

発声スタイルの違いによる母音の話者内分散の変化

杉田あおい[†] 細山友紀[†] 西村雅史^{††}
[†] 千葉大学

西田昌史^{††} 堀内靖雄[†] 黒岩眞吾[†]
^{††} 静岡大学

1. はじめに

人の発声する音声は、同じ音素列でも全く同じ音声波形となることはない。発声スタイルや込める感情の違い等で大きく変化することはもちろんだが、本人が同じように発声しようとしても音声波形は異なっている。特に、時期差がある場合は、一般的な音響特徴空間においても変化が大きいことが知られている[1]。そのため、話者認識だけでなく、現在我々が取り組んでいる、咽頭内の食物残留の検出を目的とした、音声から声道内の状態を知ろうとする研究[2]においても、話者内で変化しにくい音声を得られることが重要になっている。このような背景のもと、本稿では発声スタイルが異なる発声内の母音に着目し、対数パワースペクトル包絡の話者内分散[3]の調査を行った。

2. 分析方法

使用したデータは科警研の音声データ[4]から男性50名の短音節発声の/e/と音素バランス文(A03)の文節の切れ目の「で」から手動で切り出した/e/で、話者毎に時期を開けて2回ずつ録音した計4発声のデータである。音響分析条件はサンプリング周波数16kHz、高域強調係数0.97、フレームシフト10ms、窓幅32msのハンニング窓とした。スペクトル包絡はケプストラム分析(30次)を用いて抽出し、発声毎に平均を取ったのち正規化を行い、1発声で1つのスペクトル包絡を得た。このスペクトル包絡を用いて、話者毎に(1)式で分散を求めた。

$$\sigma_i^2(n) = \frac{1}{J} \sum_{j=1}^J \left\{ E_{ij}(n) - \frac{1}{J} \sum_{j=1}^J E_{ij}(n) \right\}^2 \quad (1)$$

ここで、 n :周波数 bin, i :話者, j :発声, $E_{ij}(n)$:話者の発声毎のスペクトル包絡である(本稿では $J = 4$)。

3. 分析結果

図1に話者1名の単母音/e/と音素バランス文から切り出した/e/のスペクトル包絡を示す。図より、単母音発声よりも音素バランス文発声の変化が少ないことがわかる。図2に、話者50名の周波数ごとの分散の平均を比較した図を示す。図2より周波数帯域で比較すると、4~7kHz範囲で分散に差があることがわかる。これより、4~5kHzの零点の変化から梨状窩における食物残留を検出する研究[2]等の、スペクトルの微細な変化に着目するような研究においては、単母音発声よりも文発声中の母音を使用する方が、安定した分析や検出を行うことが

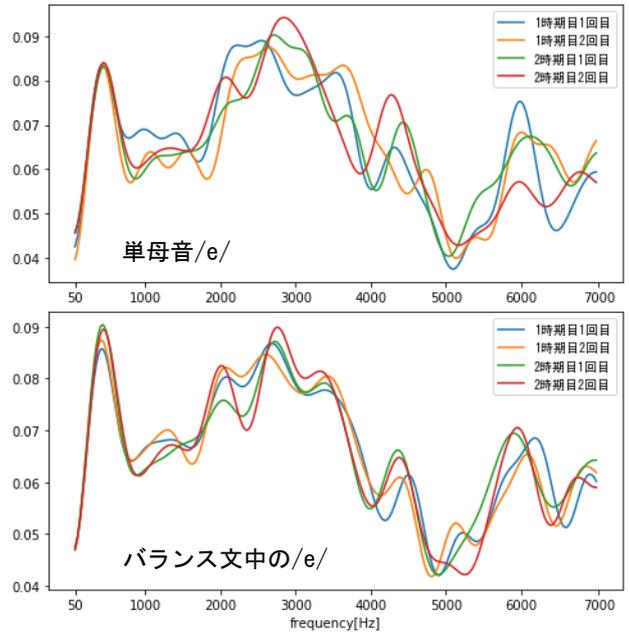


図1. 男性話者1名の単母音/e/と音素バランス文から切り出した/e/の各4発声のスペクトル包絡

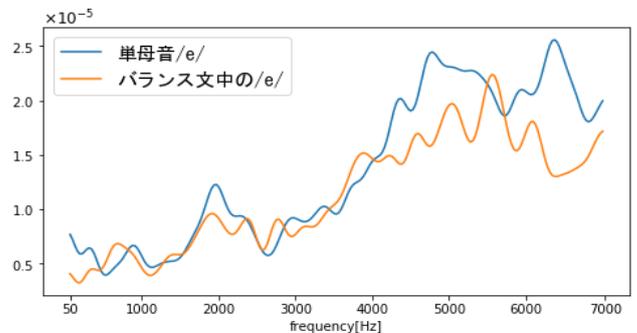


図2. 男性話者50名の単母音/e/と音素バランス文から切り出した/e/の周波数毎の分散の平均

できると考えられる。

4. おわりに

本稿では単母音発声と文発声中の、母音/e/のスペクトル包絡の話者内分散を比較した。今後は他の母音でも同様の比較を行う。

謝辞 本研究の一部はJSPS 科研費 18H03260, 18H01671, および、栢森情報科学振興財団の助成を受けた。

参考文献

- [1] 王ほか, 日本音響学会 69 巻 7 号, pp.357-364, 2013.
- [2] Hosoyama et al., KES-InMed, 171-177, 2020.
- [3] Kitamura et al., J. Acoust. Soc. Jpn.(E)16, 5, 1995.
- [4] 蒔苗ほか, IEICE Technical Report SP2007-40, 2007.