

モンゴル語の母音生成時の声道断面積関数の分析

☆加地優太, 竹本浩典 (千葉工大), 斎藤純男 (拓殖大), 前川喜久雄 (国語研)

1 はじめに

モンゴル語には母音調和が存在し、一語には男性母音 (/a/, /o/, /u/, /i/) と女性母音 (/e/, /ö/, /ü/, /i/) が混在することはない[1]。元来、男性母音は後舌、女性母音は前舌とされてきたが[1]、現代のモンゴル語では調音が変化して必ずしもこの基準が当てはまらなくなった[1]。それにもかかわらず、男性母音と女性母音の音韻対立は存続している。これは、前舌・後舌から、別の調音音声学的な対立へと移行したからではないかと考えられる。その有力な候補は舌根の前後位置による対立であり[1]、われわれは、モンゴル語の母音生成時の声道形状を MRI (Magnetic Resonance Imaging) で計測・分析し、それを確認できたが[2]、舌根の前後位置が音声にどのような対立をもたらすかは明らかになっていない。本研究では、声道形状とその音響特性を検討するための基礎資料として、MRI データから声道断面積関数を抽出し、音声を生成したので報告する。

2 材料と方法

2.1 実験参加者と MRI データ

実験参加者はモンゴル語 (巴林方言) 話者の女性 3 名 (F01~F03) である。実験参加者がモンゴル語の 8 母音 (/a/, /e/, /o/, /ö/, /u/, /ü/, /i/, /i/) を発声中の頭頸部の MRI 画像を ATR-Promotions の Siemens MAGNETOM Prisma fit3T を用いて撮像した。撮像は約 10 秒間の母音発声中に行われ、空間解像度 $1.0 \times 1.0 \times 3.0$ mm の母音 MRI データを得た。また、安静呼吸時に口を閉じて口唇、頬の内側、舌を歯列に密着させた状態で 15 秒間の撮像を行い、空間解像度 $0.3 \times 0.3 \times 1.0$ mm の歯列データを得た。また、MRI 実験後、F01, F03 の 2 名は静穏なホテルの部屋、F02 は内モンゴル大の録音室でモンゴル語 8 母音の録音を行った。なお、録音機材は SONY ELECTRET CONDENSER MICROPHONE ECM-44B (44.1

kHz, 16 bit) で、音声収録は仰臥位で行った。

2.2 歯列の補填

MRI では歯列も口腔も共に輝度値が 0 (黒) となる。したがって、口腔形状を正確に計測するためには母音 MRI データに歯列データから抽出した歯列を補填する必要がある。この操作には母音 MRI データと歯列データの座標系を対応させる剛体変換行列が必要である。従来法では、母音 MRI データと歯列データに共通する解剖学的なランドマークを用いて剛体変換行列を作成していた[3]。しかし、本研究の母音 MRI データは空間解像度が低く、ランドマークの特定が困難であった。そこで母音 MRI データと歯列データで下顎周辺などのほぼ対応する部位のボリュームデータを抽出し、これらの平均二乗誤差が最小になるような剛体変換行列を勾配降下法で求めた。この行列を用いて 3 名の実験参加者の 8 つの母音 MRI データに歯列を補填した。Fig 1 はその 1 例である。



Fig. 1 歯列を補填した F02 の/o/

2.3 声道断面積関数の抽出と調整

Takemoto *et al* [4]に従って、3 名の実験参加者の 8 母音の声道断面積関数を抽出した。しかし、空間解像度が低く、輪郭が不鮮明であったため、抽出した断面積関数は修正が必要であった。そこで、Story の手法を用いて、声道伝達関数の第 1~4 ピーク周波数が実音声の第 1~4 フォルマントと一致するように声道断面積関数を修正した[5]。これは、声道断面積関数から声道伝達関数のピーク周波数における断面積の音響感度関数 (断面積の微小変動に対するピーク周波数の変動量の関数)

*Examination of vocal tract area functions for vowels in Mongolian, by KAJI, Yuta, TAKEMOTO, Hironori (Chiba Institute of Technology), SAITO, Yoshio (Takushoku University), MAEKAWA, Kikuo (NINJAL)

を計算し、その感度が高い部分の微小変形を繰り返して行うことで、所期のピーク周波数を持つ声道断面積関数を得る手法である。

Fig. 2 上段は修正前後の伝達関数、下段は修正前後の声道断面積関数である。上段の黒い縦線は実音声から抽出した第1~4フォルマン트의周波数である。この操作により、声道形状が微修正されて所期のピーク周波数を持つ声道断面積関数が得られる。

