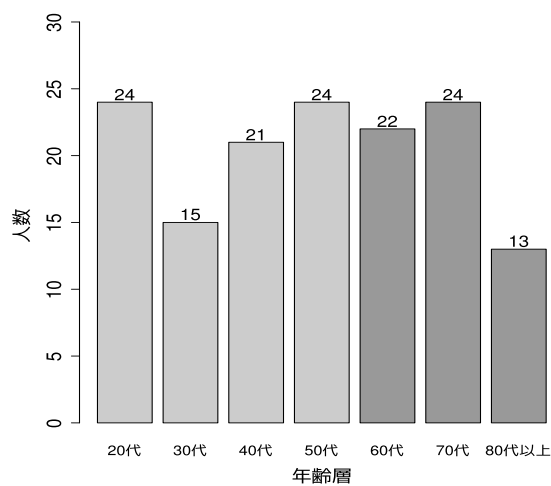
4. 研究成果

(1) 音声データベースの作成

音声の研究には、音声データベースが必須であるが、これまで高齢者音声分析のための

データベースが無かったため、新たに作成した。20歳~90歳までの男女の各543単語の発声を、男性142名、女性131名分収録した。年代別の人数バランスを下図に示す。80代以上は収録が難しいために男女とも人数が少ないが、大体人数バランスが取れている。

そして、音素別の詳細な音声分析を行うために、収録した単語音声データ毎に、目視及び自動により、単語内の各音素の時間的位置情報ラベルを付与し、研究用データベースとして整備した。



(2) 高齢者音声の聴覚的特徴の抽出

高齢者音声と非高齢者音声の聴覚的な違いを聴取実験とアンケートによって抽出し分類した。聴覚的特徴を表現する言葉を網羅的に列挙し、これを、被験者に対するアンケートと聴取実験によって絞り込んでゆき、高齢者音声の特徴を表現する言葉を抽出した。その結果を、階層的クラスタ分析を主とした統計解析手法によって整理する事で、男声では9語、女声では8語の高齢者音声の表現語を抽出した。さらに因子分析によって、これらを要因別に3つに分類した。手順を以下に示す。

辞書などから様々な発話の特徴を表す語（表現語）を収集

【男声：115語 女声：113語】

アンケートによって、特徴が想像出来る、理解性が高い語を選定

【残った語：男声61語 女声59語】

聴取実験により、高齢者音声の表現語の選定

【残った語：男声37語 女声29語】

アンケートにより、同義語を統合する

【クラスタ分析】

【男声10語 女声8語に統合】

聴取実験により、特徴別に分類する

【因子分析】

（男声3種類 女声3種類に分類）

結果として、

- ・ 声帯振動に起因する「雑音性の音質」（息漏れ、絞り出すような声、しゃがれ・ガラガラ声、震え）
- ・ 発声器官の動きの緩慢さによる「音韻的な不明瞭さ」（めりはりの無い声、暗い声）
- ・ 「発話の遅さ」（ゆっくりした声、たどたどしい声）

