

Title	連続発話音声中に含まれる男声・女声知覚に寄与する音響特徴量に関する研究
Author(s)	柴田, 武志
Citation	
Issue Date	2008-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/4356
Rights	
Description	Supervisor: 赤木正人, 情報科学研究科, 修士

連続発話音声に含まれる男声・女声知覚に寄与する音響特徴量に関する研究

柴田 武志 (610043)

北陸先端科学技術大学院大学 情報科学研究科

2008 年 2 月 7 日

キーワード: 男声・女声知覚, STRAIGHT 分析合成法, テンポラルデコンポジション, 多次元尺度構成法, 静的・動的特徴.

人間は音声を聞くことで話者が男性か女性という性別の情報を得ることができる．これは人間が音声に含まれる男声・女声知覚に寄与する音響特徴量を知覚しているからである．これまで，男声・女声知覚は発声器官および調音器官の形状に起因する平均基本周波数やスペクトルである静的特徴が重要であるといわれてきたが，発声器官および調音器官の運動による基本周波数やスペクトルの時間変化パターンである動的特徴についてはあまり議論されていない．本稿では，静的特徴である平均基本周波数やスペクトルと動的特徴である基本周波数の変化やスペクトルの変化が男声・女声知覚にどのような順序で寄与しているかを明らかにすることを目的とする．

男声・女声の知覚を静的特徴、動的特徴と分けて調べるために、静的特徴、動的特徴を変形できる声質変換モデルが必要である。この問題を STRAIGHT(Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum) と MRTD(Modified Restricted Temporal Decomposition) を用いた声質変換モデルを提案することで解決する。声質変換モデルの流れを図 1 に提示する。図 1 ではまず音声を入力する。そして、その入力音声を最初のブロックである STRAIGHT 分析を行い 3 つのパラメータとして、基本周波数、スペクトル包絡、非周期成分を出力する。そして、スペクトル包絡は LSF(Line spectral frequency) というスペクトルパラメータに変換する。そして、次に MRTD を用いてスペクトルパラメータ (LSF) をイベントターゲット (静的成分) とイベント関数 (動的成分) に時間分解する。そしてイベント関数は、男声、女声の動きに変形して、MRTD によって合成する。得られた各パラメータを同じイベント関数で記述することで静的特徴と動的特徴をわけている。一方イベントターゲットは変形を行わず MRTD 合成する。その結果、男声から女声の静的、動的な特徴を変形することが可能になる。スペクトル包絡、基本周波数などのパラメータの変形が終わったら、最後に STRAIGHT を用いて合成する。その結果として合成音声地完成する。

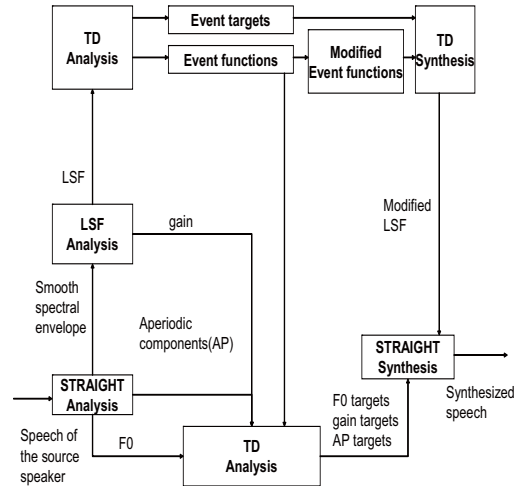


図 1: 声質変換モデルのブロックダイアグラム

声質変換モデルを用いて男声・女声音声から得られた各パラメータ値が男声・女声という分類で違いがあるかどうかを確かめるために多次元尺度構成法で分析を行った。その結果、男声と女声の音声では静的特徴（平均基本周波数、スペクトル包絡、ゲイン）および、動的特徴（スペクトル変化、音韻長、基本周波数の変化）に違いがあることがわかった。

次に分析によって、違いが見られた各特徴量を表すパラメータ値が男声・女声知覚にどう影響を与えているかどうかを調べるために、実験 1 では男声・女声音声から得られた各特徴量を表すパラメータ値を平均した音声に、男声・女声の静的特徴（平均基本周波数、スペクトル包絡、ゲイン）および、動的特徴（スペクトル変化、音韻長、基本周波数の変化）を付加した合成音声を用いて、男声・女声を判別する聴取実験を行った。実験 1 の結果、動的成分に比べて、静的成分の男声・女声知覚に対する寄与が高いことが明らかになった。男声・女声知覚について、静的な特徴である基本周波数とスペクトル包絡の影響が大きいということは、先行研究の結果を支持するものである。実験 2 では、実験 1 で影響の強い平均基本周波数とスペクトル包絡を固定し、実験 1 の結果から動的特徴の中で男声・女声知覚に影響を与えた“基本周波数の変化”、“語尾の変化”、“音韻長”といった特徴が知覚にどのような影響を与えているか調査した。実験 2 の結果から、語尾が動的成分の中で一番影響を与えていることが明らかになった。全体の傾向として、女声と判断された動的特徴を付加していくと、女声らしく知覚されるという結果が得られた。この結果は先行研究における女声らしさには話し方が影響を与えているという知見を支持するものであった。実験 1 と実験 2 の結果、男声・女声知覚には静的特徴である平均基本周波数とスペクトル包絡が大きな影響を与えており、次いで、動的特徴である基本周波数の変化と音韻長が影響を与えており、スペクトルの変化とゲインのダイナミックレンジはあまり影

響を与えていないことが明らかになった。

今後の課題として、実験2で行ったものは女声らしいについて調査したが、男声らしいについて調査していない。男声らしい特徴量はなんなのかということに対してさらに聴取実験を行い明らかにする必要がある。