

前川喜久雄, 籠宮隆之 (国立国語研究所)
 菅田雅彰, 鑄木時彦, 岡留 剛 (NTTコミュニケーション科学基礎研究所)

1 目的

音声によるパラ言語情報の伝達メカニズムの解明は今後の音声研究の重要なテーマである [1, 2]. 従来, パラ言語情報の伝達には音声のいわゆる韻律的特徴が重要な役割を果たすことが強調されてきている. 韻律的特徴の重要性は明らかであるが, いくつかの先行研究ではパラ言語情報の影響はいわゆる分節的特徴にも及んでいることが確認されている [3-5]. 我々は文献 [4, 5] で, 同一のテキストを6種類のパラ言語的意味で発話しわけた資料に基づいて母音のフォルマント構造にも変動が観察されることを報告した. 特に「疑い」と「感心」の発話の間には, 第2フォルマント周波数(F2)が「疑い」では高く「感心」では低いという一貫した関係が見出されている.

今回の報告では, 上述の差異がどのような調音運動によって生じているかを調音運動の実測データに基づいて検討する.

2 データ

調音運動の測定には NTT コミュニケーション科学基礎研究所に設置された交流磁気位置検出装置 (EMMA) を用いた. 上唇 (UL), 下唇 (LL), 舌尖 (T1), 舌端 (T2), 中舌面 (T3), 奥舌面 (T4), 下顎門歯 (MN) の7点にコイルを装着し, 毎秒250回のサンプリング周波数で正中断面上のコイルの位置変化を記録した (図1). 音声は16bit, 48KHzで録音し, 分析時に16KHzにダウンサンプリングした.

利用したテキストは「駅/eki/」「秋/aki/」「鮭/sake/」「佐賀/saga/」「笹田が/sasadaga/」の4単語1文であり, 指定したパラ言語的意味は「N: 中立=棒読み」「A: 感心」「S: 疑い」「D: 落胆」の4種類である (以下では上記のアルファベット略号でパラ言語情報に言及することがある).

被験者は5種類のテキストを指定された4種類の意味で16回繰り返して発話した. ただし測定途中でコイルを再装着したので, 以下の分析ではコイル位置に変化のない10回分を分析対象とする.

3 分析

最初にフォルマント周波数の分布を検討しよう. 図2に「笹田が」に含まれる四つの母音/a/のフォルマント平面上での分布を示した. F1, F2は次数18~22のLPC分析の結果とDFTスペクトルとを手作業で比較しながら推定した (図2ではF1, F2とも軸を逆転してある). 図2では左側 (高F2) に「S: 疑い」が右側 (低F2) に「A: 感心」が分離して分布し, その中間に「N: 中立」と「D: 落胆」が位置している. 先行研究と一致する結果である. また「駅」に含まれる母音/e/, /i/の場合

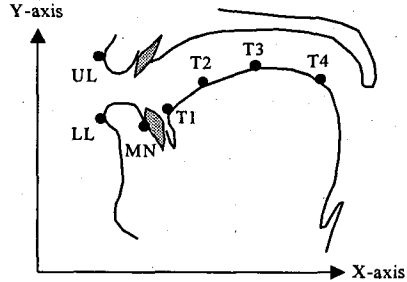


Figure 1. Placement of coils on the surface of tongue and lips. The origin of axes is on the lower left corner.

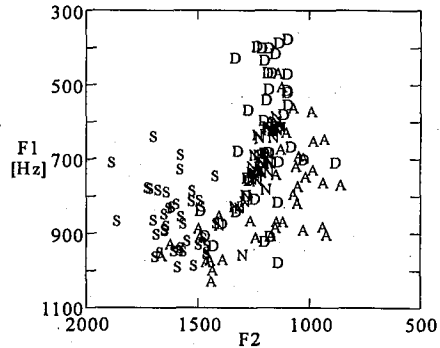


Figure 2. Scatter plot of the four vowels in /sasadaga/ uttered under four paralinguistic meanings: Letters stand for 'Admiration', 'Suspicion', 'Disappointment', and 'Neutral'.

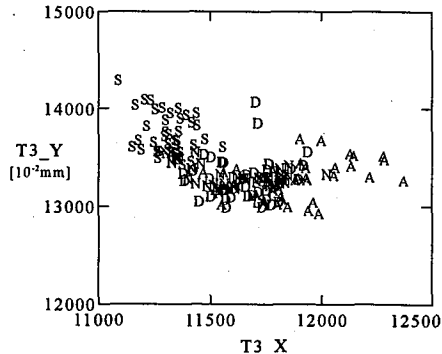


Figure 3. Positioning of the T3 coil measured at the same time point of the formant frequency data in figure 1. T3_X and T3_Y are the X- and Y-component of coil position.

* Production of paralinguistic information: From an articulatory point of view.
 Kikuo Maekawa, Takayuki Kagomiya (Nat. Lang. Res. Inst.), Masaaki Honda, Tokihiko Kaburagi, and Takesi Okadome (NTT)
 Email contact: kikuo@kokken.go.jp

合も、AとSの関係は同一であった[6].

