

非可聴つぶやき声の音響的特徴*

○ 平原達也、清水奨太、大谷真（富山県立大学・工学部）、中島淑貴（奈良先端大学院大学）

1 はじめに

音声には通常の音声のほかに、歌声、叫び声、小声、ささやき声など様々な発声様式のものがある。これらの音声は、周期性又は雑音性の音源波を声道で共鳴させ口唇や鼻口から大気中に放射した気導音である。一方、声道内で共鳴した音は、口唇から放射されるだけでなく声道壁から体内組織にも伝わる。そのため、適切なセンサーを体表に装着することによって、声道共鳴音を骨導音や肉伝導音としても検出できる。すなわち、音声の伝導様式も多様である。

このうち、微弱な呼吸を駆動音源として生成される微弱な声道共鳴音である非可聴つぶやき声 (NAM: Non-audible Murmur) は頭頸部に装着した振動センサーで検出される[1]。このNAMは、様々な音環境において音声認識装置や携帯電話などを利用したり、発声障害者の音声コミュニケーション手段として利用する新しい「音声」として注目されている。本報告では、NAMの音響的特徴を解析した結果について報告する。

2 NAMデータの収録方法と分析方法

2.1 収録方法

NAMデータの収録は、無響室あるいは防音室内で行った。Fig. 1に示すソフトシリコン型NAMマイクロフォン（mitsumi電機）をネックバンドで頭頸部に装着した話者に、ATR音素バランス文50文章を読み上げさせた。これらの音声データは、ソリッドステートレコーダ（Marantz: PMD670）にサンプリング周波数 $F_s = 48 \text{ kHz}$ 、量子化精度 16 bit で収録した。

話者は健常者 14 名と発声障害者 4 名である。健常者については、NAM信号と通常音声をNAMマイクロフォンで検出したBTOS (Body Transmitted Ordinary Speech) 信号を収録した。一部の話者については、呼吸を流さずに、

調音器官だけを運動させた場合のNAM様発声音[2]と、人工喉頭（電制：ユアトーン）を用いて発話した代用音声をNAMマイクロフォンで検出した信号も収録した。

発声障害者については、咽頭炎による嘔声患者 1 名と喉頭摘出者 3 名のNAM等を収録した。嘔声患者からは、NAM信号と嘔声をNAMマイクロフォンで検出した信号を収録した。喉頭摘出者からは、食道発声音とNAM様発声音を収録した。NAM様発声音とは、発声時の調音運動によって生じた声道共鳴音をNAMマイクロフォンで検出した信号である。喉頭摘出者では呼気が声道に流れ込まないために健常者のようなNAMは生成できない。

2.2 分析方法

収録した音声データは、通常の音声と同様の音響分析を行った。すなわち、音声波形からパワーとパワースペクトログラムを算出するとともに、パワーとデルタケプストラムを利用して、音声区間を切り出した。そして、音声区間と雑音区間の平均パワーの比を取ることによってS/Nを算出した。また、音声区間のパワースペクトルを加算平均することによって長時間スペクトルを算出した。

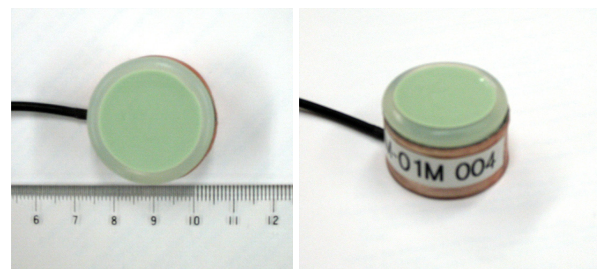


Fig. 1: Soft-silicone NAM microphone.

3 結果

3.1 音圧波形とスペクトログラム

Fig. 2に今回の分析したNAM信号の音圧波形とスペクトログラムの一例を示す。同図に示されるように、NAM信号には大振幅のポップノイズ（▼印）が重畳する。これは、NAM

* Acoustic characteristics of non-audible murmur, by HIRAHARA, Tatsuya, SHIMIZU Shota, OTANI Makoto (Toyama Prefectural Univ.) and NAKAJIMA Yoshitaka (NAIST)

センサーと皮膚との接触状態の変化や、調音運動に伴って生じる舌と口蓋や唇同士等の接触に起因するものである。これらのNAM特有のポップノイズはNAM信号と同じ帯域を持つために、スペクトル処理でノイズを低減することは困難であり、波形処理で低減した。

3.2 信号対雑音比 (S/N)

健常者が発話したNAM信号のS/Nは、 $F_s = 48\text{kHz}$ では15 dBであった。 F_s を低くしていくとS/Nは徐々に高くなるが最大でも21 dB ($F_s = 2\text{kHz}$)程度であった。

