

話速変化における音素長の変化の個人差*

☆関純香, △鍋島由輝, △武井俊輔, △下条未来, 竹本浩典 (千葉工大)

1 はじめに

話速は、1秒間あたりのモーラ数（モーラ毎秒）で表される。健常者の「普通」の話速とは平均 6~7 モーラ毎秒、「速い」話速とは平均 9~10 モーラ毎秒、「遅い」話速とは平均 5 モーラ弱毎秒程度とされており、個人差がある[1]。この個人差は、個人の脳機能や調音器官の運動機能に由来すると考えられるが、未解明である。

同じ文章を異なる話速で読み上げると、各音素長は非線形に伸縮する[3]。つまり、話速を変化させることで音素長の変化が大きい音素もあれば変化が小さい音素もある。また、挟母音/i/, /u/が前後を無声子音で挟まれた場合に母音の無声化が生じる[4]のように、音素環境によっても音素長は変化する。しかし、話速変化に伴う各音素長の伸縮パタンの個人差などについては知られていない。そこで本研究では、15人の実験参加者が本人にとって「普通」、「速い」、「遅い」の三段階の話速で文章を読み上げた音声を分析し、話速の変化に伴う音素長の変化などについて、基礎的な検討を行ったので報告する。

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）は日本人成人男性9名、女性6名である。話者は、自身にとって普通（normal）、速い（fast）、遅い（slow）の3段階で朗読文「北風と太陽」（7文、195モーラ）を読み上げた。録音は、千葉工業大学の無響室でSONY リニアPCMレコーダーPCM-D10（サンプリング周波数 48kHz, 24 bit）を用いて行った。

録音した音声にJulius-segmentation-kit [5]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[6]のラベリングを参考に、音声分析ソフト Praat[7]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、閉鎖音の前の無音区間（閉鎖）、声帯振動の継続、共鳴の継続、200 ms 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にラベルを付与し区別している。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。

2.3 音素長の分析

音素ラベリングに基づいて朗読音声の各音素の継続時間長（単位：秒）を抽出し、音素ごとの総和時間を求めた。そして、話速の変化に伴う各音素長の変化量を求めた。本研究では、普通に話すときに比べて速く話すときの時間変化（fast-normal）を f-n、遅く話すときの時間変化（slow-normal）を s-n で表す。

Table 1 上：f-n, 下：s-n（単位：s）

f-n	総発話時間	実発話時間	母音	閉鎖	その他	間
平均	-4.47	-2.35	-1.23	-0.39	-0.72	-2.12
標準偏差	1.55	1.04	0.74	0.29	0.55	1.06

S-n	総発話時間	実発話時間	母音	閉鎖	その他	間
平均	4.37	2.12	1.20	0.31	0.60	2.25
標準偏差	2.08	0.93	0.58	0.27	0.31	1.64

3 結果と考察

3.1 総発話時間・実発話時間の変化

朗読文の normal の総発話時間と実発話時間はそれぞれ平均 28.63 s, 22.15 s であった。Table 1 は f-n, s-n の発話時間や音素長などの変化の平均値を示す。ここで、音素長の変化を 2.2 節で上げたすべてのラベルについて検討したが、母音、閉鎖以外は、どれも平均 0.13 s 以下の変化しかなかったため、「その他」としてまとめた。

* Individual differences in phoneme length changes due to speech rate, by SEKI, Sumika, NABESHIMA, Yuki, TAKEI, Shunsuke, SHIMOJO, Mirai, and TAKEMOTO, Hironori (Chiba Institute of Technology).

