

風邪引き声の知覚について*

◎網野加苗, 菅原勉, 荒井隆行 (上智大)

1 はじめに

自動話者認識などの音声技術においては、特に冬期に、話者が風邪を引くことによる認識率低下の問題が指摘されている^[1]。また、鼻音などの有声音は話者の個性を反映しやすいと言われていた^[2-3]、風邪によって音響的に影響を受けやすいという点で、話者の判断における信頼性が不安定であることが指摘されている^[1,2,4]。

Tull ら^[5,6]は風邪引き声と通常音声と比較し、風邪引き声では調音点の変化や雑音成分の増加が見られることを確認した。また聴取実験によって、人間は風邪引き声と通常音声を区別できることを示した。さらに、メルケプストラム係数が風邪引き内、通常内で安定し、風邪引き-通常間で異なることを指摘した。

しかし、より具体的な風邪引き声の音響特徴については未解明である。本研究ではその第一歩として、風邪引き声の印象を与える特徴量が音声のどの帯域に含まれるのかを聴取実験を中心に調べた。

2 実験

実験条件を表1に示す。本実験の2名の話者は両者とも喉と鼻の両方に風邪の症状が見られた。風邪引き声の録音は最も症状が重い時に行い、通常音声の録音は完治後(ただし両者とも風邪引き声の録音から2週間以内)に同様の手順で行われた。

聴取実験は話者別にパートを分け、同じ帯域を阻止した風邪引き声と通常音声をペアで呈示した。音声ペアは11種類の異なる帯域を阻止した音声と原音声の計12種類とし、それぞれの繰り返し回数は10回であった。音声ペアはランダムに呈示され、各ペア内での風邪引き声・通常音声の前後の順は10回中それぞれ5回ずつとした。

被験者は防音室でヘッドフォンを通じて刺激音を両耳聴取した。実験の始めに各話者の風邪引き声及び通常音声のサンプル(実験文とは異なる文)を聴取させ、両者を区別できることを確認した。被験者には聴取した音声ペアに対して、表2に示す選択肢から回答するよう指示した。5つの選択肢を設けてあるが、実際に聴取させたペアには選択肢の1か2に該当するものしか存在しない。

3 結果

図1に各帯域10標本に対する話者別の正答率(被験者13名の平均)を示す。左端のcont.は原音声(control)である。2名の話者間で正答率に差が見られるが、これは一方の話者(Sp2)が聴取者のうちのより多くに通常音声

表1 実験条件

音声資料	2名の話者の風邪引き声と通常音声
サンプリング	20kHz, 16bit
実験文	ある時北風と太陽が力比べをしました。
被験者	13名(日本語母語, 健聴者)
聴取回数	12種×各10回(各話者120標本)
帯域阻止フィルタ	0-6000Hzの間を500Hz刻みで、1000Hzの幅を阻止
使用ソフト	Praat Ver.4.3.1.6

表2 聴取実験の回答の選択肢

1	音声ペアのうち前者が風邪引き声
2	音声ペアのうち後者が風邪引き声
3	音声ペアの両方が風邪引き声
4	音声ペアの両方が通常音声
5	音声ペアは別人の声の組み合わせなので比較不可

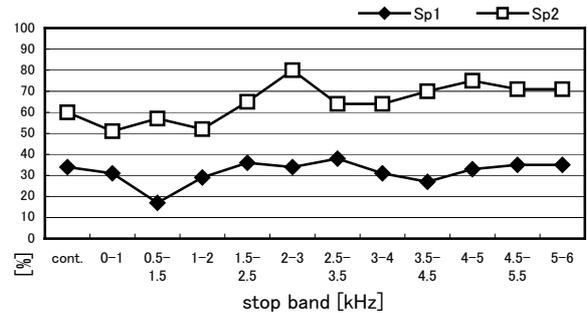


図1 選択肢1及び2と正答した割合(話者別)