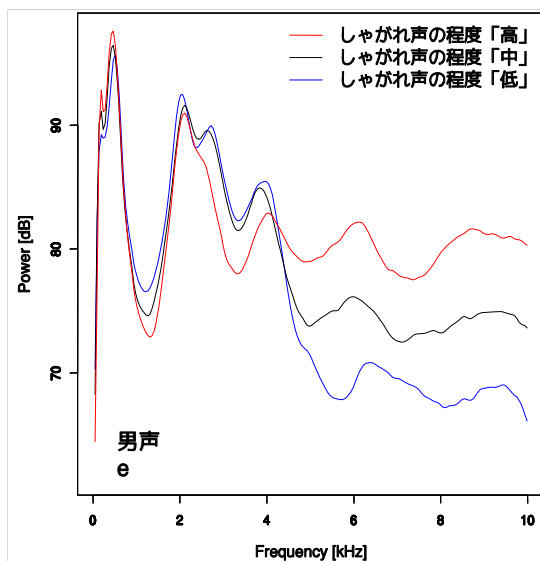
が高齢者音声の特徴として分類された。

(3) 「雑音性の音質」に対する音響的特徴

聴感的に雑音性の音質と感ずる話者の音声进行分析し、次の音響的特徴があることがわかった。

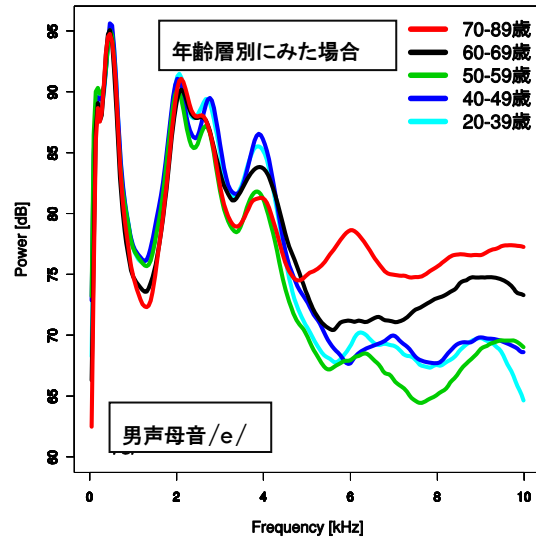
① 周波数領域での特徴

高齢者話者をしゃがれ度合い別に分類し母音部のスペクトル进行分析した結果、下図のように、しゃがれの度合いが高いほど、



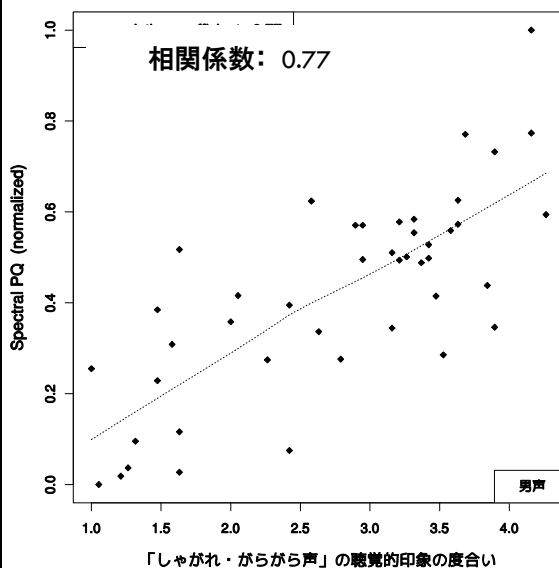
4 kHz以上の高域成分の割合が大きいたことがわかった。したがって、しゃがれ度は全スペクトルパワーに対する4 kHz以上の成分の割合で定量化できる。

次に、年代別にスペクトル特徴を観察すると下図のように、60歳以上の高齢者は高域成分が大きいたことがわかった。これは、高齢者にしゃがれ声が多いことに符合する。



② 時間軸方向の特徴

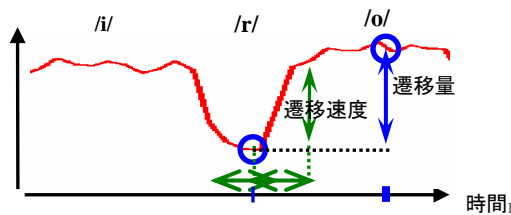
雑音性音質の話者は時間軸方向にも揺らぎがある。揺らぎの度合いをSpectral PQと言う尺度で定量化し、聴感的特徴との関係をプロットすると、下図のように強い相関があることが判明した。また、Spectral PQは年齢とともに上昇するので、時間軸方向のゆらぎは高齢者音声の特徴である。揺らぎが大きいと、ガラガラ声に聞こえる。



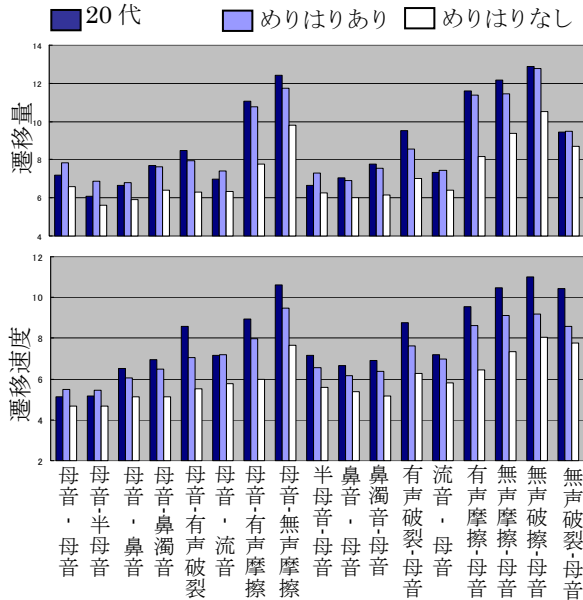
(4) 「音韻的不明瞭さ」に対する音響的特徴
 高齢者は筋肉の動きが緩慢になるため、調音の速い動きに対応できなくなる。それによって、各音韻の調音があいまいになるので、不明瞭な発声になると考えた。したがって、スペクトルの時間的な動きに着目し、2つの物理量を定義した。

- ・ 遷移量：隣接する2音素間の変化量
- ・ 遷移速度：2音素間の遷移部における単位時間当たりの変化量

下図は、例として「顔色」の/iro/の部分の遷移量、遷移速度の概念を示したものである。



高齢者を発声が明瞭なグループ（発声のめりはりあり）と不明瞭なグループ（めりはりなし）に分け、この2グループに20代の話者（非高齢、めりはりあり）を加え、3グループの遷移量、遷移速度を比較した。下図は、音素結合ごとの物理量の大きさである。

