

A constitutive representation on the production of glottal waves

田部, 洋祐

<https://doi.org/10.15017/459601>

出版情報 : 九州大学, 2006, 博士 (芸術工学), 課程博士
バージョン :
権利関係 :

第1章

緒論

1.1 はじめに

音声は、声帯振動などによって生成された音源波が、声道を音響的に伝搬し、口唇または鼻孔で放射されることで生成される。これより、音声の生成機構は、音源の生成、声道の形による調音、口唇または鼻孔からの放射の3作用に分離してモデル化することができる [1]。現在の音声情報処理の基礎にあるのが、このモデルを更に簡略化して音源と調音を完全に分離した、図 1.1 に示す線形分離等価回路モデルである。これは、音源 (sound source) と調音フィルタ (articulatory filter) の電氣的等価回路が、相互作用を伴わずに接続された形で音声 (speech) の生成モデルを表現するものである。

図 1.1 の線形分離等価回路モデルに基づいて、1968 年には、Itakura [3] らや Atal [4] らによって、音声の線形予測分析法 (LPC) が開発された。これは、音声を音源パルス列 (pulse sequence: 図 1.1) によって励振された全極型の声道調音フィルタの出力であると考え、実音

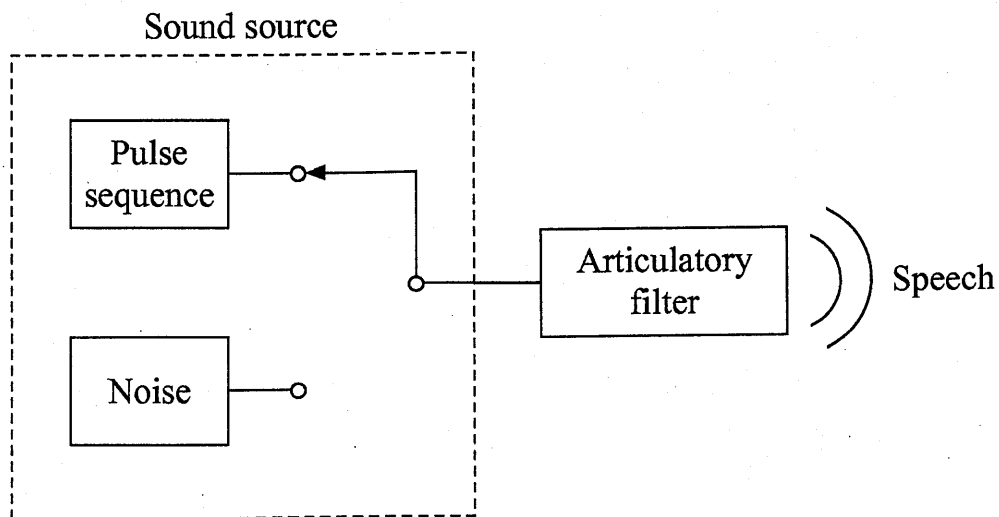


図 1.1 音声生成機構の線形分離等価回路モデル [2]

声からフィルタ係数をパラメトリックに推定する手法である。この手法は、音声研究の歴史にとって極めて重要な出来事であった。その重要性は、この手法が音声波形やそのスペクトルの性質を少数のパラメータで効率的に表現でき、しかもそれらのパラメータが比較的簡単な計算で求まるという点にある [5]。これより、LPC は音声のデジタル符号化と圧縮伝送の技術への利用がなされ、1970年代には Itakura によって PARCOR [6] や LSP [7] の理論へと発展し、現在の携帯電話の伝送方式の基礎となった。また、これらの手法は、パラメータの操作による音声合成や、実音声からのパラメータの推定による音声認識にも利用されるようになった。

