

音響音声学の展望 Trends in Acoustic Phonetics

前川 喜久雄*
Kikuo MAEKAWA

1 はじめに

本稿では、音響音声学という用語をかなり広い意味で用いている。音響的分析手法を利用した言語研究と分析手法そのものの研究とでは、研究の方法論も研究者の所属もちがっているのが普通である。大学に例をとれば、前者は文学部、後者は工学部の研究になるだろう。

しかし、文学部的な研究と工学部的な研究とを範疇的に峻別してしまえば、音声研究の本質を見失うおそれがあるので、以下では両者に触れる。ただし紙幅の制限で、いずれについても系統的な記述をおこなうことは不可能である。分析手法については Borden & Harris (1980), Kent & Read (1992), 日本音声言語医学会 (1994), 古井 (1992), 城戸 (1985) などの概説書を参照していただきたい。後に挙げたものほど数学的記述が多用されている。

以下では「分節音」と「韻律特徴」に大別して筆を進める。分節音に関しては分析手法に重点を置き、韻律特徴に関しては言語学的研究に重点を置く。後者では、特にイントネーション研究に重点を置いているが、これは編集委員会からの依頼でもあり、また昨今の研究動向を反映するものでもある。

2 分節音

2.1 母音

現代的な母音の生成理論の嚆矢となったのは Chiba & Kajiyama (1941) である。彼らは X 線写真とパラトグラムを利用して声道の断面積を測定し、声道の共鳴特性を電気的等価回路に置換して研究する途を拓いた。この礎のうえに発展し 1960 年前後に Fant (1960) 等において集大成されたのが現代の音響学的音声生成理論である。その根幹をなす

音源フィルター理論 (source-filter theory) は声帯の振動により生成される喉頭音源のスペクトル S が声道の伝達関数 T による影響を被って音声スペクトル W として生成される過程を S と T が相互に独立という仮定のもとに $S \cdot T = W$ と定式化する。 T のパラメータには声道の固有共鳴としてのフォルマントおよび反共鳴としてのアンチフォルマントがある。これらは概略スペクトルの山および谷に相当すると理解してよいが、それぞれ共鳴周波数およびバンド幅を与えることで記述される。音声生成モデルのうちフォルマントのみを考慮するものを全極モデル、アンチフォルマントも考慮するものを極零モデルと呼ぶ。

母音のスペクトルのうち音韻に關与する情報の形成に重要なのは T であるので声道伝達関数の推定方法については多くの方法が提案されている。サウンドスペクトログラムの視察による方法に始まり、音源フィルター理論に沿って各種パラメータを最適化手法により推定する Analysis-by-Synthesis 法、音源と声道のスペクトル特性を分離するのに適したケプストラム法、時系列データのスペクトル分解の理論に基礎をおく線型予測分析法などがある。現在では計算量の少ない線型予測分析法が広く普及しているが、これは全極モデルに分類される。

Peterson & Barney (1952) は英語母音のフォルマントを大量に分析した研究として有名である。日本語母音のフォルマント周波数の測定例としては粕谷他 (1968) がある。ただしいずれもサウンドスペクトログラムによる測定であり、現在の水準からすれば測定精度が高いとはいえない。

母音の生成に際してアンチフォルマントを考慮する必要がある典型的なケースは鼻母音ないし鼻音化母音である。これらの母音では鼻腔の接続により特定の周波数領域の音響エネルギーが鼻腔

* 国立国語研究所 (The National Language Research Institute)

に吸収されることによって鼻音性が生じる (Hattori et al., 1958)。アンチフォルマントを含む母音のスペクトルの推定には現在でも困難が大きい。

2.2 子音

音源フィルター理論の枠組みで子音の生成を説明すると、音源が喉頭に存在するとは限らない点が母音との根本的な相違点である。無声破裂音・摩擦音では子音の音源は声道中の狭窄点にあり、破裂音ではパルス的な雑音が、摩擦音では持続性の雑音が生じる。これらの雑音源も声道による共鳴を受けるが、雑音源よりも前方の声道形状の影響が強い。有声子音の場合には雑音源のほかに喉頭にも音源が存在することになる。破裂子音の音響特徴とその知覚における役割についてはハスキンス研究所における一連の研究が有名である。代表的報告として Liberman (1957) がある。摩擦音のスペクトル特徴については Strevens (1960) 参照。

