

# パラ言語情報の調音運動への影響—「疑い」と「感心」の比較—

前川 喜久雄(国立国語研究所言語行動研究部)

龍宮 隆之(国立国語研究所非常勤研究員)

菅田 雅彰(NTTコミュニケーション科学基礎研究所)

鈴木 時彦(NTTコミュニケーション科学基礎研究所)

岡留 剛(NTTコミュニケーション科学基礎研究所)

## 1 目的

音声は言語情報の他にパラ言語情報をも伝達している。音韻情報や統語情報などの言語情報は離散的記号の組み合わせによって伝達される情報である。これに対して発話の意図や話者の心的態度などのパラ言語情報は基本的には範疇的な対立を構成していると考えられるものの、単一カテゴリーの内部で量的な変化が可能である点で言語情報と相違している[1,2]。

我々は近年パラ言語情報の伝達に関与する音声特徴についての検討を音響的側面と知覚の側面から行ってきた[3,4]。その結果、パラ言語情報の影響はピッチ、パワー、分節音持続時間長などの韻律的特徴のほか、喉頭原音の特性や母音のスペクトル特徴にも組織的な影響が観察されることが確認されている。

文献[3,4]に報告した研究では、「山田さんですか」などのテキストを話者3名が6種類のパラ言語的意味で発声しわけた資料の音響的分析の一環として母音のフォルマント周波数を測定した。

その結果、「感心」と「疑い」の発話を同一テキストの同一母音において比較すると、第2フォルマント周波数(F2)に組織的な差異が観察され、「感心」の発話のF2は「疑い」よりも統計的に有意に低い(あるいは「疑い」のF2が有意に高い)ことが確認された。この傾向はフォルマント分析を実施した男性話者2名に共通して観察された。

またフォルマント平面上において「感心」と「疑い」の発話における母音の重心間距離を計算すると、発話の冒頭と末尾では音響距離が大きく、中間部では音響距離が減少する傾向も観察された。

このような音響上の差異は音声生成過程上、どのような差異に起因しているのだろうか。先行研究では、調音運動に対する内省に立脚

して「感心」の発話では前よりに、また「疑い」の発話では後よりに母音の調音が変化しているものと推測した。以下本稿では調音運動の観測データを利用して、この仮説を定量的に検討する。

## 2 データ

調音運動の観測には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。同装置の原理は文献[5]に詳しいが、被験者の頭部に発生させた交流磁界の強度を調音器官の表面に装着した微少なコイルによって検出することによって、正中断面上における調音器官の運動軌跡を記録する装置である。今回の実験では、上唇(UL)、下唇(LL)、下顎門歯(MN)、舌尖(T1)、舌端(T2)、中舌面(T3)、奥舌面(T4)の7箇所のコイルを装着して各コイルの位置を測定した。図1はコイルの装着位置を模式的に示したものである。

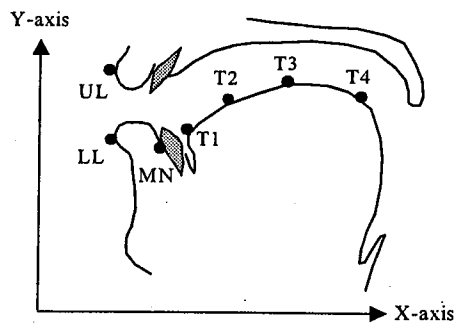


Figure 1. Placement of coils on the surface of tongue and lips. The origin of axes is on the lower left corner.

被験者は男性1名。テキストは「駅/eki/」「秋/aki/」「鮭/sake/」「佐賀/saga/」「笹田/sasadaga/」の4単語1文である。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種類であり、先行研

究で指定したパラ言語的意味のサブセットになっている（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種類の意味で繰り返し発話した。被験者は1枚につき一種のテキストとパラ言語的意味が指定された20枚のカードを目前において、カードをめくりながら指定された発話をおこなった。繰り返しは16回行なったが、測定途中でコイルを再装着したので、以下の分析ではそのうち連続した10回分だけを分析の対象にする。各コイルの運動軌跡はサンプリング周波数250Hzで記録し音声は48KHz(分析時は16KHzにダウンサンプリング)、16bitで記録した。