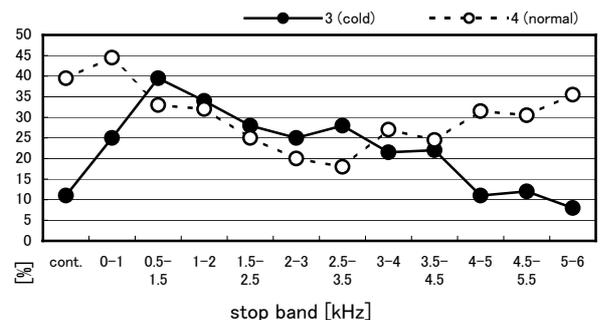


図2 選択肢3または4と答えた割合(話者平均)

を知られていたことと、風邪引きの程度に違いがあったこと等によると考えられる。阻止帯域別には目立った特徴はないが、Sp1では0.5-1.5kHzを阻止した場合に正答率が低く、Sp2では2-3kHzを阻止した場合に正答率が高い。

図2には選択肢3及び4の回答率(2名の話者の平均)を示す。実線が選択肢3、点線が選択肢4と答えた割合である。この図から、原音声では通常音声を風邪引き声と誤答するよりも風邪引き声を通常音声と誤答する率の

*A study on perception of cold-affected speech, by AMINO, Kanae, SUGAWARA, Tsutomu, and ARAI, Takayuki (Sophia University).

方が高いことが分かる。また、0.5~1.5kHz の比較的低い帯域を阻止した場合に通常音声（風邪引き声とする誤答率が最も高く、阻止帯域が高くなるほどこの誤答率は減少する。逆に風邪引き声を通常音声とした誤答率は低域及び高域を阻止した場合に高く、中間の2~3.5kHz 付近で低かった。なお、音声ペアは別人（選択肢5）と答えた割合は、Sp1, Sp2 でそれぞれ2.7%, 1.2%であった。

4 考察

4.1 スペクトル分析

聴取実験の結果は、風邪を引くと (1) 0.5~1.5kHz 付近で通常見られる何らかの音響特徴が阻害され、(2) 2~3.5kHz 付近の帯域に何らかの変化が現れるという可能性を示唆している。これらをふまえて、2名の話者の実験文と、それとは別に録音した母音、鼻音のスペクトルを観察した。その様子を図3, 図4に示す。図4の母音のスペクトルは日本語5母音の単独発話各6標本分、鼻音は /m/, /n/ の6標本分を平均したものである。

図3の実験文のスペクトルでは、上述の(1)に関して目立った特徴は見られないが、(2)に関しては両話者に共通して2~3kHz 付近で風邪引き声のスペクトル包絡が平坦になっていることが分かる。また図4で、母音の包絡には大きな変化は見られないが、鼻音では特に Sp2 において通常音声と風邪引き声で大きく異なっている。また両話者に共通して1.5~3kHz (Sp2 では1.5~4kHz) の帯域で風邪引き声のスペクトル包絡が平坦になっていることが分かる。逆に、母音・鼻音の0~1kHz と5~6kHz の帯域では両話者ともに風邪を引いても包絡はあまり変化していないと言える。

4.2 風邪引き声と個人性

本実験の話者のように、鼻腔・副鼻腔の炎症により鼻閉（鼻詰まり）が生じた場合、鼻腔に狭窄あるいは閉塞ができるため、一時的な閉鼻声（hyponasal）となる^[7]。閉鼻声では鼻腔共鳴が不足する上に軟口蓋挙上も不十分となるため^[8]、母音や半母音、そして特に鼻音が明らかに正常音声とは異なる音質を持つようになる^[7-10]。図4の鼻音のスペクトルが一部平坦になっているのは、鼻腔共鳴の不足によるものと考えられる。また、通常音声の非鼻音化母音や有声閉鎖子音でも鼻孔放射が起こることが知られている^[11-12]、閉鼻声になることによって、鼻孔放射が顕著な母音及び話者では、そのスペクトルも影響を受けることが予想される。

