

喉の衰えが音声にもたらす影響の音声セグメントを用いた推定

石井 航一[†] 井本 智明[†] 秀島 雅之^{††} 和田 淳一郎^{††} 松浦 博[†]
[†] 静岡県立大学経営情報学部 ^{††} 東京医科歯科大学歯学部

1. はじめに

喉の機能は嚥下(エネルギー摂取のために食べ物を飲み込む)、呼吸(酸素を肺に取り入れ二酸化炭素を排出する)、発声(他者とコミュニケーションをする)と極めて重要である。加齢等による喉の衰えを発話から早期に推定できれば、可逆的に対応できる^[1]。本報告では、独自開発の音声セグメント^[2]を用いて求めたパラメータから喉の衰えを推定する試みについて述べる。

2. データと分析結果

実験データには、20代~50代の話者が発話した日本音響学会新聞記事読み上げ記事コーパス(以下、JNAS)と、60代以上の高齢者が発話した新聞記事読み上げ高齢者音声コーパス(以下、S-JNAS)を用いた。ここでは音素バランス文の最初の「あらゆる現実をすべて自分の方へ捻じ曲げたのだ」を用いた。JNASは男性15名と女性16名、S-JNASは男性25名と女性28名である。各データを聴取し、明瞭度を5段階で評価した。

有声音長はアルファベット2文字で示される音声セグメントラベルに有声を示す文字を含むフレームの数とした。その部分のF0の分散が有声音部のF0分散である。分散とは隣り合ったフレームでのF0の差を2乗して足し合わせた値である。F0変化幅はF0の最大値から最小値を引いた値であり文全体での抑揚の大きさを示す。母音定常部は音声セグメントがAA,II,UU,EE,OOとなり、安定した母音とみなせる部分である。音声区間長は音声パワーの最大値の5%より大きい場合を音声区間としたフレーム数である。妥当セグメント比率は文中で出現しうる音声セグメントをあらかじめ設定して、その音声セグメントが現れたフレーム数の、音声区間長に対する割合である。妥当セグメントポイントとは、各モーラを構成する音声セグメントの組合せをあらかじめ設定し、そのラベルの組合せがいずれかのモーラで比率0.4以上

表1 各パラメータと明瞭度との相関係数 (JNASについて)

相関係数	有声音長	有声音部のF0分散	F0変化幅	母音定常部の比率	母音定常部のF0分散
明瞭度	0.25	-0.56	0.21	0.34	-0.11
相関係数	音声区間長	妥当セグメント比率	母音領域比率	母音領域F0分散	妥当セグメントポイント
明瞭度	0.29	0.36	0.31	-0.02	0.44

表2 各パラメータと明瞭度との相関係数 (S-JNASについて)

相関係数	有声音長	有声音部のF0分散	F0変化幅	母音定常部の比率	母音定常部のF0分散
明瞭度	-0.22	-0.35	0.27	0.06	-0.30
相関係数	音声区間長	妥当セグメント比率	母音領域比率	母音領域F0分散	妥当セグメントポイント
明瞭度	-0.22	0.00	-0.06	-0.36	0.36

の場合をポイントとして足し合わせた値である。

表1と表2に各パラメータと明瞭度との相関係数を示す。JNAS、S-JNAS共に比較的高い相関係数を得られた有声音部のF0分散と妥当セグメントポイントについて各データごとプロットし、それぞれ図1、図2に示す。

図1では明瞭度5の有声音部のF0分散は7000以下にすべて納まっているが、明瞭度1ではほぼ10000以上である。図2では明瞭度5の妥当セグメントポイントはすべて7を超えているが、明瞭度1では、ほぼ7以下となっている。しかし、S-JNASでは例外となる値もありIDを付記して確認すると、混同しているデータも他方のパラメータでは区別できていた。

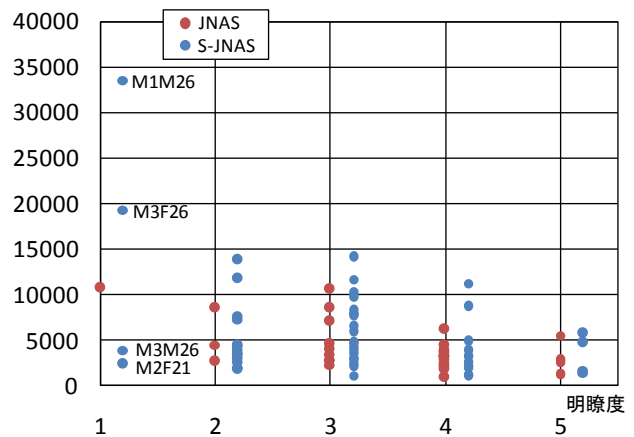


図1 明瞭度と有声音部のF0分散の分布

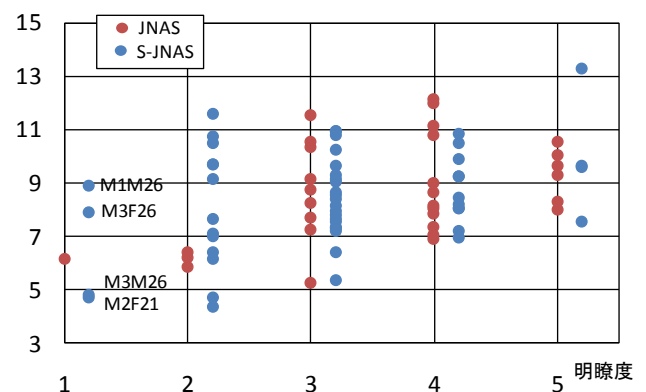


図2 明瞭度と妥当セグメントポイントの分布

5. まとめ

喉の衰えと考えられる声のかすれなどによって、明瞭度が低くなる発話について、音声セグメントを用いて求めた有声音部のF0分散や妥当セグメントポイントから推定できる可能性が得られた。

参考文献 [1] <https://www.jda.or.jp/enlightenment/oral/about.html> [2] 松浦博他, 情処学論, Vol. 46, No. 5, pp. 1165-1175 (2005).