一方、流音・半子音などのいわゆる sonorant な子音群のスペクトルは母音に類似しており、母音と同様フォルマントによる特徴づけが可能である。英語を対象とした Lehiste (1964) によれば、半子音の第1第2フォルマント周波数は対応する狭母音に比して低く、第3フォルマントにも相違がみられる。また半子音の持続時間は母音に比して一般に短く、定常的なスペクトルを示すことがない。鼻子音のスペクトルも母音や半子音に類似した定常的な性質をそなえているが、アンチフォルマントによる選択的なエネルギーの吸収と鼻腔に起因する極の存在を考慮する必要がある (Fujimura, 1962)。

先に言及したハスキンス研究所における研究は、子音の研究ではスペクトルの静的特性に加えて、スペクトルの時間変化特性の解明が不可欠であることを明らかにした。特に obstruent な子音においては、重要な音響特徴量が複数存在することが稀ではない。例えば無声破裂子音の調音位置の知覚には、burst 自体の音響的性質と隣接する母音の第2フォルマントの軌跡が共に重要である。同じく破裂子音の有声/無声の対立の知覚に寄与する特徴としては低周波数領域における周期性成分の存在、閉鎖区間の持続時間、Voice Onset Time (子音の破

裂時点から声帯振動の開始時点までの時間差)、VOT 区間に含まれる氣息音のスペクトル特徴、後続母音の第1フォルマントの軌跡、先行・後続母音の持続時間などが指摘されている。これらの音響特徴の知覚への寄与は子音の環境によって変動し、さらに音韻体系を反映した言語ごとの差異も存在するようである。

子音の音響学的研究のうち主要なものは Lehiste (1967), Fry (1976) に採録されている。また Kent & Read (1992) にも子音に関する丁寧な解説がある。

2.3 喉頭音源

ファイバースコープによる高速映画などによって観察される現実の声帯振動は非常に複雑であるが、これを一定限度内で近似するモデルとして石坂とフラナガンによる二質量モデル (Ishizaka & Flanagan, 1972) がある。このモデルは声帯振動の垂直方向の位相差を考慮に入れたもので、音声合成システムの音源として広く利用されてきた。近年では病的音声を含む多様な声質のシミュレーションのために、声帯筋と声帯表面の粘膜との位相差も考慮に入れた3質量モデル (Story & Titze, 1995) も提案されるようになった。

声帯振動によって生成される声門音源波形(体積流)をパラメトリックに記述するモデルも重要である。音源波形を直接録音することは不可能なので、声道の伝達特性 T と逆の特性をもつフィルタによって声道の影響を除去する逆フィルタ法が利用される。

Rosenberg (1970) は三角関数や多項式によって音源波形を近似した合成音を用いた知覚実験をおこない、母音の自然性には、声帯振動の一周期に占める声門の開放区間の比率 (open quotient) と、開放期から閉鎖期に移行する時点での数学的不連続性とが大きく貢献することを見出した。

その後 Fant (1986), Fant et al. (1985), Klatt & Klatt (1990) 等のモデルが提案されているが、女声や子供の声を含む多様な合成音を実現するためには T の制御と並んで音源波形の適切な制御が不可欠である (Klatt & Klatt, 1990; 粕谷・楊, 1995)。また喉頭音源のスペクトル特性がパラ言語的情報の伝達に利用されている可能性も示唆されている

(Fujimura et al., 1995)。

声門音源波形の研究には上記の逆フィルタ法が必須であるが、最近、音声波形から声道と音源のパラメータを同時に推定する方法が開発された(Ding et al. 1995)。音源研究への寄与が期待される。

2.4 喉頭音源と声道の相互作用

音源フィルター理論の前提であるSとTの独立性は常に成立しているわけではない。例えば母音の生成時には声門の開放区間では声門下の空洞(気管)が声門上の声道に連結されるためTに変化が生じる。Miki et al. (1994)は声門開放区間では周波数の低域に声道に起因する極(subglottal formant)が生じると同時に高域に位置する極の位置にも複雑な変化が生じることを報告している。

この種の音響的相互作用の他に生理学的な相互作用も存在する。Foの上昇下降に伴って喉頭が上下に移動することはよく知られているが、この移動によって、声道(特に咽頭部)の形状が変化するためTに影響がおよぶ。さらに喉頭は舌骨を介して舌とも連結されているため両者間に力学的な相互作用が存在する(Honda et al., 1994; 平井他, 1995)。従来母音のintrinsic Foとして知られてきた現象はこの相互作用に関係している。Foの変化にともなう母音フォルマントの変化を測定した例として平井(1995)、前川(1996)がある。

2.5 音節

近年の音韻理論では音節やモーラが音韻表示の単位として重要な役割を担うようになったが、音声学的特徴に関する研究は少ない。その中で重要な研究として、音韻論的出自の異なる二重子音の性質を比較して autosegmental phonology による抽象的分析を支持する結果を得た Lahiri & Hankamer (1988)、閉音節にはほぼ普遍的な音声特徴として closed syllable vowel shortening を指摘した Maddieson (1985) 等を挙げることができる。