図 1.1 の線形分離等価回路モデルから導かれた LPC, PARCOR, LSP といった手法は、現在の音声情報処理の基礎である。しかし、音声生成における実際の音源は、図 1.1 のモデルに示すような単純なパルス列や雑音 (noise) ではなく、それらが組み合わされたより複雑な形態を持ち、また様々な情報を有することが知られている。音源波は、その周期や振幅、スペクトル構造を様々に変化させることによって、歌声や感情を表現する際の基本周波数の変化や、裏声・だみ声・囁き声などの声質の変化をもたらす。このような音源の持つ情報の生成機構を理解するためには、音声の音源生成過程を物理的な実体に即した形でモデル化する研究が不可欠である。音源の物理モデルの研究が進めば、音声の生成機構の科学的理解に役立つばかりでなく、様々な年齢、性別、感情、声質に対応した音源の合成や認識といった工学的な応用や、喉頭摘出者のための代替音源の作成といった医療分野への貢献などが期待される。

本論文では、人間の音声の基本的な音源波である、声門波の物理的な生成モデルを検討する。声門波は、声帯振動によって声門流^{*1}が変調された結果として生じる。したがって、その物理的な生成過程をモデル化するには、声帯の振動機構と、声門流の流体力学的な挙動を表現する必要がある。特に、声門流の表現に際しては、流れの非線形性や声帯振動による境界の移動といった困難な問題を取り扱わなければならない。そこで本論文では、流れの境界層近似を導入し、声門流を複数の構成要素を用いて表現することで、問題の解決を図る。この方法は物理的な意味が明確であり、多様な発声条件の下で生成される声門波を柔軟かつ安定に表現することが期待される。本論文では、境界層近似に基づいた声門波の生成モデルを提案し、その妥当性と意義について論じる。

1.2 声門波の生成過程の概要

本論文では、声門波の生成機構を物理的な実体に即してモデル化する。本節では、音声の生成過程において声門波を生成する機構を担う喉頭の構造について説明し、声門波の物理的な生成過程を概説する。

^{*1} 本論文では、声門を通過する呼気を声門流と呼ぶ。

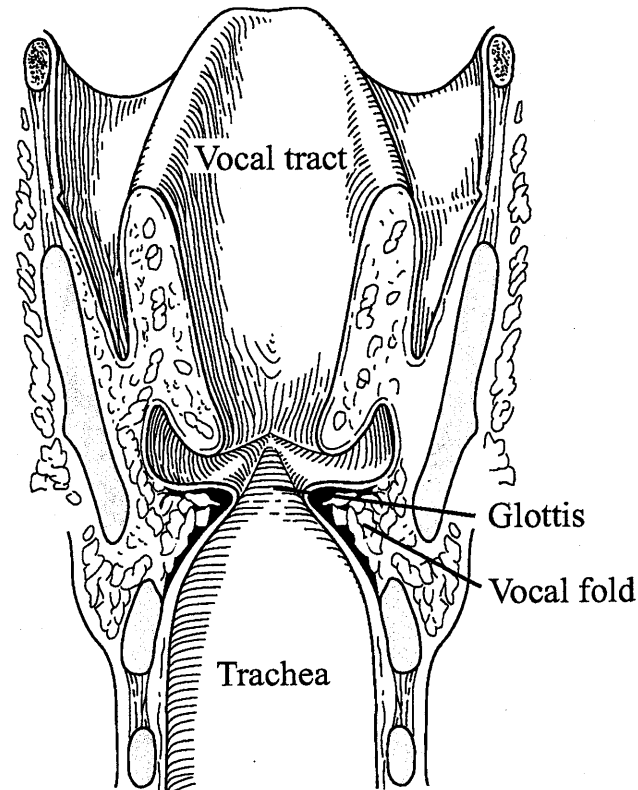


図 1.2 喉頭の前頭断面 [8]