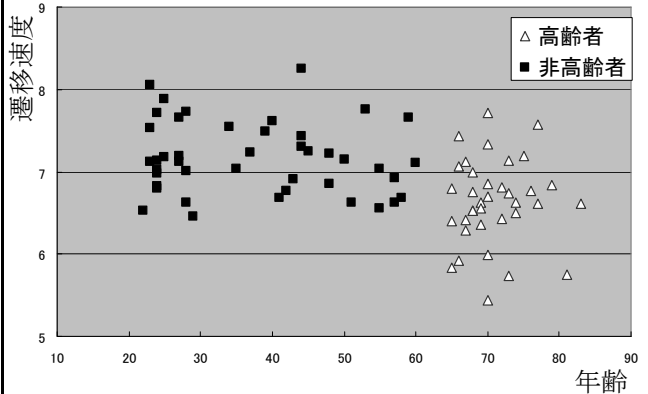


音素結合によって大きさは異なるが、相対的な大きさには、ほぼ同様の傾向が見られる。

- ・ 遷移量は、20代成人と高齢者めりはりありはほぼ同じ大きさであるが、高齢者めりはりなしの値は小さい。

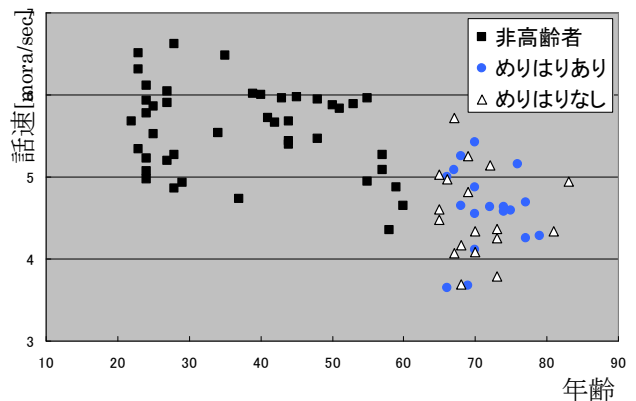
20代≧めりはりあり>めりはりなし
 すなわち、遷移量は発声の明瞭性と関係のある物理量である。

- ・ 遷移速度は、20代>めりはりあり>めりはりなしの関係がある。すなわち、遷移速度は発声の明瞭性のみでなく、年齢との関係がある物理量である。下図のように、個人差はあるが、高齢者の遷移速度は非高齢者よりも小さくなる傾向が見える。年齢とともに調音器官の動きが緩慢になっていることの証左である。



(5) 「発話の遅さ」に対応する音響的特徴

話速は、1秒間に話すモーラ（音節）数で表現される。下図は年齢と話速の関係を話者ごとにプロットしたものである。これを見ると、高齢者は明らかに話速が遅いことがわかる。すなわち、話速が遅いのは、高齢者の特徴であり、mora/sec で定量化できる。



高齢者は、発声のめりはりがある人と無い人を区別しているが、めりはり度と話速の間には相関が認められない。

ここで、上記(4)で述べた遷移速度と遷移量の関係を考察する。遷移量は、遷移速度と遷移時間（音素遷移に要した時間）の積に比例する。 $遷移量 \propto 遷移速度 \times 遷移時間$
 遷移時間は、発話速度が遅ければ、長くできる。すなわち、(4)で述べた めりはりのある高齢者は、遷移速度の遅さを遷移時間の長さ（発話の遅さ）でカバーし、遷移量を確保していると考えられる。一方、めりはりの無い高齢者は、話速を遅くして遷移時間を長

くしても、なお十分な遷移量に到達しないほど遷移速度が遅い話者であると言える。ここではデータを示さないが、破裂音⇄母音、無声摩擦音⇄母音のようにスペクトル変化が大きい音素結合部において、遷移速度の影響が顕著に見られた。

(6) 評価システム

上記、高齢者と非高齢者の特徴の違いを表現する物理量のいくつかに対して、それらの適用による音声認識性能の改善および聞きやすさの改善への効果を調査した。

① 音素認識率の評価

高齢者音声の「全パワーに対する 4 kHz 以上の割合」を非高齢者音声の平均値に合わせてしゃがれの補正を行い、高齢者の音素認識率を評価した。その結果、補正前の値が 86.30% であったものが、補正によって 87.50% に向上した。

ここで述べたのは基礎的な評価のみであるが、今後本研究成果を現行の音声認識システムに組み込めば、高齢者の認識率の改善が期待できることが示された。

② 聞きやすさの評価

分析合成システム STRAIGHT を用いて、高齢者の音声を分析してパラメータ化し、物理量によってパラメータを補正し、再び合成して試聴実験を行い、音質を評価した。その結果、以下のことがわかった。

