

# 無発声ないし微発声音声認識への活用を目的とした 表面筋電波形の調査

永井 秀利<sup>†</sup>, 竹下 舞<sup>†</sup>, 中村 貞吾<sup>†</sup>, 野村 浩郷<sup>†</sup>  
九州工業大学 情報工学部 知能情報工学科<sup>†</sup>

## 1 研究の目的

社会の情報化の進展に伴い、小型の情報機器を持ち歩く人も多くなっている。そうした携帯型情報機器で問題となるのは、操作その他のための入力手段である。本来、入力装置として最も確実なのはキーボードであるが、携帯型機器ではかさばって邪魔になりがちである。他の手段として有効なのは音声入力であろう。音声認識には多くの研究者が取り組んでおり、実用に供することが可能な程度の認識精度が得られるようになってきている。しかし実際に携帯型情報機器を街中で操作することを考えた場合、所構わず携帯電話で会話することで生じている社会的迷惑の発生状況や、情報機器の操作という個人のプライバシーや機密性の高い情報を第三者のいる前に出せるかということからすると、音声操作が期待通りに有効に使われるかは少々疑問である。

ここで発声という行為について考えた場合、たとえ実際には声を出さなくても、発声に関係する筋肉には発声時と似通った緊張が生じるとされる。そこで本研究ではこの筋活動に着目し、それに基づいて無発声または微発声(すなわち、マイク等で獲得できるような明確な発声がない)の状況での発声内容認識を行い、計算機への入力とすることを旨とする。本稿では、この目的のための筋電波形データの獲得を試みた結果について述べる。

## 2 発声内容認識への筋電活用

発声には、表情筋、舌筋、声帯筋といった複数の筋群が協調して働く。これらすべての筋の動きを正確に調べることができるなら、発声を再現することもかなり容易となる。筋の活動状況は筋電位で捉えることができるが、発声に関わるすべての筋の電位を正確に計測することは極めて困難である。針電極を使えば目標の筋の電位を直接に精度良く計測できるが、電極を目標筋へ刺入する必要があるため、これを一般の利用に供することは難しい。

侵襲のない筋電位の計測方法としては、体表に装着した一対の電極により計測する表面筋電がある。表面筋電は、複数の筋の活動電位が重なったものであり、目標筋の活動だけを正確に取り出すことはできない。また、深層の筋に対しては浅層の筋の活動が重なって得られるため、深層の筋の活動を正しく捉えるには困難が伴う。電極の接触抵抗や、動作に伴う装着位置の変位などによるノイズの多さも問題となる。しかし、発声内容の認識という観点からは、必ずしもすべての筋の正確な活動状態を知らねばならないというわけではない。そこで本研究では、利用上の負荷の小ささを鑑み、表面筋電に基づいて認識を行うものとする。

## 3 口裂周辺の表面筋電

日本語の5母音の発声においては、口唇の形状に比較的に顕著に特徴が現れる。また、両唇破裂音のように、発声の際に口唇形状が大きく変化する子音もある。口唇形状は主として口裂周辺の表情筋に支配されているため、口裂周辺の表面筋電を計測することにより、口唇に特徴が現れる発声内容の認識が可能であると考えられる。この観点から、口唇形状に関わるとされるいくつかの筋を対象として表面筋電の計測を行った。およその電極装着位置と各位置の目標筋およびその機能を図1に示す。

