

WAVEによる話速を変化させた時の舌運動の観測*

☆関純香, △下条未来, △武井俊輔, 竹本浩典 (千葉工大),
能田由紀子, 前川喜久雄 (国語研)

1 はじめに

話速を変化させると各音素の継続時間長(音素長)が変化し, 結果として単位時間あたりの音素数が変化する。一般に, 話速を大きくすると音素長は短くなり, 小さくすると長くなる。しかし, 各音素長は非線形に伸縮し, その変化率は一様ではないことが知られている[1]。

このような話速の変化によって生じる音素長の変化は, 音素環境だけでなく, 舌や口唇などの調音器官の形態や運動の個人差を反映していると考えられるが, はっきりしたことはわかっていない。そこで本研究では, 話速を変化させて朗読した際の音素長の変化と舌運動の変化を検討した。

2 材料と方法

2.1 実験参加者と録音方法

実験参加者(話者)は, 20代の日本人4名(男性2名:M1, M2, 女性2名:F1, F2)である。朗読文は「北風と太陽」(7文, 195モーラ, 音素数369)と「宇宙の果てにせまる」(野本陽代, 1998)より5文(235モーラ, 音素数361)を抜粋したものの2つを用いた。以下, それぞれ「北風」, 「宇宙」と呼称する。話者は, 自身にとってnormal(普通), fast(速い), slow(遅い)の3段階で朗読した。

音声は国立国語研究所の防音室内でコンデンサーマイクロホン(c-357)とリニアPCMレコーダー(PCM-D10)を用いてサンプリング周波数22kHz, 量子化ビット数16bitで録音した。

録音した音声にJulius-segmentation-kit[3]を用いて自動で音素ラベリングを行った。その後, CSJコーパス[4]を参考に, 音声分析ソフトPraat[5]を用いて手動で修正し, 補助ラベルを付加した。

2.2 話速の分析

CSJコーパスのラベリングでは, 母音ラベ

ル(5種), 子音ラベル(9種)以外に, 閉鎖音の前の無声区間(閉鎖), 声帯振動の継続, 共鳴の継続, 200ms以内の発話休止区間をポーズ, それ以上の発話休止区間をインターバルとし, ラベルを付与して区別している。本研究では, ポーズとインターバルの時間を合わせて「間」とした。そして, 「間」を含めた発話時間を総発話時間, 「間」を除いた発話時間を実発話時間とした。また, 閉鎖音の持続時間は除外した。

2.3 音素長の分析

音素ラベリングに基づいて各音素の継続時間長(単位:秒)を抽出し, 音素ごとに総和時間を求めた。そして, 話速の変化にともなう各音素長の変化量を求めた。

2.4 舌運動の計測

発話中の音声収録と同時に舌運動をNorthern Digital Inc.のWave speech research system(以下WAVE)[6,7]で計測した。話者の舌の正中矢状面に約1cmの間隔をあけて前からT1, T2, T3の3つのセンサを設置し(Fig.1), 各センサの3次元座標を100Hzサンプリングで記録した。

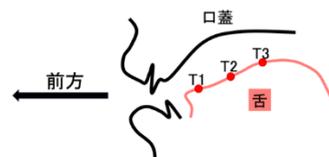


Fig. 1 WAVEのセンサ位置

3 結果と考察

3.1 話速の変化とWAVEセンサの影響

実験参加者の総発話時間における話速は, normal, fast, slowでそれぞれ444.1, 502.5, 386.7モーラ/分であった。これらはいずれもWAVEセンサを装着せずに話速を変化させた先行研究[8, 9]で得られた値の範囲内であった。またこれは, センサの装着が話速に与える影響は無視できることを示唆する。

* Observation of tongue movements while changing speech rate by using the WAVE system, by SEKI, Sumika, SHIMOJO, Mirai, TAKEI, Shunsuke, TAKEMOTO, Hironori (Chiba Institute of Technology), NOTA, Yukiko and MAEKAWA, Kikuo (NINJAL).

3.2 各センサの移動距離

Table 1 は各話者の3つのセンサの総移動距離と実発話時間である。normal より移動距離が減少した場合を青字，増加した場合を赤字で示す。全ての話者で話速にかかわらずセンサの総移動距離は T1, T2, T3 の順で大きかった。これは，舌先ほどよく動くことを示す。また，T1 に着目すると舌の移動速度は，normal, fast, slow ではそれぞれ，平均 9.3, 10.4, 8.9 cm/s であった。

センサの移動距離の変化パターンは話者，朗読文によって異なっていた。速く話すと移動距離が短く，遅く話すと長くなったのは M2, F1 の北風のみであった。速く話しても遅く話しても移動距離が大きくなる場合（M1 の北風，F2 の宇宙）もあれば，逆の傾向がみられる場合（M1, M2 の宇宙）もあった。