日本語の音韻論では通常音節の他にモーラも認定するが、日本語にはモーラをもたない方言(シラビーム方言)も存在すると言われている。この種の方言の中から前川(1984)は秋田方言、高田(1985)

は青森方言の促音の持続時間を検討しているが、Maddiesonの指摘する平音節における母音短縮はいずれの方言においても観察されていない。反対に促音におわる音節中の母音の持続時間はむしろ増大する傾向にあり、標準語と同様である。音声的にみる限り、特殊モーラによって形成される日本語の「閉音節」は、シラビーム方言においても、英語・韓国語などに見られる典型的な閉音節と同一視することは困難である。

3 韻律特徴

韻律特徴は発話のタイミングに関するものと喉頭音源の特性に関するものとに大別できる。前者には分節音の持続時間や等時性(isochronism)の問題があり、後者には基本周波数(Fo)や音声振幅の問題がある。

韻律特徴のなかにはアクセント、語強勢、分節音の長短の対立など語レベルでの意味の対立に関与するものもあるが、その他に、句・文・談話レベルの言語構造、さらにバラ言語的、非言語的な情報なども伝達している。

従来、語レベルの現象に比べて、非中核的な意味の伝達に関する研究はなおざりにされてきたうらみがあったが、近年、状況に変化が生じてきた。音声工学では、合成音声の自然性を向上させ、多様な音声を実現するためには音韻論的に非関与的な特徴もふくめた韻律全体の制御が重要であることが認識され、活発な研究がおこなわれはじめた(Klatt, 1987; 勾坂, 1993; 藤崎, 1994)。

3.1 持続時間長

分節音の持続時間長はサウンドスペクトログラフの普及ともななって早くから研究がおこなわれた。持続時間の決定要因には各分節音に固有のものと種々の環境によるものがある。前者に関しては、母音の持続時間長が開口度が狭いほど短いなどの普遍的な傾向が認められる(Lehiste, 1970)。後者に関する一般化は容易でないものの、音韻的な韻律境界の影響は諸言語に共通して認められると言ってよい(Lehiste, 1970; Umeda, 1975, 1977; Klatt, 1976; 川崎, 1983; Kaiki et al., 1992)。また品詞や語の意味内容の影響も認められる(Kaiki et al.,

1992)。ポーズ(無音区間)の有無および持続時間長に関する研究も重要である(東・津熊, 1989; Uyeno et al. 1981; 杉藤, 1989; 海木・匂坂, 1992)。

持続時間に関する現象のうち、言語学者の興味を強くひきつけたものに、言語のリズム単位もしくは等時性に関する研究がある。諸言語は *stress-timed language* と *syllable-timed language* とに大分類されるが、日本語は後者の下位区分としての *mora-timing* に分類される(Han, 1962, 1994; Port et al., 1980, 1987; Homma, 1981; Otake, 1989)。

外国語(特に英語)の音声学習における韻律の重要性は夙に指摘されており、実験的にも支持されているから(Suzuki et al., 1989), 言語学者・語学教師がリズムの問題に注目するのは当然である。しかし、等時性の研究は現在では一種の停滞に陥っているように思われる。リズムの範疇的差異の存在を明証する実験結果は得られておらず、等時性単位としてのモーラを否定する見解(例えば Beckman, 1982)に対する十分な反駁もおこなわれていない。

リズムの研究には今後の解明をまつ基礎的研究課題が種々残されており、持続時間の測定以外に広汎なアプローチを試みる必要がある。例えば知覚面からの検討は従来ほとんどおこなわれてきていないが、加藤他(1993)は持続時間を加工した音声の自然性評定スコアの分析から分節音よりも大きな単位の関与を示唆するとともに隣接する母音子音間のラウドネス差が評定スコアの変動要因として重要であることを指摘している。さらに持続時間長測定における分節音境界の定義にも再検討の余地がある。従来、音響境界はスペクトログラム上でのスペクトル急変点として定義されてきたが、このように変化の結果に着目する定義の他に音声の生成過程を重視して変化の開始点に着目する定義も可能であり、その方が中枢での運動企画をより直裁に反映する可能性が高い(渡辺・相谷, 1995)。Vatikostis-Bateson & Kelso (1993) のように調音運動データの中にリズムを探る研究もある。

3.2 基本周波数 (Fo)