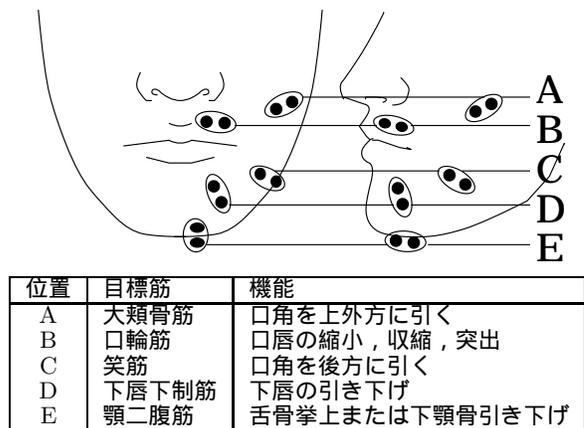


図1: 口裂周辺への電極装着位置の概略図

真鍋らの研究 [5][6] では、大頰骨筋(図のA)、口輪筋(B)、顎二腹筋(E)の3つを選定した実験の結果、日本語5母音の認識にかなり高い精度で成功している。だが、我々が調査した限りにおいては、大頰骨筋はあまり有効とは言えないという結果が得られた。

大頰骨筋は、意識的にかかなり不自然なくらい大きく口を動かした場合には確かに明確な反応が得られ、5母音の識別にも有効と思えた。しかし、無発声ないし微発声をより自然に行う程度の口の動かしかたをした場合、その反応は表面筋電波形上にほとんど確認できなかった。言い換えれば、ノイズに埋もれる程度にしか反応が得られなかった。大頰骨筋を対象とするのは、主として「い」や「え」の発声の際の口の形状を捉えることを意図したものであるが、軽く発声する場合にはほとんど口角を引き上げることはないように見受けられた。それに対して、笑筋の場合は、大頰骨筋よりは識別に利用できそうな反応が得られた。笑筋を計測する目的も大頰骨筋と類似しており、大頰骨筋を緊張させて笑筋を緊張させないという発声の仕方は考えづらいので、大頰骨筋は計測せず、笑筋で代表させればよいと思われる。

顎二腹筋を対象とするのは、主として「あ」や「え」の発声の際の口の開放を捉えることを意図したものであるが、反応が乏しくて発声していない状態との識別の用をなさない場合も見受けられた。これは、弱く発声する際

など、筋を弛緩させて顎の自重で下がるに任せたのに近い程度にしか口の開放を行わない場合があるためである。ただ、口を大きく開けた場合には明確な反応が出るため、発声の強弱を判定する(無発声の場合でも、強く発声しようとするれば筋も強く動かされる)際に有効であろう。

なお、下唇下制筋では、さして大きくは口を動かさない場合でも大きく動かした場合でも、識別に寄与できそうな良好な反応が得られた。

以上により我々は、口裂周辺の筋としては口輪筋、笑筋、下唇下制筋、顎二腹筋の4つを計測対象とすることを提案する。

## 4 頸部の表面筋電

口裂周辺の筋電のみを用いた場合、得られるのは口の形状だけであると考えてよいであろう。母音については、口の形状だけでもかなりの精度が得られるものと期待できるが、舌の動きを一切無視したのでは、子音の認識にはかなりの困難が伴う。また、口唇の形状変化のタイミングと発声開始、終了とは必ずしも一致していないため、実際の発声開始点、終了点の認識も問題である。読唇術というものが存在する以上、口唇の動きだけで発話内容を認識することも不可能ではないと言えるが、これを精度良く行うためには極めて高度な言語知識を駆使する必要がある。しかし、現在の自然言語処理技術はそこまで成熟してはいない。

舌の動きを支配する筋肉はほとんど表層には出ないため、正確な舌の動きを表面筋電によって精密に得ることは極めて困難である。しかしながら、おおよその舌の動きの有無だけでも得ることができれば、これを手がかりとして認識精度(特に子音認識精度)を向上させることができると予測される。

試してみればわかるが、呼吸を一切出さないようにして発声時と同様に口と舌とを動かすようにすることは少し難しい。そのため通常は、無発声の場合であっても微弱な呼吸吐出がなされ、それに伴って声帯周りの筋肉に発声時に類似した緊張が生じる可能性が高いであろう。よって、これを捉えることができれば発声の開始点、終了点の認識の手がかりとすることができると考えられるが、声帯は喉頭の軟骨の中にあるため、その動きを表面筋電で直接には捉えることができない。しかし、喉頭部の体表に電極を装着して喉頭部の活動に関わる筋の筋電を得たならば、そこから声帯の活動の有無を推測できる可能性がある。