鼻音の第1~第5フォルマント周波数は話者の個人性を反映することが指摘されており^[9]、図4から本実験の2名の話者でも、通常音声では母音や実験文よりも鼻音のスペクトルにおいて明らかに話者間の違いが大きいことが分かる。しかし同時に、図4の風邪引き声のスペクトルでは個人の特徴が出にくくなっており、先行研究^[12]における予想が明示される結果となった。

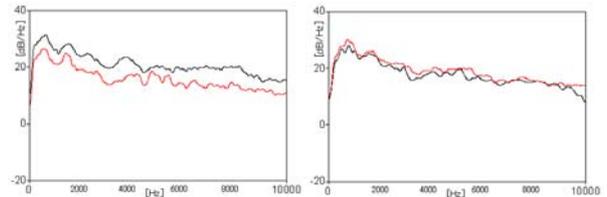


図3. 実験文の平均スペクトル

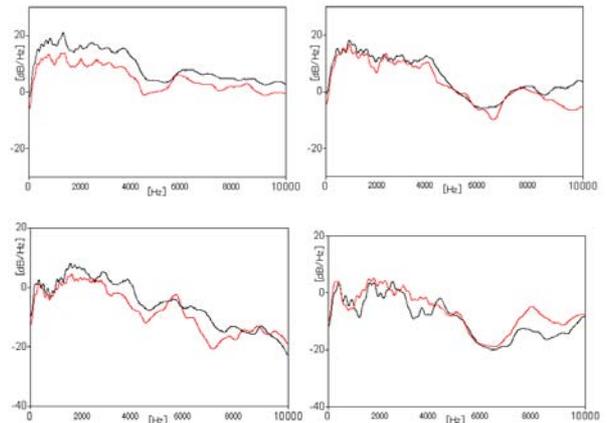


図4. 母音（上段）と鼻音（下段）の平均スペクトル
（図3,4共に、左: Sp1, 右: Sp2, 黒（濃）: 通常, 赤（薄）: 風邪引き）

前述の通り、自動話者認識では個人性を多く含みつつ、風邪などによって影響されない特徴量を用いるのが理想であり、また実音声のみではなく電話音声も利用できることが望ましい。本研究では、鼻音のスペクトルの0~1kHz 及び5~6kHz が比較的通常音声に近いという結果が出たが、これについては今後さらに複数の話者に関して再現性を検証する必要がある。

5 おわりに

風邪引き声の特徴を調べる第一歩として聴取実験を行い、風邪引き声らしさは2~3.5kHz の帯域に現れる可能性が示唆された。また、スペクトルの分析から、風邪を引くと1.5~3kHz 付近の帯域の包絡が平坦になることが分かった。さらに、鼻音のスペクトルは風邪引きによって話者の個人性を反映しにくくなることが示された。

謝辞

本研究は日本学術振興会特別研究員奨励費（17・6901）の助成を得た。

参考文献

- [1] Hardcastle and Laver (ed.), Blackwell, 1997.
- [2] Sambur, IEEE Trans. ASSP, vol.23 (2), 176-182, 1975.
- [3] 網野, 信学技報, SP2004-37, 49-54, 2004.
- [4] Nolan, Cambridge Uni. P., 1983.
- [5] Tull, and Rutledge, JASA, 99 (4), 2549, 1996.
- [6] Tull et al., JASA, 100 (4), 2760, 1996.
- [7] Laver, Cambridge Uni. P., 1980.
- [8] 内山 (監修), 内須川, 高野 (編集), 岩崎学術出版社, 1979.
- [9] パーマー (著), 田邊 (監修), 三田地 (訳), 学苑社, 2001.
- [10] Warren et al., Huffman and Krakow (ed.), Academic Press, 1993.
- [11] 党ら, 信学論 J74-A, 7, 1121-1124, 1991.
- [12] 党ら, 音響学会誌, 49 (5), 313-320, 1993.