

超音波画像法にもとづく舌の相対的サイズ計測の試み*

◎孫静 (神戸大), 北村達也 (甲南大), 能田由紀子 (国語研), 牧野桃子, 談曉鋒,
山根典子 (広島大), 林良子 (神戸大)

1 はじめに

発話の仕方には個人差が見られ、滑舌や明瞭度に影響を及ぼすことがある。特に、舌の動きが鈍い場合には、構音障害の特徴として現れることもある。もし健常者においても舌の動きに差異を生じさせる要因が存在するとすれば、その一因として、口腔や咽頭腔の大きさに対する舌の相対的なサイズ (RTS: Relative Tongue Size) の違いが考えられる[1]。RTS を計測することで、こうした差異の原因を数値的に捉えることが可能と推測される。Feng ら[1]では MRI の正中矢状断面を用いて RTS を測定しているが、MRI を用いる方法は高コストであるという課題がある。本研究では、超音波画像法を用いて RTS を計測し、MRI による RTS との比較をする。また、RTS が舌の運動速度とどのように関連しているかについても詳細に検討を行う。

2 データ収集

2.1 実験参加者

日本語母語話者 8 名 (男性 3 名, 女性 5 名) と、中国語を母語とする日本語学習者 2 名 (男性 1 名, 女性 1 名) の計 10 名が対象である。学習者はそれぞれ日本語能力試験 (JLPT) の N1 および N2 の合格者である。超音波実験と rtMRI 実験は同一の話者で実施した。

2.2 超音波データの収集

超音波ゼリーを塗布したマイクロコンベックスプローブ (8C-Rs) を実験参加者の下顎にあて、超音波断層撮像装置 (LOGIQ e Premium) によって正中矢状断面における舌の運動を B モードで撮像し、動画 (60 fps) として記録した。プローブは UltraFit ヘッドセットにて固定し、発話時には頭を動かさないよう指示した。発話資料は日本語汎用版「北風と太陽」[2]を使用し、計 327 モーラの音声収録した。

2.3 rtMRI データの収集

rtMRI データは 3 種類あり、「リアルタイム MRI 日本語調音運動データベース (rtMRIDB)」

[3]から抽出した 3 名、「神戸大 日中対照調音動態 MRI 動画コーパス (KUJC-MRI)」[4]から抽出した 1 名、筆者らが撮像した 6 名である。これらのデータから日本語母音 /i/ 発話時のフレームを抽出して利用した。

筆者らのデータは脳活動イメージングセンタの 3T MRI 装置 (MAGNETOM Verio, Siemens) を用いて発話中の正中矢状断面の rtMRI を撮像した。撮像条件は、画像サイズ 256×256 ピクセル、解像度 1 mm×1 mm、スライス厚 10 mm、フレームレート 13.79 fps である。

3 RTS の計測方法

3.1 rtMRI データによる RTS の計測方法

RTS は、Feng ら[1]が提案した以下の数式で求めた。

$$RTS = \frac{A_t}{A_t + A_o}$$

ここで、 A_t は母音 /i/ 発話時の舌の面積、 A_o は口腔+咽頭腔の面積である。

3.2 舌輪郭抽出法

DeepLabCut (DLC) [5]を用いて超音波データと rtMRI データから舌輪郭点を抽出するモデルを作成し、舌輪郭点を自動的に抽出した。Fig. 1 には日本語母音 /i/ 発音時の舌輪郭点の一例を示す。

3.3 超音波データによる RTS の計測方法

本稿では超音波データから求めた RTS を uRTS と表す。超音波データでは口腔や咽頭腔が直接見えず、舌のみが観察可能である。そこで、様々な子音や母音 (頭子音 /-, k, s, t, h, r, ɕ/ (-は頭子音を持たない時) + 日本語単母音 /a, i, u, e, o/ を組み合わせた 35 語) の発話時における舌の輪郭座標を利用し、その凸包の面積 A_c を計測する。この凸包面積 (mm^2) は、日本語発話時に舌が動く範囲 (可動域) を表す。母音 /i/ 発話時の舌の面積 A_t は定常部の中央の 1 フレームから計算した。本研究では uRTS を以下の式にて定義する。

* An Attempt at Measuring the Relative Tongue Size Based on Ultrasound Imaging, by SUN, Jing (Kobe Univ.), KITAMURA, Tatsuya (Konan Univ.), NOTA, Yukiko (NINJAL), MAKINO, Toko, TAN, Xiaofeng, YAMANE, Noriko (Hiroshima Univ.), HAYASHI, Ryoko (Kobe Univ.).

$$uRTS = \frac{A_i}{A_c}$$

なお、 A_c と A_i を計算するとき、ともに座標(0, 0)を含めた。Fig. 2に凸包と母音/i/のトレースを示す。

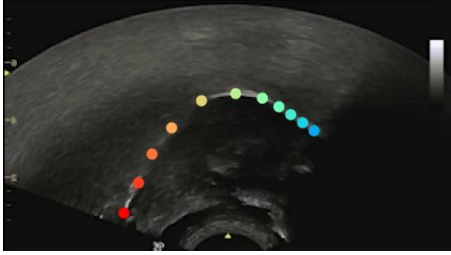


Fig. 1 Tongue shape for the Japanese vowel /i/ extracted using DLC.