### 3 分析

以下では最初に音響的特徴を概観し、続いて調音運動の特徴を分析する。

#### 3.1 音響分析

図2にテキスト「笹田が」を例として、パラ言語情報の音声基本周波数(F<sub>0</sub>)への影響を示す。Nと比較すると、A、Sではピッチレンジが拡大し、Dでは反対に縮小する。Aでは発話の後半においてF<sub>0</sub>が緩やかに下降して発話末に至るのに対して、Sでは発話冒頭から上昇が持続し、最終モーラで急激に上昇する。これらの特徴は先行研究において指摘した特徴と一致している。

図3には持続時間長への影響を示した。テキストを構成する四つのモーラの平均持続時間長の比率をNの平均を1.0とおいて表示している。どのモーラにおいても、A、D、Sの持続時間長はNよりも顕著に長い。発話冒頭モーラにおけるA、D、Sの延長率は約2.0であるが、この比率は第3モーラまで徐々に低減していった後、最終モーラで大幅に増大する。これらもまた先行研究と一致する結果である。

図4にはテキストに含まれる四つの母音/a/の第1、第2フォルマント周波数(F<sub>1</sub>, F<sub>2</sub>)の分布を「A感心」と「S疑い」について示した(X軸、Y軸ともに逆転してあることに注意。また参考のために「N中立」と「D落胆」の平均値もプロットしている)。いずれの母音においても、「A感心」と「S疑い」の間には明瞭な分布の差が認められ、先行研究と一致する。ただし、図4

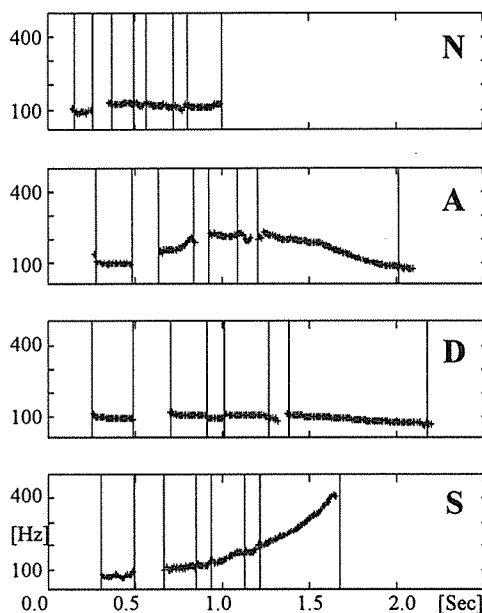


Figure 2 Typical pitch contours of /sasadaga/ under four different paralinguistic renditions: Neutral, Admiration, Disappointment, and Suspicion, from top to bottom. Vertical lines denote phoneme boundaries.

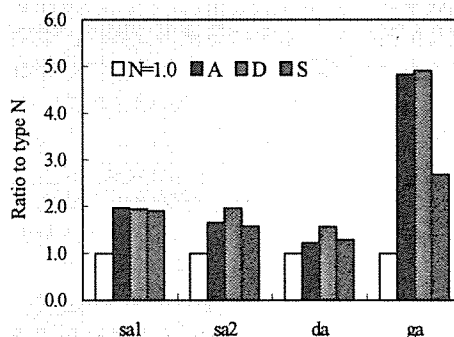


Figure 3 Mean duration of the four morae in /sasadaga/ as a function of paralinguistic information types. The mean of neutral utterances was set to 1.0.

では、AとSの分布間の距離が発話の後に位置する母音ほど増大する傾向にあるが、先行研究では発話冒頭モーラにおける距離がやや大きく、第2モーラ以降に減少し、最終モーラで大幅に増大するという（持続時間長と同一の）パターンを示しており、この点では先行研究結果との相違が認められた。

### 3. 2 調音運動の分析

図4におけるAとSの分布の相違は、主としてF2の相違である。以下ではこの点に留意して調音運動との関係を検討する。F2に顕著な影響を与える調音運動の候補としては、1) 調音位置の前後方向の移動、と、2) 唇の状態(開口面積)のふたつが考えられる。調音位置が前よりに移動すればF2は上昇し、後退すれば低下する。一方、開口面積とF2の間には、正の相関が存在することが知られている[6]。