Foの抽出手法にはさまざまなものが提案されており、波形にもとづく零交差法、スペクトルにも

とづくケプストラム法、波形の自己相関による方法などがある。自己相関による分析が普及しているが、Foの抽出を完全に自動化することは難しい。一般に有声区間の冒頭と末尾では声帯振動の非周期性が生じがちでありFoの誤抽出が生じやすい。特に発話末尾では著しい非周期性が観察されやすく、喉頭音源のスペクトル自体にも変化が生じる。

既に述べたように、Foは言語によっては語彙レベルでの意味の対立に関与する他に、句レベルでの修飾構造に関連して組織的に制御されることがあり、平叙文と疑問文の対立のように文レベルでの意味の対立にも関与する。談話レベルでは焦点(フォーカス)の所在が専らFoによって表出される。また、新/旧情報の区別や談話の階層構造の表出にFoが関係していることも、ほぼ確実である。発話意図などのパラ言語情報、感情などの非言語情報の伝達ではFoが主役を果たす。

日本語のアクセントについては戦前にも佐久間鼎らによる実験がおこなわれていた。しかし、本格的な研究は杉藤とその共同研究者によって実施され、杉藤(1982)にまとめられた。アクセントの関連量としてFoが最重要であることが再確認されると同時に、音韻論的なアクセントの指定とその音響学的実現形態とが乖離するケース(無声化母音に核がおかれる発話、おそ下がり現象など)について両者の対応関係が検討された(前川, 1993)。

3.3 イントネーション

句よりも大きな言語単位におけるFoの挙動に関する研究を一括して紹介する。句レベルの現象として最もよく研究されているのは名詞句内部もしくは名詞句間の修飾関係とFoの関係である。特に日本語の左枝わかれ構文と右枝わかれ構文のFoの相違に関してはUyeno et al. (1981), 上野(1988), 窪菌(1992), Kubozono (1993), Fujisaki & Kawai (1988), 郡(1992), Selkirk & Tateishi (1991) など多くの研究がおこなわれているが、右枝わかれ構文の冒頭に生じるFoの上昇を音韻構造のなかにもどう位置づけるかという問題にはまだ納得のゆく解決が与えられていない。

疑問文のイントネーション上の特徴は、これを

句末ないし文末に局在するFo特徴と見る立場もある。しかし実際には平叙文と疑問文との間では発話全体のFo形状が組織的に変化していることが普通である。Maekawa (1991)は単純疑問文と疑問詞疑問文のイントネーション上の相違が、疑問詞疑問文における述語句全体のFo形状の強い抑制にあることを示した。この抑制の結果、疑問詞疑問文では一見有核述語句のアクセントが削除されるかに見えることがあるが、実際にはアクセント削除は生じていない(Maekawa, 1994a; 前川, 1995)。修辭疑問文のイントネーションを分析したMiura & Hara (1995)の結果も、文末イントネーション観の枠組みにはおさまらない性質のものである。

フォーカスに関する研究には郡(1988), Fujisaki & Kawai (1988), Pierrehumbert & Beckman (1988; 以下P&Bと略記)などがある。日本語ではフォーカスがもつばらピッチレンジの拡大によって実現されることは間違いないが、詳細に観察すると持続時間や母音の調音にも組織的な変化が生じており(前川, 1996), さらに声質も変化していると予想される。フォーカスは韻律構造の理論にとっても重要である。P&B (1988)は後述するdownstep規則の適用範囲の決定要因としてフォーカスを重視している。

3.4 イントネーションのモデル

ところで発話全体のFo形状を論じるためにはイントネーションのモデルを考えることが必要である。現在提案されているモデルは全体的なモデルと局所的なモデルに大別することができる。全体的モデルの代表は藤崎の対数線型重畳モデル(藤崎・須藤, 1971; Fujisaki & Hirose, 1984)である。このモデルは、所与のFoデータを時間に沿ったFoの下降を表現するフレーズ成分とアクセントなどの言語情報の表現に参与するアクセント成分とが対数周波数軸上で加算されたものとみなしている。フレーズ成分は持続時間が長く、アクセント成分は短い。また、これらの成分の存在は喉頭における声帯振動制御機構の特性を反映するものと考えられている(Fujisaki, 1988)。

局所的モデルの代表はPierrehumbert(1980)に端を発するtoneモデルである。このモデルは音韻論

的な発想にもとづくもので、発話の要所要所に離散的に指定された音韻論的toneが音声実現規則によって補間されることにより発話のFo形状が計算されると想定する。Foのいわゆる自然下降に関しては発話の構造に依存しないdeclinationの他に、アクセントによって引き起こされるピッチレンジの狭窄現象であるdownstep (catathesis)と発話末尾でのfinal loweringを想定している(P & B, 1988)。