次いで図3にフォルマント測定時点におけるT3コイルの位置の分布を示す。ここでもまたSは左側(T3_Xが小さい=前寄り)、Aは右側(T3_Xが大きい=後寄り)に分布している。T3_XとF2の積率相関係数を計算すると-0.826となる。四つの母音を個別に分析しても同一の傾向が見出されるが、AとSとの分布間距離は発話の末尾に向かうにつれて増大する傾向がある[6].

図3に観察された調音の前後方向の組織的変動は母音だけに生じているのではない。図4は波形から決定した分節音区間毎にその区間に属する測定値の平均値と標準誤差を計算した結果である。子音区間においても、AとSの差は明瞭に保たれていることがわかる。同一のパタンがT3とT1に観察されることから、AとSとは舌全体が前後に移動していると推測される。

F2の上昇/下降に顕著な影響を及ぼす調音パラメタには、舌の前後位置の他に唇の開口面積がある[7]。図5には図4と同じ様式で上唇(UL_Y)と下唇(LL_Y)のY軸成分の差を示した。発話冒頭の子音/s1/を唯一の例外として、Sの発話は常にAよりも唇間距離が大きいことがわかる。唇の開口面積とフォルマント周波数の間には正の相関が存在することを考えると、AとSとのF2の相違には唇の制御も貢献していると推定できる。

4. まとめ

パラ言語情報生成時の調音運動を観測することによって、先行研究において発見された「A:感心」と「S:疑い」との間の母音フォルマント周波数の差異が、舌の前後方向位置および唇間距離の異同によって生じていることを確認した。また、この差異は母音だけでなく、子音を含む発話の全体におよんでいることを確認した。

今後の課題として、まず被験者の増員、今回報告した分節的特徴の変動と韻律的特徴の変動との関係の解明、音声生成モデルにおけるパラ言語情報の位置づけ、などが挙げられる。

謝辞 本研究にはCREST/JSTおよび文部省科学研究費基盤B(No.10410110)の支援を受けました。

文献

- [1] 藤崎博也「韻律研究の諸側面とその課題」日本音響学会講演論文集, 287-290, 1994.
- [2] 前川喜久雄「音声学」講座言語の科学2『音声』, 岩波書店, 1998.
- [3] Abe, M. "Speaking styles: Statistical analysis and synthesis by a Text-to-Speech system." van Santen et al. eds. *Progress in Speech Synthesis*. Springer, 1996.
- [4] Maekawa, K. "Phonetic and phonological characteristics of paralinguistic information in spoken Japanese", *Proc. ICSLP98*(CD-ROM), Paper #0997, 1998.
- [5] 前川喜久雄・北川智利「パラ言語情報の生成と知覚」電子情報通信学会技術研究報告, SP99-10, 1999.
- [6] 前川 他「パラ言語情報の調音運動への影響」日本音響学会平成11年度全国大会予稿集, 1999.
- [7] Fant, G. *Acoustic Theory of Speech Production*. Mouton, 1960.

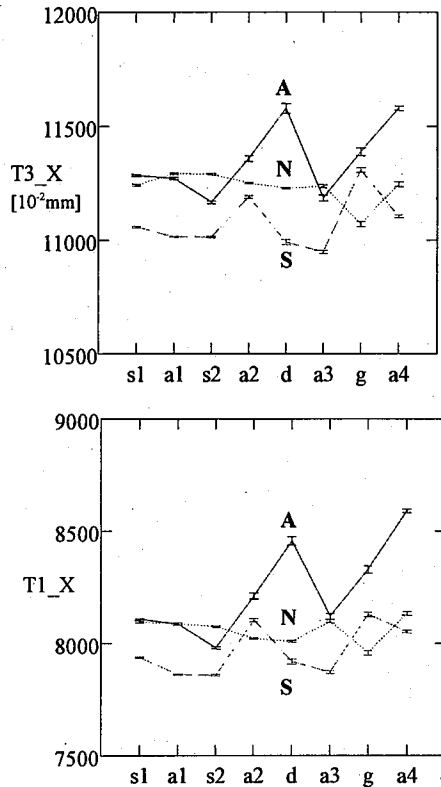


Figure 4. Mean value of the x-component of T3 (top) and T1 (bottom) coils as a function of segment and paralinguistic information type. Short vertical lines show standard errors for each segment. Letters 'A', 'N', 'S' stand respectively for 'Admiration', 'Neutral', and 'Suspicion'.

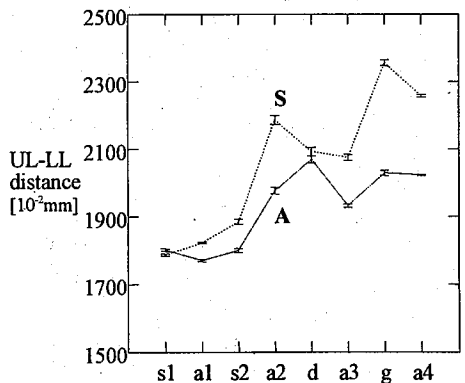


Figure 5. Distance between the UL (upper lip) and LL (lower lip) coils as a function of segment and paralinguistic information type. 'A' and 'S' stand respectively for 'admiration' and 'suspicion.'