これらの観点に基づいて舌および声帯の活動の手がかりを得るために、頸部の複数個所に電極を装着して、表面筋電を計測した。おおよその電極装着位置を図2に示す。

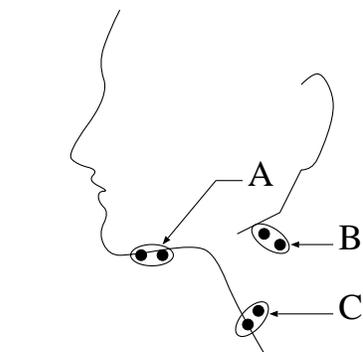


図2: 頸部への電極装着位置の概略図

図のAとBとは、舌の活動の手がかりの獲得を目的とした電極配置である。いずれも浅部の筋の谷間から、その奥にあって舌の動きを支配する舌外筋の一部の活動状態の計測を狙うものである。Aは図1のEの位置に類似しているが、実際には左右の顎二腹筋の間隙となるオトガイ下三角の位置から、その奥にあるオトガイ舌骨筋やそのさらに奥のオトガイ舌筋の活動を捉えることを目指したものである。Bは顎下三角に近い位置で、顎二腹筋や胸鎖乳突筋の影響をできるだけ避けて、深部の舌骨舌筋の活動を捉えることを目指している。

それに対し、図のCの位置は、声帯の活動の手がかりを獲得することを目的としている。声帯筋は甲状軟骨の内側にあるため、表面筋電で直接に捉えることは不可能と思われるが、喉頭外筋の活動は計測できる可能性が高い。この観点から、甲状軟骨の下部にあって声帯を緊張させる機能を持つ輪状甲状筋の活動を捉えることを目指したのがCである。もちろんこの位置は性差、年齢差など個人差が大きいいため、被験者に応じて位置を調整する必要がある。

五十音を1文字ずつ発声して計測した結果、発声の際の舌の動きに呼応した信号を筋電波形上に見ることができた。また特に子音の一部については、発声準備のための筋活動が実際の発声に先行して現れることも波形の上で確認できた。

## 5 まとめ

獲得された筋電波形を見る限りにおいては、少なくとも母音に関しては認識の手がかりとなる差異が得られているように見受けられる。また、舌の動きの一部を得ることができている点から子音の認識についても手がかりを得ていると考えられる。その波形の有意性や識別能力がどの程度かについてはさらなる分析が必要であり、これから研究を進めねばならない課題と言えよう。今後、波形特徴の獲得を目指すと同時に、測定した筋電波形データに基づく認識実験を行う予定である。

謝辞 本研究は、文部科学省/日本学術振興会平成14年度萌芽研究課題「筋電による計算機への自然言語インターフェースに関する研究」(課題番号14658103)での科学研究費補助金交付の下で行われた。ここに記して謝す。

## 参考文献

- [1] W.Kahle, H.Leonhardt, W.Platzer (越智淳三訳): “解剖学アトラス 第3版”, 文光堂(1990)
- [2] J.H.Worfel (矢谷令子, 小川恵子訳): “図説 筋の機能解剖 第4版”, 医学書院(1993)
- [3] 上條雅彦: “図説 口腔解剖学”, アナトーム社(1966)
- [4] 平山亮, Eric V. Bateson, Vincent L. Gracco, 川人光男: “神経回路モデルによる筋電信号からの発話時口唇運動軌道の生成”, 電子情報通信学会論文誌 A Vol.J79-A No.2, pp.371-382(1996)
- [5] 真鍋宏幸, 平岩明, 杉村利明: “無発声によるコミュニケーション技術”, NTT DoCoMo テクニカル・ジャーナル Vol.10 No.3, pp.43-47(2002)
- [6] 真鍋宏幸, 平岩明, 杉村利明: “無発声音声認識のための指輪型電極の提案”, 第1回情報科学技術フォーラム(FIT2002) 論文集, K-27(2002)