藤崎のモデルは所与のFoデータからモデルのパラメータを推定できる点にすぐれた特徴が認められる。反面、適応が困難と思われる言語の存在が指摘されており(前川, 1992, 1994; Maekawa, 1994b), モデルの適用範囲を制限するか、対象言語によってはモデルを拡張することが必要である。

Toneモデルは対象言語によらず適用可能であると思われるが、現状ではFoデータからtoneの配置を自動的に推定することは困難であり、Foの再現性も高くない。また、音声実現規則を簡素化するために、フォーカス等の効果をできる限り局所的にとらえようとしている点にも問題がある。現実には観察される非局所的かつ双方向的な変動を韻律構造の理論中に取り込む必要がある(前川, 1996)。

4 音韻理論との接点

4.1 量子理論

定量的な音声研究と記号操作にもとづく定性的な音韻研究とは本来相互補完的な関係にあるはずである。しかし、イントネーション研究を稀な例外として、音声研究者と音韻研究者との交流はあまり活発とはいえないのが現状であり、この溝を埋めることが望まれる。

音声学と音韻論の架橋として最も著名な試みは調音動作とその音響的帰結間の非線型性に関する量子理論quantal theoryであろう(Stevens, 1972, 1988)。この理論は、弁別素性の成立基盤を音響学的・生理学的な基盤に立って解明しようとするもので、声道モデルの音響特性の解析にもとづいて、調音動作(例えば舌による狭窄)とその帰結として生じる音響的变化との関係が声道中の位置に依存して大幅に異なることを重視する。

量子理論については様々な反論も提出されているが(Journal of Phonetics Vol. 17, No.1-2の特集参

照), 子音の調音位置が声道中で非一様に分布する事実については量子理論による説明が最も説得力に富んでいる。量子理論とはやや観点が異なるが Lindblom (1986) は母音体系の成立基盤を知覚の側面から探っており、興味ぶかい。

4.2 調音結合

音声研究と音韻研究の立場の相違が鮮明にあらわれる問題のひとつは調音結合の問題である。記号レベルでは単一とみなされる音韻が環境に依存して異なったかたちで実現される調音結合の現象は、音韻研究のなかでは、きわめて顕著な変動が異音現象として記述される他は、物理・生理レベルの平滑化現象として研究対象外に追いやられてきた。この点に関してはヤコブソンらの古典的二項弁別素性理論から今日の多層的な音韻理論まで基本的に同一である。しかし生理的・音響的な実験研究は調音結合の言語依存性を示唆しており、調音結合は音韻研究の対象となる資格をそなえていると考えられる。実際、音素を単位として規則による音声合成システムを構築しようとする、調音結合に関して言語ごとに大幅に異なる制御規則を用意する必要がある。しかも調音結合の実態は複雑であり、これを十分な精度で表現できるモデルはまだ構築されていない。そのため、現在実用化されている規則音声合成システムは、CVC, VCV, demi-syllable など分節音よりも長い素片を合成単位に採用することによって、調音結合の効果を合成単位内部に取りこんでいるものが主流である(広川, 1993)。

調音結合が生じる根本的原因は音声諸器官がかなりの程度まで独立して並行的に動作する事実に求められる。従って調音結合を音韻表示にとりこむためには音韻表示自体を調音運動に倣って多次元化するのがひとつの方法である。

これを極端に推進したのが articulatory phonology (Browman & Goldstein, 1989, 1992) のモデルである。このモデルでは個々の調音体の運動軌跡そのものを音韻表示の最終レベルとして採用しており、時間に沿って運動軌跡が記述される。このような表示では複数の調音体の運動軌跡が時間的に重畳して与えられるから、調音結合の効果もその

なかに自然に表現される。これが articulatory phonology の本質的な利点であると言えるだろう。

類似したモデルに藤村の C/D model (Fujimura, 1992, 1996) がある。調音体の運動軌跡を生成する点は articulatory phonology と同じであるが、運動軌跡を音韻表示とはみなしていない。音声実現規則の入力となるのは metrical tree に類した韻律構造であり、その終端には音節を単位として、弁別素性が指定されている。

C/D model では、韻律構造はまず抽象的なパルス列に変換される。このパルス列は音節および韻律境界の強度を時系列として表現したもので、metrical phonology における grid に該当する。次いで、パルス列は弁別素性の指定に従って個々の調音体に配分される。調音体の運動軌跡はパルス列を入力とするシステム応答として計算され、最終的には調音的な音声合成システムが駆動されることになっている。