音声の生成過程において、声門波を生成する機構を担うのが、図 1.2 に示す喉頭である [8]。図 1.2 は喉頭を前頭断面より示したものである。ここで、声門波の生成にあたり中心的な働きをする部位は声帯 (vocal fold) である。声帯は気管 (trachea) と声道 (vocal tract) をつなぐ中継点にあたり、肺から送られてくる呼気の通路を開閉する弁の役割を果たす。その大きさは、上下方向の厚さが成人男性で約 3 mm、声帯縁の前後方向の長さが 15 mm の程度である。声帯は粘弾性の特性をもつ組織であり、表面から深部に向かって階層的な構造をなしている。また、左右の声帯の間の呼気の通り道は、声門 (glottis) と呼ばれる。

声門波は、声帯の振動によって声門流が変調された結果として生じる。図 1.3 の上段に、声帯振動のサイクルを示す。声帯の振動は、以下のような機構によって生じるものと考えられている。

1. 初期的に、声門がほぼ閉じられた状態にあるとする。このとき、肺から送られる呼気の流れが妨げられ、声門下部の空気圧 (声門下圧) が上昇することによって、声帯粘膜が押し上げられ、声門が開く (図 1.3: a, b, c)。
2. 呼気から声帯粘膜に働く圧力が減少すると同時に、声帯自身の弾性復元力によって声帯はもとの状態に戻ろうとする。これにより、声門は閉じる動作をする (図 1.3: d, e)。
3. 声帯が初期的な状態に戻り、声門が閉ざされる (図 1.3: f, a)。以上の過程を繰り返すことで、声帯の振動が維持される。