まず、全体的な傾向として、総発話時間をはじめとする各変化量は速く話しても、遅く話しても同程度の短縮と伸長がみられた。しかし、各変化量の標準偏差の大きい値が示すように、ばらつきは非常に大きかった。Fig.1 はすべての話者の f-n, s-n の総発話時間、実発話時間の変化を示す。f-n の総発話時間の変化は F01 が最大で 7.39 s, M05 が最小で 1.78 s と大きいばらつきがあった。一方、s-n の総発話時間の変化は F03 が最大で 8.99 s, M05 が最小で 1.90 s と f-n よりもさらに大きいばらつきがあった。

さらに、同一話者でも、f-n と s-n の総発話時間の変化は非対称であった。例えば、F01 は、f-n は 7.39 s, s-n は 4.51 s と遅く話すより速く話すときの変化量が大きかった。逆に F03 は、f-n は 3.74 s, s-n は 8.99 s と速く話すより遅く話すときの変化量が大きかった。このような非対称性は多くの話者に見られた。

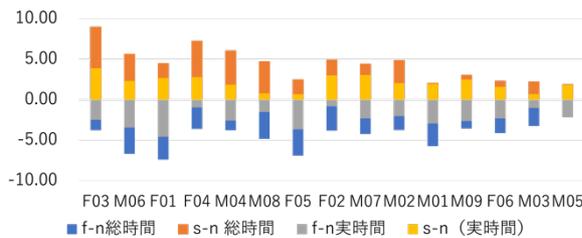


Fig. 1 総発話時間と実発話時間の変化

Table 2 f-n の各変化量の相関

	総発話時間	実発話時間	母音	閉鎖	その他	間
総発話時間	1.00					
実発話時間	0.74	1.00				
母音	0.66	0.74	1.00			
閉鎖	0.06	0.47	-0.01	1.00		
その他	0.50	0.68	0.07	0.40	1.00	
間	0.73	0.09	0.24	-0.39	0.05	1.00

Table 3 s-n の各変化量の相関

	総発話時間	実発話時間	母音	閉鎖	その他	間
総発話時間	1.00					
実発話時間	0.65	1.00				
母音	0.57	0.92	1.00			
閉鎖	0.27	0.52	0.26	1.00		
その他	0.63	0.81	0.66	0.21	1.00	
間	0.90	0.25	0.20	0.04	0.34	1.00

Table 2, 3 は総発話時間、実発話時間、母音、閉鎖、その他、間に対する相関分析の結果である。f-n, s-n に共通して、総発話時間は実発話時間、母音、間と高い相関が見られた。特に、s-n の総発話時間と間、実発話時間と母

音には非常に強い相関が見られた。相関分析の結果から、人は話速を速くする時には、主に母音、間を短縮しているのに対し、話速を遅くする際には、母音を伸長するよりも間を積極的に使い発話時間を伸長していると考えられる。

4 まとめ

本研究では、15名の「北風と太陽」の読み上げ音声を分析し、話速の変化にともなう音素長の変化の基礎的な検討を行った。その結果、母音、閉鎖、間に対してその他の音素の変化量は非常に小さいことが明らかになった。発話時間と各音素長を相関分析したところ、速く話すときは母音と間を利用して発話時間を短縮していたが、遅く話すときは速く話すときより間を積極的に使い発話時間を伸長していることが明らかになった。本稿では全体的な傾向について報告したが、今後、どのような音素環境のどの音素の継続時間長がどのように変化するか詳細に分析してゆく必要がある。

謝辞

本研究は JSPS 科研費 20H01265 の助成を受けて実施した。

参考文献

- [1] 鈴木誠史, 小倉俊彦, 話す速さの検出の基礎実験, 電波研究所季, pp.123-127, 1970
- [2] 遠山仁美, 松原茂樹, 同時通訳コーパスを用いた通訳者の発話速度の分析, 電子情報通信学会総合大会講演論文集, p.S87-S88, 2005.
- [3] 古井貞熙, 計測と制御, pp. 688-693, 1986.
- [4] 河津宏美, 前川喜久雄, 子音の調音様式が母音の無声化率におよぼす影響:「日本語話し言葉コーパス」の分析, 日本音響学会講演論文集, pp.443-444, 2009.
- [5] 李 晃伸, 大語彙連続音声認識エンジン Julius ver.4, 情報処理学会研究報告, SLP-69-53, 2007.
- [6] 前川喜久雄, 菊池英明, 藤本雅子, 米山聖子, 『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』付属電子文書, 2007.
- [7] Paul Boersma & Vincent van Heuven, Praat, a system for doing phonetics by computer, Glot International 5(9/10): 341-345, 2001.