C/D model は(そして articulatory phonology も) 現状では多分に概念的なモデルにとどまっている。しかし、発話の韻律構造が分節音も含めた音声生成過程の全体を支配しているという発想は、多くの研究者に共有されつつある認識である。韻律構造が分節音の調音運動に与える影響に関しては Pierrehumbert & Talkin (1990), Sproat & Fujimura (1992) などの興味ぶかい研究がおこなわれている。

5 むすび

近年の音声研究では容易に多量のデータを取得することができるようになった。しかし、改めて指摘するまでもなく、データは最終的に何らかのモデルにまとめられる必要がある。そして、モデルを提供する役割は、当然、言語学にも期待されている。この点で、過去に日本音声学会が音韻に関する理論的研究を排斥した時期があったことはきわめて遺憾であった。

日本の音声学界にとって今何よりも必要なのは、記述音声学の基礎をふまえたうえで言語の形式的特性に関する理論(つまり音韻論)と自然科学的な研究方法とを身につけた新しい世代の研究者の育成である。学会全体として取り組むべき課題であろう。

参考文献

- 東 淳一・津熊良政 (1989) 「統語的あいまい文の理解において Fo とポーズが果たす役割」音声学会会報, 191, 1-5.
- Beckman, M. (1982) "Segment duration and the mora in Japanese." *Phonetica*, 39, 113-35.
- Borden, G. and K. Harris (1980) *Speech science primer*. Baltimore: Williams & Wilkins. 廣瀬 肇 訳「ことばの科学入門」メディカルリサーチセンター, 1984.
- Browman, C. and L. Goldstein (1989) "Articulatory gestures as phonological units." *Phonology*, 6, 201-51.
- (1992) "Articulatory phonology: An overview." *Phonetica*, 49, 155-80.
- Chiba, T. and M. Kajiyama (1941) *The vowel: Its nature and structure*. Tokyo: Kaiseikan.
- Ding, W. et al. (1995). "Simultaneous estimation of vocal tract and voice source parameters based on an ARX model." *IEICE Transactions on Information and Systems*, E78-D, 6, 738-43.
- Fant, G. (1960) *Acoustic theory of speech production*. The Hague: Mouton.
- (1986) "Glottal flow: Models and interaction." *J. Phonetics*, 14, 393-99.
- et al. (1985) "A four-parameter model of glottal flow." *STL-QPSR*, Royal Institute of Technology, 4/1985, 1-13.
- Fry, D. B. (1976) *Acoustic phonetics: A course of basic readings*. Cambridge Univ. Press.
- Fujimura, O. (1962) "Analysis of nasal consonants." *J. Acoust. Soc. America*, 34, 1865-1875.
- (1992) "Phonology and phonetics: A syllable-based model of articulatory organization." *J. Acoust. Soc. Jpn. (E)*, 13-1, 39-48.
- (1996) "The C/D model as a dynamic, non-segmental approach." *Technical Report*, TR-H-184. ATR Human Information Processing Laboratories.
- Fujimura, O. et al. (1995) "Voice quality control within a sentence." In O. Fujimura et al. eds. *Vocal Fold Physiology 8: Voice quality control*. San Diego: Singular Publishing Group.
- Fujisaki, H. (1988) "A note on the physiological and physical basis for the phrase and accent components in the voice fundamental frequency contour." In O. Fujimura ed. *Vocal fold physiology: Voice production, mechanisms and functions*. New York: Raven Press.
- and K. Hirose (1984) "Analysis of voice fundamental frequency contours for declarative sentences of Japanese." *J. Acoust. Soc. Jpn. (E)*, 5-4, 233-42.