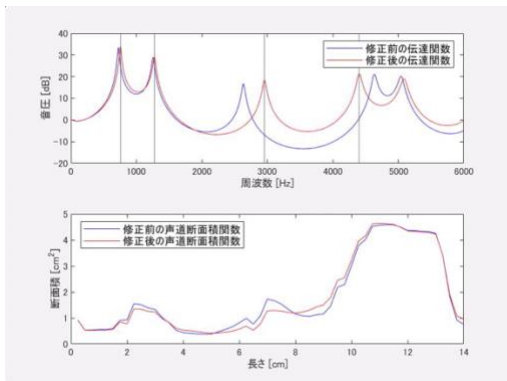


Fig. 2 声道断面積関数を自動修正した結果

2.4 音声の生成

Rosenberg 波[6]を用いて声帯音源波を生成した。その振幅、基本周波数、声帯振動に伴う乱流雑音は実音声から STRAIGHT[7]で抽出して与えた。そして、この声帯音源波に声道断面積関数から計算した声道インパルス応答を畳み込むことによって音声を生成した。

Table.1 音声のフォルマンと伝達関数のピークの誤差

		/a/	/e/	/o/	/ø/	/u/	/ü/	/i/	/y/
F01	F1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	F2	0.00%	-0.08%	0.09%	0.00%	0.00%	0.00%	0.07%	0.07%
	F3	0.00%	-0.03%	-0.07%	-0.03%	0.07%	0.00%	-0.08%	-0.05%
	F4	0.00%	-0.09%	-0.08%	0.03%	-0.08%	0.00%	0.07%	0.02%
F02	F1	-0.13%	-0.14%	-0.13%	0.00%	0.00%	0.00%	0.00%	0.00%
	F2	0.08%	0.08%	0.08%	-0.09%	0.00%	0.00%	-0.08%	-0.07%
	F3	0.03%	-0.04%	0.00%	-0.04%	0.00%	0.00%	0.03%	0.08%
	F4	0.07%	-0.02%	0.00%	0.00%	0.00%	0.00%	-0.02%	-0.09%
F03	F1	-0.10%	0.00%	0.00%	0.00%	0.00%	-0.50%	-0.24%	0.00%
	F2	0.07%	-0.06%	0.09%	0.11%	0.00%	0.61%	-0.07%	0.07%
	F3	0.04%	0.00%	0.04%	0.00%	0.00%	-0.04%	-0.06%	0.00%
	F4	0.03%	0.02%	0.00%	0.00%	0.00%	0.03%	-0.08%	-0.02%

3 結果・考察

Table 1 は実音声から抽出したフォルマントに対する伝達関数のピークの誤差を示す。全ての誤差が1%未満となった。Table 1 の赤で示した音韻は、モンゴル語（巴林方言）母語話者1名によって生成した母音の音韻性が不十分であると評価されたものである。音韻性は24の音声のうち16で十分であったが、8

で不十分であった。誤差が非常に小さかったことから、声道断面積関数の修正の問題ではなく、実音声から抽出したフォルマント周波数に問題があった可能性が高いが、本研究では明らかにすることができなかった。

Fig. 3 は F03 の男性母音/u/と女性母音/ü/をトレースして重ねたものである。女性母音で舌根が前方に変位して咽頭下部が拡大していた。これはほぼすべての男性母音と女性母音のペアでみられた[1]。しかし、F01の女性母音/e/では男性母音/a/より喉頭が下降することにより、咽頭下部を拡大していた。

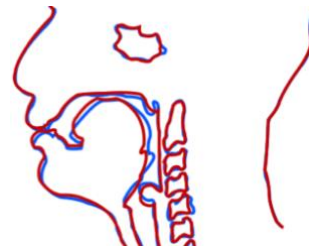


Fig. 3 F03 の/u/ (青) と/ü/ (赤)

4 まとめ

本研究ではモンゴル語の母音調和の声道形状と音響特性を検討するために、MRI データから声道断面積関数を抽出し、その伝達関数のピークと実音声のフォルマンとの誤差が1%以下になるように断面積関数を修正した。そして、生成した母音をモンゴル語母語話者が評価した結果、24母音中16母音で音韻性が十分であることが明らかになった。また、女性母音では、咽頭下部が拡大していることが明らかになった。

謝辞

本研究では、JSPS 科研費 20H01265 の支援を受けた。

参考文献

- [1] Jan-Olof Svantesson *et al.*, *The phonology of Mongolian*, Oxford University Press, 2005.
- [2] Y. Saitō *et al.*, *Proceedings of ICPHS 2019*, 1431-1434, 2019.
- [3] H. Takemoto *et al.*, *AST.*, 468-474, 2004.
- [4] H. Takemoto *et al.*, *JASA.*, 1037-1049, 2006.
- [5] B. H. Story, *JASA.*, 715-718, 2006.
- [6] A. E. Rosenberg, *JASA.*, 583-590, 1971.
- [7] H. Kawahara, *Speech Communication*, 187-207, 1999.