

隣接子音が母音声道形状に及ぼす影響の分析

☆天野沢海, 後藤翼, 竹本浩典 (千葉工大),
北村達也 (甲南大), 能田由紀子, 前川喜久雄 (国語研)

1 はじめに

われわれはリアルタイム MRI 動画による日本語調音運動データベースを構築してきた[1]。そして、動画の各フレームから機械学習を用いて半自動的に発話器官の輪郭を点群として抽出する研究も行なってきた[2,3]。現在、20名の話者の動画から、フレームごとに発話器官の5部位の輪郭を手動と同程度の精度で抽出することが可能になった[4]。

このデータベースでは、5部位の輪郭を、すべての話者の全てのフレームで解剖学的に相同な始点と終点の間を等しい点数の点群として抽出している。そこで、点群を輪郭に沿って等間隔に再配置すれば、発話器官の形状を話者間やフレーム間で定量的に分析できる。例えば、すべての話者の安静呼吸時における輪郭データを用いれば、平均声道形状やその個人差を分析できる[5]。

このような分析は、音素環境が声道形状に与える影響など動的な成分を検討する場合にも用いることができる。Cを子音、Vを母音とすると、本稿では、2モーラ語 $C_1V_1C_2V_2$ の発話において、先行子音 C_1 と後続子音 C_2 がそれぞれどれだけ V_1 に影響を与えたか、声道の形態距離を試験的に定義して評価したので報告する。

2 材料と方法

2.1 輪郭データ

データベースには、各話者がキャリア文「これが〇〇型」によって発話した2モーラ語 $C_1V_1C_2V_2$ の動画が含まれている。そこでまず、東京方言話者男女各1名(M1とF1)が $C_1=/m/$ or $/k/$, $C_2=/m/$ or $/k/$, $V_1=V_2=/a/$ or $/i/$ or $/u/$ の組み合わせで発話した12種の動画(例: $/maka/$)を選択した。次に、これらの動画の各フレームからすでに抽出されている発話器官の5部位の輪郭を、それぞれ終点から始点まで等間隔の点群となるように変換した。最後に、2モーラ語の動画から C_1 , V_1 , C_2 に対応するフ

レームを目視で選択した。なお、このデータベースの全ての動画は13.8 frame毎秒で記録されており、すべての話者の全てのフレームから舌を40点、上唇を15点、下唇・下顎を25点、軟・硬口蓋を30点、咽頭後壁を28点、喉頭を30点で輪郭を抽出している。

2.2 形態距離を用いた先行・後続子音の母音に対する影響の評価

2つの声道形状の形態的な距離は、対応する点同士の平均距離で評価することができるかと仮定した。本稿では、これを形態距離とする。もし、2つの声道形状がよく似ていれば形態距離は小さくなり、あまり似ていなければ形態距離は大きくなると考えられる。なお、輪郭を表す点群の座標は pixel で表されているので、単位も pixel である。

本稿では、2モーラ語 $C_1V_1C_2V_2$ における C_1 と C_2 のどちらが V_1 により大きな影響を与えているか、形態距離を用いて評価する。 V_1 は C_1 と C_2 のうち、より大きな影響を与えている方と発話器官の形態が類似し、形態距離が小さいと仮定する。すなわち、 C_1 と V_1 の形態距離を $|C_1V_1|$ 、 V_1 と C_2 の形態距離を $|V_1C_2|$ とすると、 $|C_1V_1| - |V_1C_2|$ が負なら C_1 の方が、正なら C_2 の方が、影響が大きいとする。

例として M1 の $/maka/$ を取り上げる。図1は、 $/m/$, $/m/$ と $/k/$ に挟まれた $/a/$, $/k/$ の輪郭である。 $/m/$ と $/a/$ の形態距離は 3.573 pixel, $/a/$ と $/k/$ の形態距離は 4.428 pixel なので、前者と後者の差は -0.855 pixel と負の値となる。よって、 $/a/$ は $/m/$ に形態的に近く、 $/m/$ の影響をより大きく受けているといえる。

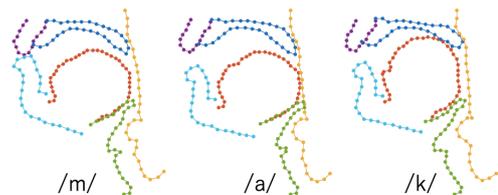


図1 $/maka/$ における発話器官の輪郭

* Analysis of effects of adjacent consonants on vowel vocal tract shape, by AMANO, Takumi, GOTO, Tsubasa, TAKEMOTO, Hironori (Chiba Inst. of Tech.), KITAMURA, Tatsuya (Konan Univ.), NOTA, Yukiko, and MAEKAWA, Kikuo (NINJAL).