- and H. Kawai (1988) "Realization of linguistic information in the voice fundamental frequency contour of the spoken Japanese." *Ann. Bull. RILP*, 22, 183-92, Univ. Tokyo.
- 藤崎博也 (1994) 「韻律研究の諸側面とその課題」日本音響学会講演論文集, 2-5-11 (平成6年10月), 287-90.
- 藤崎博也・須藤 寛 (1971) 「日本語単語アクセントの基本周波数パターンとその生成機構のモデル」日本音響学会誌, 27, 445-53.
- 古井貞熙 (1992) 『音響・音声工学』近代科学社.
- Han, M. (1962) "The feature of duration in Japanese." *Study of sounds*, 10, 65-80.
- (1994) "Acoustic manifestations of mora timing in Japanese." *J. Acoust. Soc. Am.*, 96-1, 73-82.
- Hattori, S., K. Yamamoto and O. Fujimura (1958) "Nasalization of vowels in relation to nasals." *J. Acoust. Soc. Am.* 30, 267-274.
- 平井啓之 (1995) 「Fo 変化に伴う母音のフォルマント周波数の遷移」電子情報通信学会技術報告, SP94-102.
- 他 (1995) 「舌と喉頭の相互作用を考慮した発話器官の生理学的モデル」日本音響学会誌, 51-12, 918-28.
- 広川智久 (1993) 「規則合成における音声合成単位及び音声合成法」日本音響学会誌, 49-12, 847-53.
- Homma, Y. (1981) "Durational relationship between Japanese stops and vowels." *J. Phonetics*, 9-3, 273-281.
- Honda, K. et al. (1994) "A physiological model of speech production and the implication of tongue-larynx interaction." *Proc. ICSLP94*, Yokohama, 1, 175-78.
- Ishizaka, K. and J. Flanagan (1972) "Synthesis of voiced sounds from a two-mass model of the vocal cords." *Bell system Technical Journal*, 51-6, 1233-68.
- Kaiki, N. et al. (1992) "Linguistic properties in the control of segmental duration for speech synthesis." In G. Bailly et al. eds. *Talking machines: Theories, models and designs*. Elsevier Science.
- 海木延佳・匂坂芳典 (1992) 「局所的句構造に基づくポーズ長の分析」電子情報通信学会技術報告 SP91-130.
- 粕谷英樹・楊長盛 (1995) 「音源から見た声質」日本音響学会誌, 51-11, 869-75.
- 他 (1968) 「性別による日本語5母音のピッチ周波数とホルマント周波数の変化」日本音響学会誌, 24-6, 355-364.
- 加藤宏明他 (1993) 「単語内音韻変動に対する許容度」電子情報通信学会技術報告, SP92-145.
- 川崎春子 (1983) 「音声の時間制御に関するモデルと実測データ」日本音響学会誌, 39, 389-96.
- Kent, R. and C. Read (1992) *The acoustic analysis of speech*. San Diego: Singulara Publishing Group.
- 城戸健一 (1985) 『デジタル信号処理入門』丸善.
- Klatt, D. (1976) "Linguistic use of segmental duration in English: Acoustic and perceptual evidence." *J. Acoust. Soc. Am.* 59, 1208-21.
- (1987) "Review of text-to-speech conversion for English." *J. Acoust. Soc. Am.* 82-3, 737-93.
- and L. Klatt (1990) "Analysis, synthesis and perception of voice quality variations among female and male talkers." *J. Acoust. Soc. Am.* 87, 820-57.
- 郡 史郎 (1988) 「強調とイントネーション」杉藤 編 (1989) 収録.
- (1992) 「プロソディーの自律性: フレージングを定める規則について」月刊言語, 大修館, 21-9, 31-37.
- Kubozono H. (1993) *The organization of Japanese prosody*. Tokyo: Kurosio Publishers.