図5に、図4に示したフォルマント周波数の計測時点におけるコイルT3の位置をプロットした。T3\_XがX軸成分、T3\_YがY軸成分である。計測時点としては母音の調音のために下顎がおこなう上下運動のピーク時点、すなわち下顎(図1のコイルMN)が局所的にもっとも下降した時点を選択した。この時点は音響的には母音の定常部分に該当する。

図5ではAが後寄りに(すなわちX軸の右寄りに)Sが前寄りに(同じく左寄りに)分布しており、調音運動においてもAとSは組織的に相違していることがわかる。また図4で観察されたF2の分布との関係も先述の予想に従ったものとなっている。

ちなみに図4におけるF2値と図5におけるT3\_Xの積率相関係数を計算すると-0.826であり、強い負の相関が認められる。T3\_XからF2への回帰分析をおこなうと、 $P < .001$ で有意との結果を得る。

以上は調音運動にもパラ言語情報の影響が及ぶことを明らかに示す結果であるのだが、この段階で検討しておくべき問題が二つあると思われる。ひとつは、このような調音運動の変動が母音/a/に限って観察されるのか、それとも他の母音にも広く生じているのかという問題である。

この問題を検討するためにテキスト「駅」について上と同様の分析を施してみた。図6にフォルマント周波数を、図7にコイルT2の位置を示す。これらの図からは /e/, /i/ の場合にも/a/と同様の変化が生じていることがわかる。

もうひとつの問題は、図4と図7に観察される調音上の変化が母音にだけ生じているのか、それとも子音にも及んでいるのかという

問題である。さらにこれと関連して、T3以外のコイルにも同様の影響が認められるのかという問題も検討する必要があるだろう。これらの問題を検討するために図8を作成した。この図は、テキスト「笹田が」について、サウンドスペクトログラムの視察によって決定した各分節音区間に属するT1とT3のすべての測定値のX成分の平均と標準誤差を折線グラフの形で示したものである。図には比較のために「N中立」の場合も示してある。

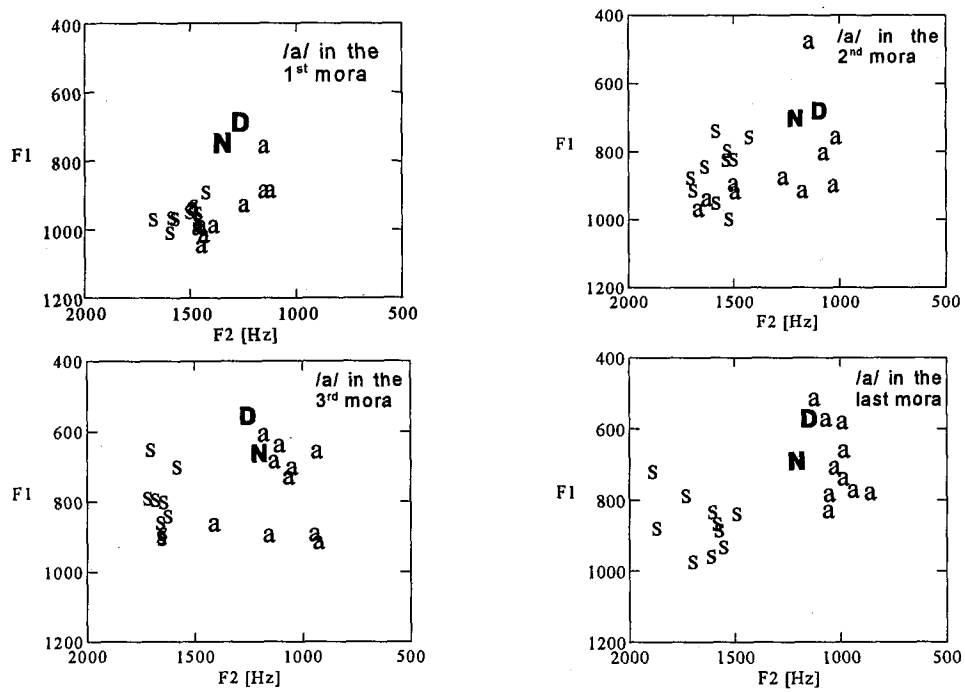
図8には母音区間と同様、子音区間(/s1/, /s2/, /d/, /g/)においてもAの測定値の方がSよりも高いX成分値をとること、つまり調音が相対的に後寄りであることが示されている。さらに、上述の傾向はT1にもT3にも一貫して観察されており、舌全体が前後に位置を変動させていることがわかる。「駅」の分析結果にも同一の傾向があらわれていた。