・平滑化によってスペクトルの微細変動を除去すると、がらがら声が軽減され聞きやすくなる。

・音素の持続時間を調整してリズムを整えるとたどたどしさが軽減する。

・話速を上昇させると、めりはり度が改善し、たどたどしさも軽減された。そして、声年齢が若返るという評価が得られた。

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) Takeshi Miyazaki, Mitsunori Mizumachi, and Katsuyuki Niyada, “Acoustic analysis of breathy and rough voice characterizing elderly speech”, Journal of Advanced Computational Intelligence and Intelligent Informatics, 査読有, pp.135-141, Mar. 2010

[学会発表] (計 14 件)

国際学会

- (1) Daisuke Harada, Kenta Hamasaki, Takeshi Miyazaki, Mitsunori Mizumachi, and Katsuyuki Niyada, “Analysis of ill-articulated speech by elderly speakers

based on temporal transition of amplitude spectrum”, International Workshop on Nonlinear Circuits and Signal Processing 2010, Mar. 2010 (Honolulu Hawaii USA)

- (2) Takeshi Miyazaki, Mitsunori Mizumachi, and Katsuyuki Niyada, “The relationship between hoarseness and acoustic features for characterizing elderly speech”, Proc. Youngnam-Kyushu Joint Conference on Acoustics 2009, pp.89-92, Feb. 2009 (Pusan Korea)

- (3) Kenta Hamasaki, Takeshi Miyazaki, Mitsunori Mizumachi, and Katsuyuki Niyada, “Investigation on acoustic feature to characterize ill-articulated voice uttered by elderly people”, International Workshop on Nonlinear Circuits and Signal Processing 2009, Mar. 2009 (Honolulu Hawaii USA)

国内学会

- (1) 濱崎健太, 原田大輔, 宮崎健, 水町光徳, 二矢田勝行, “高齢者の「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年3月(東京)

- (2) 宮崎健, 水町光徳, 二矢田勝行, “孤立単語の語頭母音部にみられる高齢者音声の音響的特徴”, 音講論(春), 2010年3月(東京)

- (3) 濱崎健太, 原田大輔, 宮崎健, 水町光徳, 二矢田勝行, “高齢者の「めりはりのない声」に対応する物理量の検討”, 音響学会九州支部第8回学生のための研究発表会講論集, pp.25-28, 2009年11月(熊本)

- (4) 原田大輔, 濱崎健太, 宮崎健, 水町光徳, 二矢田勝行, “高齢者音声の減り張りのなさとしてスペクトルの時間遷移の関係”, 第62回電気関係学会九州支部連合大会論文集, 03-2P-08, 2009年9月(飯塚)

- (5) 宮崎健, 水町光徳, 二矢田勝行, “高齢者音声にみられる「しゃがれ声」の医学的分類に基づく音響的特徴”, 音講論(春), Vol. 1, pp.323-324, 2009年3月(東京)

- (6) 宮崎健, 水町光徳, 二矢田勝行, “高齢者音声の聴覚的特徴を形容する語の抽出に関する検討”, 信学技報, SP2008-64, pp.47-52, 2008年10月(熊本)

- (7) 宮崎健, 西将輝, 水町光徳, 二矢田勝行, “高齢者音声の聴覚的特徴と対応する音響的特徴に関する検討”, 第61回電気関係学会九州支部連合大会論文集, 03-2P-05, 2008年9月(大分)

- (8) 濱崎健太, 宮崎健, 水町光徳, 二矢田勝行, “高齢者の「めりはりのない声」を表す音響的特徴の検討”, 音講論(秋), Vol. 1, pp.355-358, 2008年9月(福岡)

- (9) 宮崎健, 水町光徳, 二矢田勝行, “高齢者音

声を印象付ける聴覚的特徴に関する検討”，音講論(秋)，Vol. 1，pp.283-286，2008年9月(福岡)

(10)宮崎健，水町光徳，二矢田勝行，“高齢者音声の特徴付ける表現語の抽出に関する検討”，音講論(春)，Vol. 1，pp.427-428，2008年3月(東京)

(11)井手康裕，宮崎健，原田勝也，水町光徳，二矢田勝行，“バズ音の長さに着目した高齢者音声の特徴解析”，第60回電気関係学会九州支部連合大会論文集，05-1A-16，2007年9月(那覇)

6. 研究組織

(1) 研究代表者

二矢田 勝行 (NIYADA KATSUYUKI)
九州工業大学・工学研究院・教授
研究者番号：50363396

(2) 研究分担者

水町 光徳 (MIZUMACHI MITSUNORI)
九州工業大学・工学研究院・助教
研究者番号：90380740