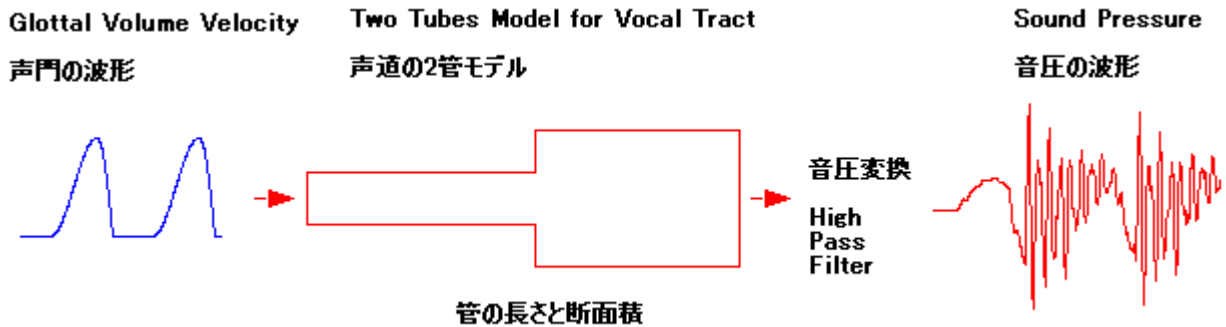


# <音声の波形の生成を理解する>

## 声門波形と2管声道モデルによる音声波形の生成

人間の声の発生の過程を単純化した模型をつかって理解してみよう。

声の発生は、まず、のどの下にある声門で声帯が振動することで、下図の左側の青色のような形の波形が発生する。そして、声帯から喉そして口の中から唇までの間を下図の真中のように2つの管(チューブ)が繋がぎ合わさった物で置き換えてみる。先ほどの青色の波形がこの2つの管に入力されるとある特定の音の組み合わせで共鳴しよく響く。そして、口から出てくる空気の流れ、つまり音声を、音圧を検出するマイクで音を拾うと、下図の右側の赤色のような波形になる。



実際の声門の波形はもっと複雑だし、声帯から喉そして口の中から唇までは途中曲がっていて歯も舌も鼻もあるのに、たった2つの管(チューブ)のつながぎ合わせで置き換えるのはちょっと飛躍しすぎている感があるが、こんな単純化した模型でも、何となく「あ」とか「え」とか聞こえるのである。

この模型では、主に、2つの管(チューブ)の長さや断面積を変化させて、「あ」とか「え」などの音の調整を行う。管(チューブ)と言っても想像がつかないかもしれないが、例えば、水道管、ガス管、または車やバイクについているマフラーも長い管である。下の右側の図は、長さが $L_1$  cmで断面積が $A_1$   $\text{cm}^2$ の管に、長さが $L_2$  cmで断面積が $A_2$   $\text{cm}^2$ の管をつながぎ合わせたイメージ図を示している。このつながぎ合わせた第1の管の左側(声門側に相当)から、声門からの空気の流れが入ると、2つの管でその長さ分の流れの遅延が生じる、また、断面積の違う2つの管のつながぎ目で音が反射する、その結果、いろいろな音の共鳴が起きる。そして、第2の管の右側(口または唇に相当)から出て行く。

入力となる声門の波形は、声門が閉鎖しているとき、声門が開くとき(つまり肺からの空気が出てくるとき)、また、声門が閉じるとき(肺からの空気の流れが止めようとするとき)、の3つの時間を可変できる。この3つの時間を足し合わせたもの(1区間)が、音の高さを決める、いわゆるピッチ周期に相当する。ピッチ周期が短いと高い音に聞こえ、ピッチ周期が長いと低い音(低音)に聞こえる。そして、声門の閉鎖そして開いて閉じるを繰り返して、下図の左側の青色のように山の形が幾つか続く波形となる。