- 窪園晴夫(1992)「プロソディーの普遍性」月刊言語, 大修館, 21-9, 23-30.
- Lahiri, A. and J. Hankamer (1988) "The timing of geminate consonants." *J. Phonetics*, 16, 327-38.
- Lehiste, I. (1964) "Acoustic characteristics of selected English consonants." *Int. J. Am. Ling.*, 30, 1-97.
- (1967) *Readings in acoustic phonetics*. Cambridge: The MIT Press.
- (1970) *Suprasegmentals*. Cambridge: The MIT Press.
- Liberman, A. (1957) "Some results of research in speech perception." *J. Acoust. Soc. Am.*, 29, 117-123.
- Lindblom, B. (1986) "Phonetic universals in vowel system." In J.J. Ohala and J.J. Jaeger eds. *Experimental phonology*. Orlando: Academic Press.
- Maddieson, I. (1985) "Phonetic cues to syllabification." In V. Fromkin, ed. *Phonetic linguistics*. NY: Academic Press.
- Maekawa, K. (1991) "Perception of intonational characteristics of WH and non-WH questions in Tokyo Japanese." *Proc. 12th ICPhS*, 4, 202-5, Aix-en-Provence.
- (1994a) "Is there 'dephrasing' of the accental phrase in Japanese?" *OSU working papers in linguistics*, 44, 146-165.
- (1994b) "Intonational structure of Kumamoto Japanese: A perceptual validation." *Proc. ICSLP94*, Yokohama, 1, 119-22.
- 前川喜久雄 (1984) 「秋田方言促音の持続時間」広島方言研究所編『方言研究年報』27, 231-47, 和泉書院.
- (1992) 「熊本無アクセント方言のイントネーション」月刊言語, 大修館, 21-9, 66-74.
- (1993) 「書評: 杉藤美代子著 日本語アクセントの研究」国語学, 174, 42-47.
- (1994) 「イントネーションの言語学的理論とその問題点」日本音響学会講演論文集, 2-5-12, (平成6年10月), 291-94.
- (1995) 「[アクセント句の dephrasing について] 日本音響学会講演論文集, 1-1-8, (平成7年9月), 213-14.
- (1996) 「フォーカスが発話の時間構造と母音フォルマントにおよぼす影響」日本音響学会講演論文集, 1-P-21 (平成8年3月), 203-4.
- Miki N. et al. (1994) "Vocal tract model and 3-dimensional effect of articulation." *Proc. ICSLP94*, Yokohama, 1, 167-70.
- Miura, I. and N. Hara (1995) "Production and perception of rhetorical questions in Osaka Japanese." *J. Phonetics*, 23-3, 291-303.
- 日本音声言語医学会 (1994) 『声の検査法(第2版)基礎編』. 医歯薬出版.
- Otake, T. (1989) "A cross linguistic contrast in the temporal compensation effect." *The Bulletin*, 191, 14-19, The Phonetic Society of Japan.
- Peterson, G. and H. Barney (1952) "Control methods used in the study of vowels." *J. Acoust. Soc. Am.*, 24, 175-84.
- Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation*. Diss., M.I.T.
- Pierrehumbert, J. and M. Beckman (1988) *Japanese tone structure*. Cambridge: The MIT Press.
- Pierrehumbert, J. and D. Talkin (1990) "Lenition of /h/ and glottal stop." In G. Docherty and D. Ladd, eds., *Papers in laboratory phonology II*. Cambridge: Cambridge Univ. Press.
- Port, R. et al. (1980) "Temporal compensation and universal phonetics." *Phonetica*, 37, 235-52.
- (1987) "Evidence for mora timing in Japanese." *J. Acoust. Soc. Am.*, 81-5, 1574-85.
- Rosenberg, M. (1970) "Effects of glottal pulse shape on the quality of natural vowels." *J. Acoust. Soc. Am.*, 49-2, 583-90.
- 勾坂芳典 (1993) 「韻律制御研究の現状と課題」日本音響学会誌, 49-12, 854-59.
- Selkirk, E. and K. Tateishi (1991) "Syntax and downstep in Japanese." In Georgopoulos et al. eds. *Interdisciplinary approach to language*. Kluwer.
- Sproat, R. and O. Fujimura (1992) "Allophonic variation in English /l/ and its implications for phonetic implementation." *J. Phonetics*, 21-3, 291-312.
- Stevens, K. (1972) "The quantal nature of speech: Evidence from articulatory-acoustic data." In E. David Jr. and P. Denes, eds., *Human communication: A unified view*. NY: McGraw-Hill.
- (1988) "On the quantal nature of speech." *J. Phonetics*, 17-1/2, 3-45.
- Story, B. and R. Titze (1995) "Voice simulation with a body-cover model of the vocal folds." *J. Acoust. Soc. Am.* 97-2, 1249-60.
- Stevens, P. (1960) "Spectra of fricative noise in human speech." *Lang. Speech*, 3, 32-49.
- 杉藤美代子 (1982) 『日本語アクセントの研究』三省堂.
- 編 (1989) 『講座日本語と日本語教育 第2巻 音声・音韻(上)』, 明治書院.
- (1989) 「談話におけるポーズとイントネーション」杉藤 編 (1989) 収録.
- Suzuki, H. et al. (1989) "A comparison of evaluations by American and Japanese listeners of English spoken by Japanese speakers." *Ann. Bull. RILP*, 23, 51-7, Univ. Tokyo.
- 高田正治 (1985) 「促音の調音上の特徴について」研究報告集, 6, 17-40, 国立国語研究所.
- Umeda, N. (1975) "Vowel duration in American English." *J. Acoust. Soc. Am.*, 58-2, 434-45.
- (1977) "Consonant duration in American English." *J. Acoust. Soc. Am.*, 61, 846-58.
- Uyeno, T. et al. (1981) "Syntactic structures and prosody in Japanese." *Ann. Bull. RILP*, 15, 91-108, Univ. Tokyo.
- 上野田鶴子 (1988) 「文法とイントネーション」杉藤 編 (1989) 収録.
- Vatikostis-Bateson and S. Kelso (1993) "Rhythm type and articulatory dynamics in English, French and Japanese." *J. Phonetics*, 21-3, 231-66.
- 渡辺義尚・粕谷英樹 (1995) 「生成過程を考慮した朗読及び対話音声の時間構造の検討」電子情報通信学会技術報告, SP94-103.