

発話スタイルの違いによるVOTの性質*

○粕谷英樹（国際医療福祉大・保健医療），森大毅（宇都宮大・工）

1 はじめに

Lisker & Abramson [1]がいろいろな言語における閉鎖子音の音韻論的，音声学的，音響学および知覚的特徴として，有声開始時間（VOT, voice onset time）の重要性を指摘して以来，測定が容易なこともあって，VOTは多くの言語や健常/音声言語障害について調べられてきた．日本語についても，音韻論および音声学の観点から調査した Homma [2]，清水[3]の研究や音声言語障害について調べた Itoh らの研究[4]がある．しかし，これらは特定の単語を（有声/無声の最小対）キャリア文に埋め込んで発音したものであった．

有声/無声の対立や閉鎖子音の種類と VOT の関係については，孤立発音やキャリア文挿入という統制した実験下で，舌の運動や声帯振動の空気力学的な観点から説明されてきたが，多様なパラ言語情報を付与した発話における変化の様相については調べられていない．

そこで，本論文では，童話を感情を入れずに中立的に読んだ場合と，子供に読み聞かせることを想定して感情を豊かに入れて読んだ場合の VOT の変化の性質について調べた．

2 実験方法

2.1 音声資料

1名の成人女性（声優）が，童話「人魚姫」の一部を感情を込めずに読んだ音声（neutral）と児童に読み聞かせるように読んだ音声（emotive）を音声資料とした．

今回は閉鎖子音として，有声/無声破裂音だけを対象にした．無声破裂音/p, t, k/は語中でも VOT の測定に問題はないが，有声破裂音/b, d, g/の場合，語中では先行する音（多くの場合有声の母音）から破裂まで声帯振動が持続する（ボイスパー）ので，イントネーション句の先頭にある有声破裂音だけについて測定した．従って，有声破裂音の標本数は非常に少ない．分析した破裂音の標本数を neutral

と emotive それぞれについて，Table 1 に示す．

2.2 測定方法

PC のディスプレイ上に表示した音声波形とデジタル・サウンドスペクトログラムを参考にして，破裂時点と声帯振動の開始時点の時間間隔を測定した．実際の測定では，L-Voice を用いた[5]．

Table 1 Number of samples analyzed for VOT.

stops	Neutral	Emotive
	No. samples	No. samples
/b/	5	5
/p/	12	11
/d/	16	16
/t/	409	395
/g/	5	5
/k/	353	349

3 結果

測定結果を Table 2 に示す．表中，AV, SD はそれぞれ，破裂音ごとの測定値の平均値（ms）および標準偏差（ms）である．括弧内の数字は清水の測定結果[3]である．

Table 2 The average values of VOT of the Japanese stop consonants in running speech read by a female in two different speaking styles, neutral and emotive. The figures in parentheses are from Shimizu [3].

stops	Neutral		Emotive	
	AV (ms)	SD (ms)	AV (ms)	SD (ms)
/b/	-31.4 (-89)	10.5 (28.5)	-25.6	22.8
/p/	12.1 (41)	3.8 (17.1)	17.8	7.2
/d/	-31.9 (-75)	16.5 (32.7)	-42.1	24.3
/t/	14.5 (30)	3.9 (12.7)	17.4	6.7
/g/	-14.2 (-75)	32.3 (27.0)	-42.4	30.6
/k/	35.5 (66)	13.2 (12.1)	46.3	18.4

* Characteristics of voice onset time (VOT) of Japanese stop consonants in running speech read in different speaking styles, by KASUYA, Hideki (International Univ. of Health and Welfare), MORI, Hiroki (Utsunomiya Univ.)

また, *neutral* と *emotive* の発話スタイルでの有声破裂音 (赤色) と無声破裂音 (青) の VOT の分布図を, それぞれ Fig. 1, Fig. 2 に示す. 図では, 有声破裂音の標本数が少ないので, 頻度は 10 倍にして表示し, 見やすくしてあるので, 注意されたい.

4 考察

Table 2 で, 本研究で得られた測定値と清水の測定値の平均値 (括弧で示した値) [3] を比べると, 無声破裂音の VOT では清水の値が 2~3 倍ほど長く, 有声破裂音でも負側に 2 倍以上長いなど, 大きな違いがある. これは, 清水らが “これは... です” というキャリア文の中に “場 (ば) 一羽 (ば)” というような 1 音節語 (最小対) を入れて発音したため, 焦点が置かれ, 特徴が一層強調されたためと推察される. また, 標準偏差も清水の方が本研究で得られた値よりかなり大きい, これは, 本研究では 1 名の話者であるのに対して, 清水の研究では男女各 3 名であり, 個人差が反映している考えられる.