最後に唇の運動を検討する。図9に「笹田が」における下唇位置のX成分(LL\_X)を図8と同一の方法で示す。この図で注目されるのは、AとSの関係が図8とは逆転しており「A感心」は小さな値を「S疑い」は大きな値をとっていることである。この傾向は発話の冒頭ではほとんど観察されないが、末尾に向かうにつれて顕在化している。

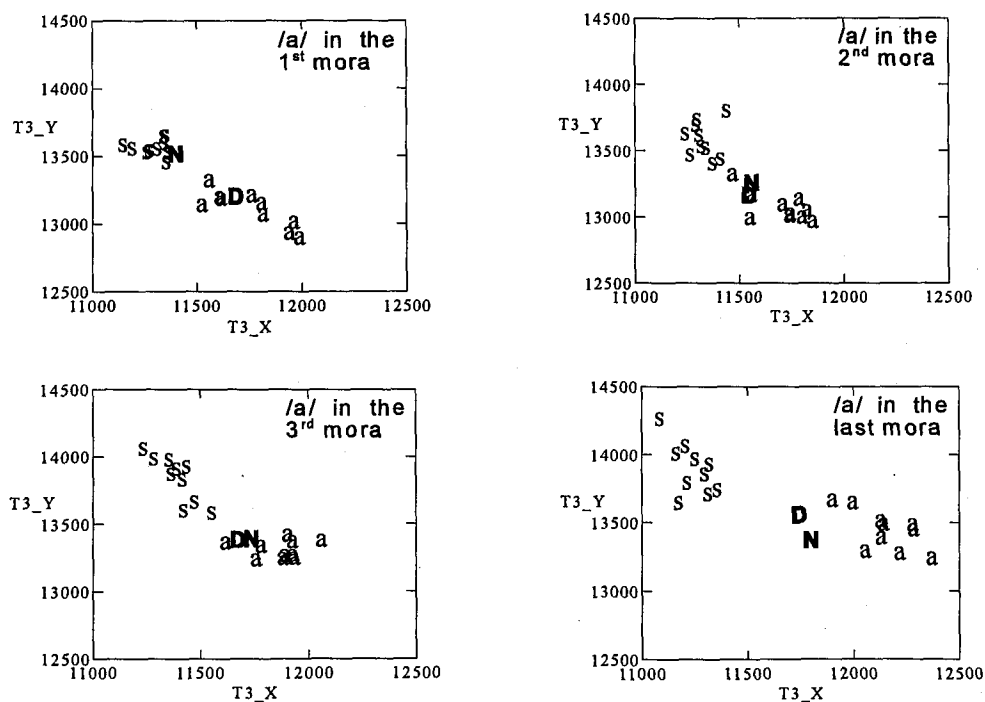
図9にあらわれた唇の位置変化は、唇における開口面積の調整のための変化であると推測される。つまり、Aの発話では上下の唇を接近させて開口面積を低下させることによってF2を低下させ、Sでは反対に開口面積を増大させてF2を上昇させるという調整である。

図10は「笹田が」における上唇コイルと下唇コイルのY軸成分の差の平均値の推移を示している。Aにおけるコイル間距離はSよりも顕著に小さく、上記の仮説の妥当性を示唆する結果であると言える。