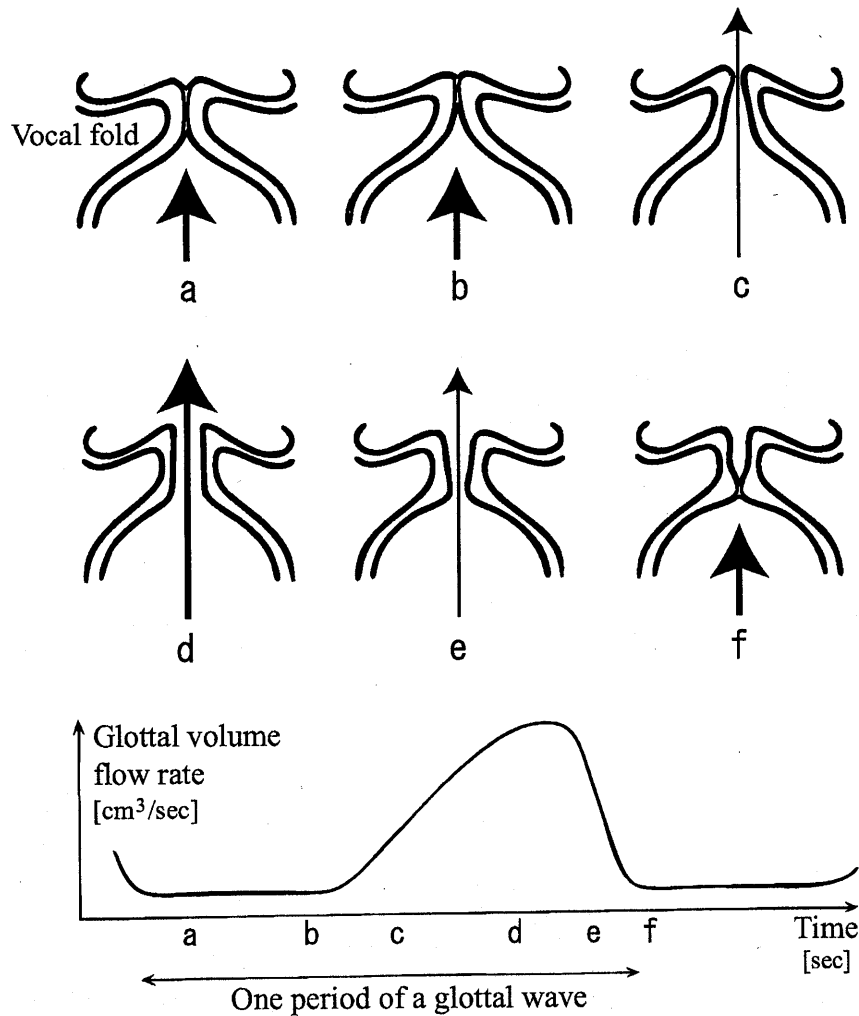


図 1.3 声帯振動のサイクル (上段) と、声門波の様子 (下段) [9]。

このような声門流と粘弾性の声帯との相互作用による声帯振動を、声帯の自励振動と呼ぶ。声帯振動により呼気は変調され、声門を通過する呼気の体積流量 (声門体積流量, glottal volume flow rate) は時間的に変動する。図 1.3 の下段に声門体積流量の時間波形を示す。ここで、横軸は時間、縦軸は単位時間あたりの流量を表している。この声門体積流量の時間波形のことを、本論文では声門波 (glottal wave) と呼ぶ。声門波の時間的な変化が空気の密度変動となり、音波として声道を伝搬することで音声が生産される。

1.3 声門波の生成に関する従来研究

前節では、声帯の自励振動のメカニズムと、声帯振動によって声門波が生成される過程について概説した。本節では、このような声門波の生成過程に関する従来研究について述べる。声門波の生成に関する研究は、基本的概念としての声帯の自励振動説が受け入れられてからは、主として声帯の振動機構に関する検討と、声門流の流体力学的挙動に関する検討が行われてい

る。声帯の振動機構の検討に関しては、声帯の粘弾性を低自由度の機械系としてモデル化する手法によって、声帯の自励振動説の理論的裏付けが行われている。一方、声門流に関しては、声帯の自励振動を持続させる外力としての働きと、声門波のスペクトル構造を決定する要因について、模型実験や流れの数値シミュレーションによって検討されている。

まず、基本的概念として、声帯振動は、粘弾性声帯と声門流との相互作用による自励振動という説が受け入れられている。この説は、19世紀頃に Müller により提唱されたものである。同説は、1958年の van den Berg [10]、1966年の広戸 [11] らによる裏付けのもと、筋弾性空気力学理論として確立された。その後、実効的に振動する部位としての声帯粘膜の役割が重要視され、ボディーカバー理論へと発展した [12]。

声帯の自励振動説に基づいて、声帯の振動機構を低自由度の機械系としてモデル化することが行われた。最初の声帯のモデル化の試みは、1898年に Ewald [13] によって行われた。これは、声帯の粘弾性を1つの質量・バネ・ダンパからなる機械系としてモデル化して、声帯の自励振動説を理論的に説明するものであった。更に、声帯振動の観測によって図 1.3 に示す声帯の振動パタンの生理的特徴が知られる様になると、1968年には Ishizaka ら [14] によって声帯の2質量モデルが提案された。2質量モデルは、声帯の実効的に振動する部位を、質量・バネ・ダンパからなる機械系が連結された2自由度の振動モデルによって表したものである。その後、1995年には Story ら [15] によって、ボディー・カバー理論に基づいた声帯の3質量モデルが提案された。これらのモデルは、実際の声帯組織構造に比べればかなり簡略化されたものであるが、声帯振動の生理学的なデータと良く一致するため、現在でも声帯振動に関する研究で広く用いられている。

声門流に関しては、まず、声帯の自励振動を持続させる外力としての働きについて検討するために、声門における圧力の分布が研究された。1957年に van den Berg ら [16] は、声門の流路が平行な喉頭模型を作成して、声門における圧力の分布を決定する要因を実験的に検討した。これにより、圧力分布を決定する要因として、声門の狭めの区間における Bernoulli の効果と粘性摩擦の影響、声門の出口における流れの剥離の影響が確認された。また、計算機の演算能力の増加に伴い、流れの予測手法として数値シミュレーションが利用されるようになると、1990年には、飯嶋ら [17] によって、静的な喉頭における声門流の流れ抵抗の数値シミュレーションが行われた。数値シミュレーションの精度は年々高まり、2001年とその翌年には、Scherer ら [18-20] によって、様々な声門形状における声帯表面に働く圧力分布の詳細な数値シミュレーションが行われ、同時に行われた模型実験結果との対応が確認された。