われわれの測定で, /p, t/ に比べて軟口蓋音 /k/ の VOT が 2 倍以上長いという性質は, *neutral/emotive* を問わず一貫している. /p, t/ の VOT は *neutral* と *emotive* でそれほど変わらないが, /k/ では *emotive* の方が 10 ms 以上長くなる. 例えば, “深い深い底に” という発話には, 3 つの軟口蓋音 /k/ がある (それらを出現順に k1, k2, k3 で表す). *emotive* の発話スタイルでは, この部分は感情を込めて発話するため, 母音は長く, 破裂後の帯気部も長く制御している. *neutral* と *emotive* の差は $\Delta k1=7$ ms, $\Delta k2=6$ ms, $\Delta k3=20$ ms で, いずれも *emotive* で長くなる. k3 (/soko/) では 20 ms も長いということは, 話者が感情を込めるとき, 明らかに帯気部を長くするような制御を行っているからであると推察される. また *emotive* の方が一貫して標準偏差が大きい. このことは, Fig.1, 2 を見ても明らかである.

有声破裂音 /b, d/ では *neutral* と *emotive* で 5~10 ms の違いしかないが, /g/ では 28 ms も *emotive* の方が負の方向に長い.

有声破裂音の標本数が余り多くないので, 断定的にはいえないが, 無声/有声を問わず, 軟口蓋破裂音にパラ言語の影響がより強く反映される傾向がある.

5 おわりに

今後 VOT を有声/無声の弁別特徴としての役割だけでなく, パラ言語情報を付与する上で担っている性質についても, もっと詳細に検討したい.

謝辞

音声資料をご提供頂いた釜井孝浩氏 (パナソニック (株) 先端技術研究所), 実験に協力して頂いたストドニツキー・ヤン氏, 文献[2] をご教示頂いた前川喜久雄氏 (国立国語研究所) に感謝します. 本研究の一部は, 科研費 (19500155, 20500163) によった.

参考文献

- [1] Lisker, L & Abramson, A.S., *Word*, **20**, pp.1-28, 1964
- [2] Homma, Y, *Bulletin of the phonetic Society of Jpn.*, **163**, 7-9, 1980.
- [3] 清水, *音声研究*, **3**, pp.4-10, 1999.
- [4] Itoh, M., et al, *Ann. Bull. RILLP*, **13**, pp.123-132, 1979.
- [5] Zhu, H., Kasuya, H., *JASJ (E)*, **19**, pp.223-230, 1998.

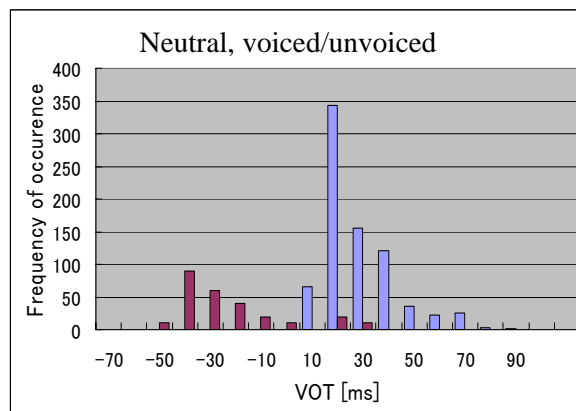


Fig. 1 VOT distributions of voiced (red) and unvoiced (blue) stops, where the frequency of occurrence of the voiced is multiplied by ten to improve the visibility.

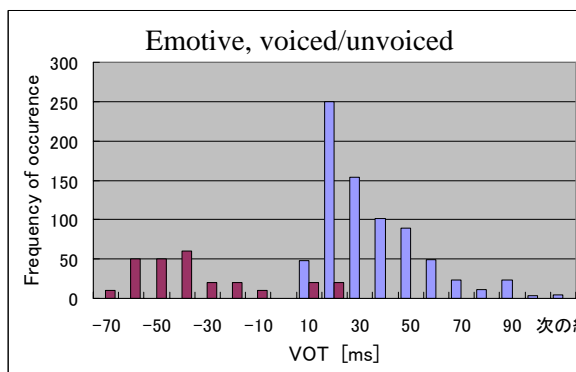


Fig. 2. Same as Fig. 1 except that speech samples are from those read in a emotive style.