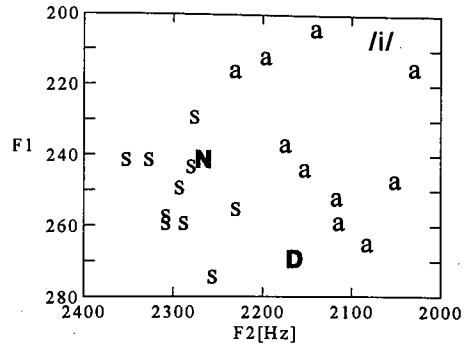
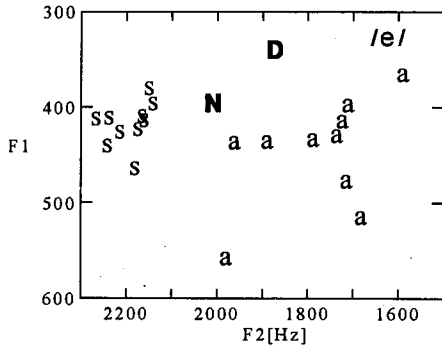
以上をまとめよう。先行研究および今回の音響分析によって存在が確認された「A感心」と「S疑い」の間のF2の差異は組織的な調音運動の差異によって生みだされている。Aでは前寄りにSでは後寄りに舌全体が移動してF2を上昇/下降させる。またこれとは独立に唇による開口面積の調整も存在していると推測される。



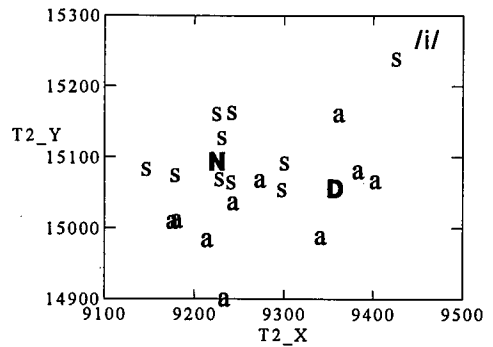
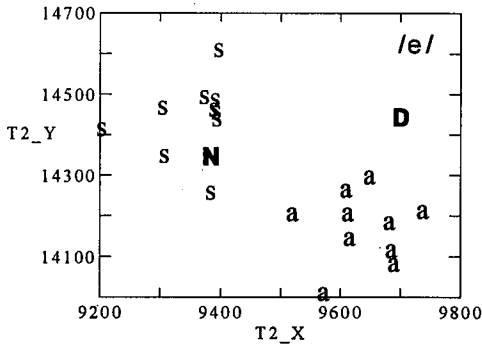
**Figure 4.** Distributions of the four vowels of /asadaga/ on F1-F2 formant plane. The letters 'a' and 's' denote respectively the samples uttered in 'admiration' and 'suspicion' utterances. The capital letters 'D' and 'N' denote the mean values of 'disappointment' and 'neutral'.



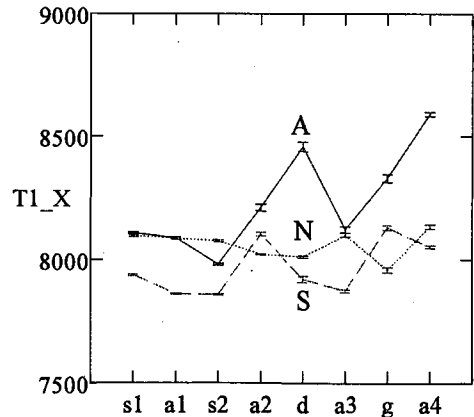
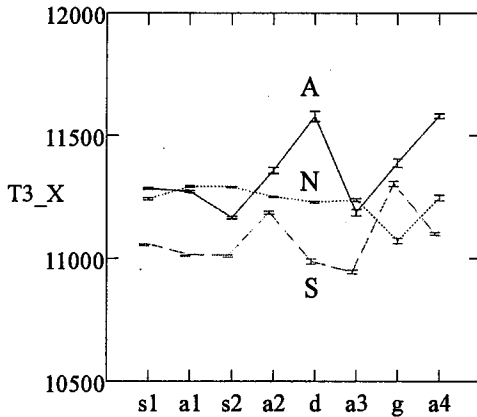
**Figure 5.** Distribution of the position of the T3 coil measured at the same time-point as in figure 4. Unit is  $10^{-2}$  mm. Capital letters denote the mean values of 'Neutral' and 'Disappointment'.



**Figure 6.** Distributions of the vowels in /eki/ on the F1-F2 formant plane. The letters 'a' and 's' denote respectively the samples uttered in 'admiration' and 'suspicion' utterances. Capital letters denote the mean values of 'Neutral' and 'Disappointment' utterances.



**Figure 7.** Distribution of the position of the T2 coil measured at the same time-point as in figure 6. Unit is  $10^{-2}$  mm. Capital letters denote the mean values of 'Neutral' and 'Disappointment'.