また、声門流に関しては、声門波のスペクトル構造を決定する要因についても検討された。1988年に McGowan [21] は、渦音 [22] の理論に基づいて、声門流の渦流のパターンが声門波のスペクトル構造に影響を与える事を予測した。1992年には、三木 [23] によって声門出口で生じる渦のスペクトルが計算され、声門波のスペクトル構造を決定する要因としての渦の効果が確認された。2002年には、Zhao ら [24] が声帯を強制振動させた状態での声門流の数値シ

ミュレーションを行い、流れの剥離により生じる渦流が声門壁に作用して2重極子的な音源として振る舞うことを示した。

1.4 境界層近似に基づく声門流のモデル表現

前節では、声門波の生成過程に関する従来研究について述べた。声帯振動に関しては、声帯の粘弾性を低自由度の機械系としてモデル化する研究によって、声帯の自励振動説の理論的裏付けが行われた。一方、声門流に関しては、模型実験や数値シミュレーションによって、発声時における声門流の挙動や機能が検討された。しかしながら、流れの数値シミュレーションでは、声帯振動による境界の移動が問題となり、数値的に安定な計算格子を作成することは、大変な困難を伴う。また、声門の閉鎖の際には、計算格子が潰れてしまうという問題も生じる。したがって、声門流の特徴を効率的に表現した決定的な声門流モデルは未だ提案されていない。そこで本節では、声門流をモデル化する際のアイデアとしての境界層近似 [25] の概念を呈示し、その利点について述べる。

境界層とは、Prandtl(1904)により提唱された流れの領域の概念である。一般に、高 Reynolds 数^{*2}の流れでは、物体表面上にのみ流れの粘性の影響が表れ、それに伴い速度の急勾配が生じる。この速度の急勾配の領域を境界層と呼ぶ。境界層は物体表面に表れる薄い層であるが、鈍い物体形状においては境界層が物体表面より剥離し、下流側に渦流を形成する。このように、流れの様相の変化を知る上で、境界層は極めて重要な概念になる。

境界層近似のアイデアは、1994年に Pelorson ら [26] によって初めて声門流のモデル化の問題に適用された。これは、声門における流れの抵抗を簡便かつ精度良く求めるために導入された。また、境界層の概念は、流れの様相の変化を検討する際の基準と成り得るため、2004年には、Alipour ら [27] が強制振動させた声門における境界層の挙動を数値シミュレーションによって検討した。これにより、声門の開運動区間 (図 1.3: c, d) では声門の出口付近、閉運動区間 (図 1.3: e) では入口付近で境界層の剥離が生じることが示された。境界層の挙動は声帯の振動現象に呼応しており、この研究結果は境界層近似が声門流のモデル表現の導出に際して有効であることを示唆した。

本研究における声門流のモデル表現は、Pelorson らの声門流モデルを改良したものである。彼らは声門流を1次元流れと仮定したが、本論文では渦流を表現できる様に、声門流を2次元流れと仮定する。これは、声門流のモデル化において渦の表現が不可欠であるためである。境界層近似の導入により、粘性の表れる領域は、声帯表面に沿って生じる境界層と、その剥離によって生じる渦流に限定され、それ以外の流れの領域は非粘性の主流として表される。このよ

^{*2} Reynolds(1883)が定義した流体の性質を調べるための無次元量。物理的には、[慣性力/粘性力]の比で定義されている。Reynolds 数が高いということは、相対的に慣性作用が強く、粘性の影響が小さいということを示している。

うに、本論文では、声門流は主流・境界層・渦流の3つ領域から構成されると考える。多様な発声条件の下で生成させる声門波は、これらの流れの要素によって柔軟に表現できると期待される。