なお、S/Nの算出には3.1で述べたNAM信号に特有のポップノイズを波形処理によって低減した信号を用いた。ポップノイズはNAM信号よりも振幅が大きいため、それらを低減しないとS/Nはさらに低下する。

気導マイクロフォンで検出される通常の気導音と比べてNAM信号のS/Nはかなり低い。これは声道共鳴音が声道壁を通じて頭頸部組織へ伝わる際の反射、頭頸部組織の伝播損失、体表とNAMマイクロフォンの接触部での反射、NAMマイクロフォンの感度などの要因によって、NAM信号自体が低レベルになっているためである。

3.3 長時間スペクトル

Fig. 3に $F_s = 16\text{kHz}$ にリサンプルした男性6名と女声7名のNAM信号から得られた長時間スペクトルの平均を示す。女性1名については録音レベルが他のものより20 dB以上低かったために、分析対象から除外した。

男性のNAM信号も女性のNAM信号も、その長時間スペクトルは500~800 Hzで最大となっており、400 Hz以下では約-10 dB/oct.、800 Hz以上では約-17 dB/oct.で減衰している。男性のNAM信号の方が女性のものよりやや低域に伸び、高域側の減衰が大きい。これは声道長の違いによるものと考えられる。

また、これらの長時間スペクトルをノイズフロアと比較すると、今回用いたソフトシリコーン型NAMマイクロフォンで検出できるNAM信号の帯域幅は3 kHz程度であることが分かる。

なお、その後開発したウレタンエラストマー型のNAMマイクロフォン[3]では、検出されるNAM信号の帯域を広げることができることを別途確認している。

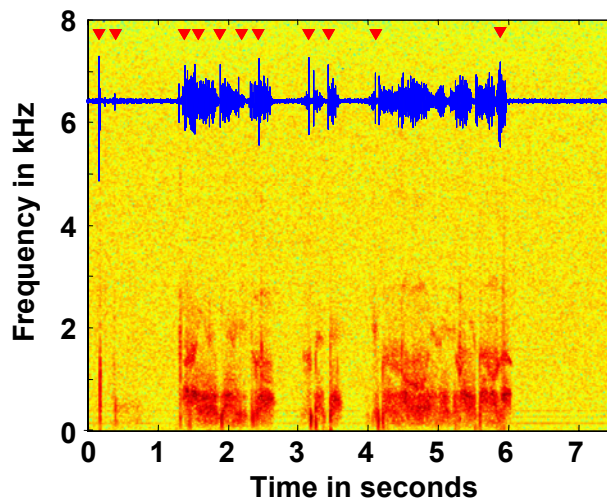


Fig. 2 Sound waveform and spectrogram of a typical NAM signal.

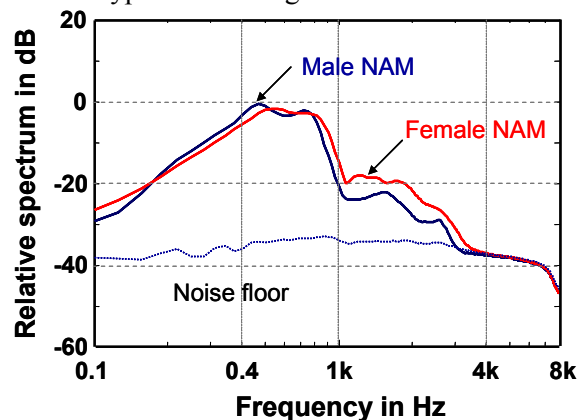


Fig. 3 Long-term spectrum of NAM signal.

4 まとめ

ソフトシリコーン型NAMマイクロフォンを用いて収録した、NAM信号の音響的特徴を解析した。その結果、NAM信号の帯域幅は3 kHz程度であること、平均S/Nは15dB程度であること、NAM信号の長時間スペクトルは500~800 Hzに山があり低域側は約-10 dB/oct.、高域側は約-17 dB/oct.で減衰していることが分かった。

本研究は総務省SCOPE-S『発声障害者の音声コミュニケーション手段の研究開発』により実施した。

参考文献

- [1] 中島淑貴 他 “非可聴つぶやきをインタフェースとするコミュニケーションのためのソフトシリコーン型NAMマイクロホンの開発,” 信学論, Vol. J89-D, No. 8, 1802-1810, 2006.
- [2] 中島淑貴 他 “喉頭摘出者のNAM様発声音” 音響学会講演論文集 2-4-11, 2006-03
- [3] 中島淑貴 他 “NAMマイクロフォンの日常使用に適する装着法とその音響的考察 -ウレタンエラストマー型NAMマイクロフォンの開発-” 音響学会講演論文集 1-Q-30, 2005-09