**Figure 8.** Mean value of the x-component of T3 (left) and T1 (right) coils as a function of segments and paralinguistic information types. Abb. 'A', 'N', 'S' stand respectively for 'Admiration', 'Neutral', and 'Suspicion'.

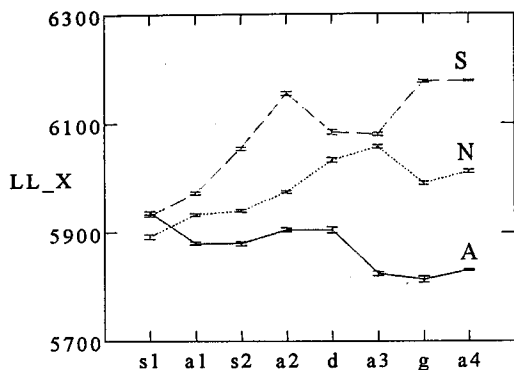


Figure 9. Mean value of the x-component of the LL (lower lip) coil as a function of segment and paralinguistic information type. Text was /sasadaga/.

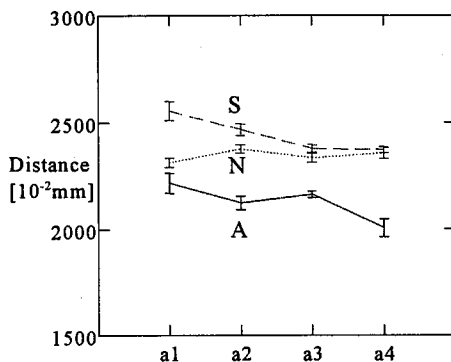


Figure 10. Mean distance between the upper lip (UL) and lower lip (LL) coils as a function of segment and paralinguistic information type. Text was /sasadaga/.

#### 4. 議論

今回の実験によってパラ言語情報の生成に際しては、従来から指摘されてきている韻律要素だけでなく、調音運動もまた組織的に変化していることが明らかとなった。

今回の実験によって存在を確認することができた舌全体の前後方向への移動を従来の調音音声学の枠組みに位置づけるならば、いわゆる tone of voice (もしくは voice-quality) の現象とみなすのが妥当である。

Laverは調音音声学の記述能力を拡張する試みの一貫として分節音の範囲を超えた大きな時間区分において定まる音声器官の位置的特徴を“setting”と呼び、その分類の枠組みを与えている[7]。これに従うならば、「S疑い」の発話は舌(tongue body)の前方向への偏移をあらわす“advanced body”という特徴によって特徴づけられ、反対に「A感心」は、後方への偏移をあらわす“retracted body”によって特徴づけられると記述できるであろう。

さらに今回発見された唇の調整も tone of voice のためのひとつの setting として把握するのが妥当であろう。Laverの分類には“lip-spread”という特徴があり、定性的な記述にはこれを利用すべきかと思われる。

#### 5. 今後の課題

従来、tone of voice に関する研究は音声研究の周辺の課題と認識されてきており、殊に日本語に関しては、ほとんど研究がおこなわれてきていない。しかし、本稿冒頭にも述べたように、パラ言語情報は話者が意図的に制御

して表出する情報であるから、その制御メカニズムの解明は音声生成過程の研究にとって欠かすことのできない研究である。

今後の課題としては、被験者を増やすことが何よりも必要である。分析手法に関する課題としては、各コイル間の相互作用の検討が必要である。特に下顎の運動は舌全体に影響をおよぼす可能性があるため、舌面上のコイルと下顎コイルの関係について検討する必要がある。また、今回の報告から除外したY軸成分についての解析結果も別の機会に報告する予定である。

謝辞 本研究の遂行には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。

#### 参考文献

- [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] Maekawa, K. “Phonetic and phonological characteristics of paralinguistic information in spoken Japanese”, *Proc. ICSLP98* (CD-ROM), Paper #0997, 1998.
- [4] 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.
- [5] 菅田雅彰「調音モデルに基づく音声合成」音声研究, 2-2, 1998.
- [6] Fant, G. *Acoustic Theory of Speech Production*. Mouton, 1960.
- [7] Laver, J. *Principles of Phonetics*. Cambridge Univ. Press, 1994.