本論文では、声門流を主流・境界層・渦流の3つの領域によってモデル化することを提案する。そして、それぞれの領域に適合した流れの表現を導入することで、柔軟かつ安定な声門流モデルの構築を行う。まず、境界層の領域に関しては、境界層の積分方程式を用いて、境界層の発達と剥離の現象を表現する。この際、境界層の外縁の主流の速度を知る必要があるのだが、本論文では、等角写像の手法を用いることで、自由な声門形状における主流の表現を得る。これにより、声帯振動に伴う境界層の挙動の変化を、適応的に予測することが可能となる。渦流の領域に関しては、離散渦法 [28] を用いて、離散的な渦糸を境界層の剥離位置より時々刻々と放出・対流させることで表現する。この方法では、乱流状態となる渦流領域においては、渦糸の存在密度が高くなるだけで、計算格子を作成する必要がなくなる。計算格子が必要な数値シミュレーション手法である差分法 [28] や有限要素法 [28] では、声帯振動による境界の移動のために、数値的に安定な計算格子を作成することは、大変な困難を伴う。一方、本論文で提案する手法はその様な問題と無縁であり、本手法を安定と述べた理由はこの点に由来する。

1.5 本論文の構成

本論文は、以下のように構成されている。

第2章では、代表的な声門波の生成モデルである、声帯の2質量モデル [14] の説明を行う。ここでは、まずはじめに、声帯の実効的に振動する部位のモデルとして、2自由度の機械モデルを呈示する。次に、声門流の支配方程式を明示した後に、2質量モデルにおける流れのモデル化の過程を明記する。本章は、声門波の生成モデルを構築する際の基本的な考え方と、本論文の生成モデルの特徴を明かにする目的がある。

第3章では、境界層近似に基づいた声門流の構成的表現についての説明を行う。ここではまず、声門流の無次元特徴量の考察を行う。これにより、声門流を主流・境界層・渦流の3つの領域によってモデル化する根拠が示される。次に、個々の流れの領域の具体的な定式化を行う。境界層に関しては、声門流の支配方程式を変形して、声帯表面に沿う境界層の積分方程式 [25] を導出する。自由な声門形状において境界層の積分方程式を解くには、境界層の外縁の主流速を知る必要がある。そこで本論文では、等角写像の手法 [29] を用いることで、自由な声門形状における主流の表現を、トポロジー的に等価な帯領域における流れを取り扱うことで導出する。渦流の領域に関しては、離散渦法 [28] を用いて、離散的な渦糸を境界層の剥離位置より時々刻々と放出・対流させることで表現する。その際、渦糸の運動は、渦度方程式に従って運動させる。また、自由な声門形状における渦糸の数学的な表現に関しては、同様に等角写像の手法を用いることで導出する。

第4章では、Schererら [18-20] の模型実験で用いられた声門形状における境界層の様子を、本モデルに基づいてシミュレーションした結果について述べる。ここではまず、声門形状に依存した境界層の挙動の変化について検討を行う。次に、シミュレーションによって得られた声帯表面での圧力分布を Scherer らの模型実験による実測値と比較をすることで、本手法の精度を検証する。また、与えられた声門形状と実験条件の下で声門を流れる流量を推定し、同様に模型実験の結果と比較することで、本手法の妥当性を検証する。

第5章では、本論の声門流モデルを用いて、声門波の生成シミュレーションを行った結果について述べる。ここではまず、Ishizaka らの2質量モデル [14] を参考に声帯の機械モデルを作成し、第3章で導出した声門流モデルと連立させることで、声門波の生成モデルを構築する。次に、具体的に声門波の生成シミュレーションを行い、発声時における境界層の挙動や、渦流の時空間パタンの時間発展についての検討を行う。また、境界層の剥離位置を固定した声門波の生成モデルを構築し、これと本法によるシミュレーション結果との比較検討を行う。これは、声門波の生成過程における境界層の挙動の重要性を明らかにする目的がある。

最後に、第6章では、本論文全体の総括として、本研究で得られた成果を要約する。また、将来的に検討すべき課題に関する記述を行なう。