Table 1 各センサの総移動距離と実発話時間

北風	総移動距離(cm)			実発話時間(s)	宇宙	総移動距離(cm)			実発話時間(s)	
	T1	T2	T3			T1	T2	T3		
M1	fast	211	170	157	22	fast	205	160	150	23
	normal	171	142	125	22	normal	207	174	155	24
	slow	209	183	161	24	slow	199	160	156	27
M2	fast	260	141	129	21	fast	211	127	115	22
	normal	252	148	130	23	normal	226	139	126	24
	slow	298	173	157	26	slow	215	134	125	26
F1	fast	261	169	154	25	fast	264	166	154	25
	normal	270	184	168	26	normal	280	181	169	28
	slow	279	198	182	29	slow	274	194	181	30
F2	fast	223	168	114	22	fast	245	187	114	22
	normal	220	162	117	25	normal	212	160	108	25
	slow	227	182	124	27	slow	244	211	137	30

3.3 各センサの移動速度

Table 2 は normal に対して fast や slow では各センサの総移動距離と実発話時間がどの程度増減したかを比率 (%) で示したものである。例えば，f-n (s-n) の各センサの値は，fast (slow) と normal の総移動距離の差を normal における総移動距離で除して 100 倍した値である。もしセンサと実発話時間の増減率が等しければ，話速の変化によらずセンサの速度は一定であることを示す。全体的な傾向としては，速く話せば実発話時間と総移動距離はともに減少し，遅く話せば逆となった。また，実発話時間の増減率に比べて総移動距離の増減率は小さかった。これは，舌の移動距離を増減させるより移動速度を増減させることで話速を変化させていることを示す。しかし，個人差も朗読文による差も大きかった。

また，センサの総移動距離は T1, T2, T3 の順で大きかったが，話速の変化による移動距離の変化量は T1 よりも T2 や T3 で大きい傾向があった。なお，総移動距離の増加は明

瞭な発話，減少は不明瞭な発話となるか音声認識ソフト（Google Speech Recognition）による認識率を比較したが，差は見られなかった。

Table 2 各センサの総移動距離と実発話時間と変化量 (%)

北風		変化量(%)			実発話時間	宇宙		変化量(%)			実発話時間
		T1	T2	T3				T1	T2	T3	
M1	f-n	23.4	20.0	25.3	-2.6	M1	f-n	1.2	7.6	3.2	-6.4
	s-n	21.9	29.1	28.4	8.6		s-n	-4.0	-7.9	0.5	10.4
M2	f-n	3.1	-4.8	-1.0	-8.0	M2	f-n	6.9	8.3	8.6	-9.5
	s-n	18.1	17.2	21.0	14.6		s-n	-5.0	-3.7	-1.0	8.5
F1	f-n	3.2	8.3	7.9	-4.4	F1	f-n	5.6	8.6	9.4	-10.4
	s-n	3.3	7.3	8.4	11.4		s-n	-1.9	6.9	6.8	4.6
F2	f-n	1.3	3.5	-2.7	-8.8	F2	f-n	15.5	17.3	4.9	-11.6
	s-n	3.1	11.7	6.7	8.9		s-n	15.2	31.8	26.5	19.6

4 まとめ

本研究では，4 名の話者が朗読する速度を変化させたときの音素長と舌運動の変化を朗読文全体で検討した。その結果，話速の変化は，舌の動きの大きさ（調音空間の広さ）よりも舌を動かす速度（調音器官の速さ）を変化させて実現していること，その程度や方法には個人差が大きいことが明らかになった。しかし，今回得られた結果は全体的な傾向であり，特定のフレーズなど局所的には異なる可能性がある。

謝辞

本研究は JSPS 科研費 20H01265 の助成を受けた。また，本稿の一部は国立国語研究所の言語資源開発センターの助成を受け，共同利用型共同研究「話速の変化にともなう調音運動と音響特性の変化の観測」（研究代表者：竹本浩典）の研究成果である。

参考文献

- [1] 古井貞熙，計測と制御，688-693，1986.
- [2] 河津，前川，音講論（春），443-444，2009.
- [3] 李，情報処理学会研究報告，大語彙連続音声認識エンジン Julius ver.4，SLP-69-53,2007.
- [4] 前川他，『日本語話し言葉コーパス』の分節ラベリング Version 1.0』，『日本語話し言葉コーパス』附属電子文書，2007.
- [5] Boersma & Heuven, Praat, a system for doing phonetics by computer, Glot International 5(9/10): 341-345, 2001.
- [6] 北村他，信学技報，89-93，2014.
- [7] Katz et al., ISCA, 1174-1178, 2014.
- [8] 田村，佐藤，新潟医療福祉学会誌，18(1), 35-35, 2018.
- [9] 小澤他，広島県立保健福祉短期大学紀要，3(1), 95-102, 1998.
- [10] 立川他，音声研究 19 (3), 50-56, 2015.