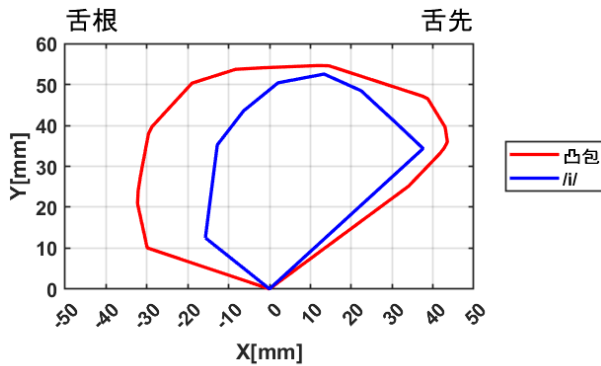


Fig. 2 Convex hull of the tongue range of motion (red) and Japanese vowel /i/ (blue).

3.4 舌の運動速度の計測方法

舌の運動速度 V は、「北風と太陽」の超音波データを用いて計算した。Takemoto ら[6]が開発した評価法を利用し、発話の最初から最後まで舌運動を対象とした。第 n フレームの各舌輪郭点から第 $n+1$ フレームの舌輪郭までの最短距離（垂直距離）を計算した。これらの距離を d_n ($n = 1, 2, \dots, N$) とし、舌の運動距離 D を以下の式で定義する。

$$D = \frac{1}{N} \sum_{n=1}^N d_n$$

ここで、 N はフレームにおける舌輪郭点の総数を表す。舌の運動速度 V [mm/s] は、以下に示すように D をフレーム周期 t で割り、さらに、1 秒あたりの発話モーラ数 s で割ることにより求める。後者は、話者間で発話速度が異なるため、これを補正する目的である。

$$V = \frac{D}{t \cdot s}$$

4 実験結果と考察

rtMRI データと超音波データで計算した RTS 値の相関を Fig. 3 に示し、uRTS と舌の運動速度の相関を Fig. 4 に示す。

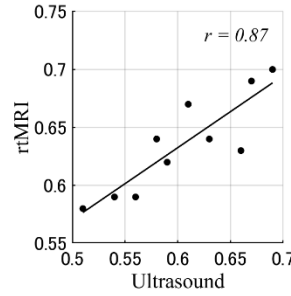


Fig. 3 Correlation between uRTS and RTS.

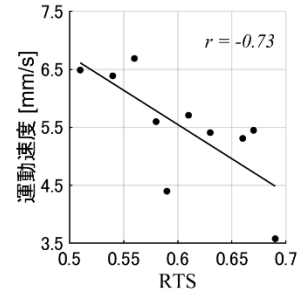


Fig. 4 Correlation between uRTS and tongue velocity.

Fig. 3, 4 を見ると、rtMRI データから求めた RTS と超音波データから求めた uRTS の相関が高く ($r = 0.87$)、両者の結果がほぼ一致していることがわかる。uRTS と舌の運動速度の間には、負の相関が見られる ($r = -0.73$)。つまり、uRTS が大きいほど運動速度が遅い傾向がある。このように、RTS が大きいほど、舌の移動空間 ($A_c - A_i$) が狭くなり、移動速度が低下し、明瞭な音声の発音が困難になり、滑舌が悪化する可能性があると考えられる。

5 まとめ

本研究では、超音波データを用いて uRTS を計測し、rtMRI データで計測した RTS と比較した。その結果、これらには高い相関が認められ、超音波データでも RTS を測定可能であることが示された。

謝辞 研究の一部は、JSPS 科研費 (Nos. 24K00071, 24K00067, 22K00621)、甲南デジタルツイン研究所の支援を受けた。

参考文献

- [1] Feng et al., *Poster presented. APCSLH*, 2017.
- [2] 比企, 音響誌, 66(10), 485-486, 2010.
- [3] Maekawa, *AST*, 46(1), 45-54, 2025.
- [4] 神戸大日中対照調音動態 MRI 動画コーパス(KUJC-MRI), 2023.
- [5] Mathis et al., *Nature Neuroscience*, 21(9), 1281-1289, 2018.
- [6] Takemoto et al. *Proc. INTERSPEECH*, 904-908, 2019.