3 結果と考察

表1は話者M1とF1の各音素環境における形態距離の差, すなわち, $|C_1V_1| - |V_1C_2|$ を計算した結果である。各列は C_1 , C_2 の順に先行子音と後続子音を示しており, 各行は母音 V_1 を示している。赤字が負の値, 青字が負の値である。つまり, 赤字は先行子音 C_1 の影響が大きく, 青字は後続子音 C_2 の影響が大きい。

表1 形態距離の差 (単位: pixel)

M1				
C_1, C_2	m, k	k, m	m, m	k, k
/a/	-0.855	2.060	0.734	0.026
/i/	-1.270	-1.572	1.426	0.402
/u/	-0.821	-0.438	-0.147	0.511
F1				
C_1, C_2	m, k	k, m	m, m	k, k
/a/	0.601	-2.124	0.387	0.008
/i/	0.778	-0.718	0.005	0.178
/u/	1.039	-0.514	0.463	0.222

まず, $C_1=/m/$, $C_2=/k/$ あるいは $C_1=/k/$, $C_2=/m/$ のときのF1の結果を考察する。表1は, F1では先行・後続に関わらず, 常に/m/より/k/が V_1 に与える影響の方が大きいことを示す。この原因を考察するために, /m/, /k/, V_1 の輪郭を重ねて比較した(図2参照)。 $/k/$ と V_1 では口の開きが同程度で, 口唇・下顎の輪郭がよく一致していた。一方, /k/と V_1 に比べて/m/では常に口の開きが小さく, 下唇から下顎にかけての輪郭, 特に下唇の輪郭が大きく異なっていた。すなわち, /k/と V_1 の形態距離が小さいのは下唇・下顎の輪郭がよく一致するためで, /m/と V_1 の形態距離が大きいのは下唇の輪郭が大きく異なるためである。

次に, M1の結果を考察する。表1は, /kama/を除いて C_2 より C_1 の方が V_1 に与える影響が大きいことを示している。この原因を考察するために, /m/, /k/, V_1 の輪郭を重ねて比較した(図3参照)。F1と同様に, /k/と V_1 では口の開きは同程度で, 口唇・下顎の輪郭がよく一致していた。そして, /m/と V_1 では下唇の輪郭が大きく異なる。そのため, /kimi/, /kumu/では/k/と V_1 の方が形態距離は小さくなる。ところが, $C_1=/m/$ のとき, F1ほど/m/の口の開きは/k/, V_1 と差がないため, /m/と V_1 の舌の輪郭はよく一致する。そのため, 下唇で輪郭が大きく異なっても, 結果として/m/と V_1 の方が形態距離は小さくなる。同様に, /kama/

では, 口の開きが/m/, /k/, V_1 でほぼ等しいため, 下唇で輪郭が大きく異なっても, 舌の輪郭が一致するために/m/との方が形態距離は小さくなったと考えられる。

最後に, $C_1=C_2=/k/$ あるいは $C_1=C_2=/m/$ の場合を考察する。M1の/mumu/を除き, C_2 の方が V_1 に与える影響が大きかった。これは, $V_1=V_2$ であるため, C_2 が同じ母音に挟まれて影響を受け, 形態が近づいたと考えられる。



図2 F1の/miki/ (左)と/kimi/ (右)の輪郭。
/m/: 青, /i/: 緑, /k/: 赤

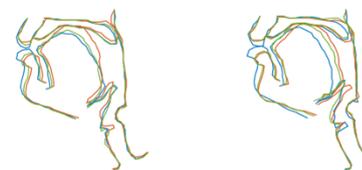


図3 M1の/miki/ (左)と/kimi/ (右)の輪郭。
/m/: 青, /i/: 緑, /k/: 赤

4 まとめ

本稿では, リアルタイムMRI動画から抽出した2名の話者の輪郭データを用いて, /k/と/m/が隣接する母音に及ぼす影響を, 形態距離の小さい子音の方が大きいと仮定して検討した。その結果, M1では先行子音が, F1では先行・後続にかかわらず/k/が, 隣接する母音により大きな影響を与えることが明らかになった。しかし, この違いが発話スタイルによるのか, 形態の差によるのかについては明らかにできなかった。なお, 形態距離により子音が母音に与える影響の大きさを評価することの妥当性は今後検討してゆく必要がある。

謝辞

本研究はJSPS科研費20H01265の助成を受けた。

参考文献

- [1] 前川ら, 音講論 (春), 1247-1248, 2018.
- [2] Takemoto *et al.*, Proc. of Interspeech 2019, 904-908, 2019.
- [3] 後藤ら, 音講論 (春), 779-780, 2020.
- [4] 天野ら, 音講論 (春), 745-746, 2021.
- [5] 後藤ら, 音講論 (秋), 655-656, 2020.