

# PRINCÍPIOS FUNDAMENTAIS DA PRODUÇÃO DE VOGAIS SEGUNDO A TEORIA ACÚSTICA DE PRODUÇÃO DA FALA

*Fundamental principles of vowel  
production according to the acoustic  
theory of speech production*

Clerton Luiz Felix Barboza\*  
Wilson Júnior de Araújo Carvalho\*\*

## 1. INTRODUÇÃO

A Teoria Acústica de Produção da Fala, comumente conhecida como teoria linear de fonte e filtro de produção da fala, teve o seu desenvolvimento marcado pelos estudos desenvolvidos por dois japoneses, Chiba e Kajiyama (1941), e um sueco, Fant (1960). Estudos que tratam da acústica na produção da fala, entretanto, são bem mais antigos como veremos a seguir.

Os primeiros passos em direção à construção de um modelo de fonte e filtro para a produção vocalica foram propostos por Müller, ainda na década de 1840 (HONDA, 2002). Seus sucessores foram capazes, após desdobramentos da pesquisa original, de produzir vogais sintéticas com a ajuda de diapasões e ressonadores. Tais estudos provocaram um grande aumento de interesse em pesquisas sobre os sons vocálicos e, na segunda metade do século XIX, duas grandes correntes teóricas acerca da produção de vogais

\* UERN - Universidade do Estado do Rio Grande do Norte.

\*\* UECE - Universidade Estadual do Ceará.

foram propostas: a Teoria Inarmônica e a Teoria Harmônica. A primeira teoria afirmava que a excitação produzida pelas pregas vocais servia apenas para inserir o ar nas cavidades superiores. O amortecimento dessa oscilação é que caracterizaria o som vocálico. A segunda, mais elaborada, afirmava que:

As pregas vocais geram uma onda sonora complexa possuidora de uma frequência fundamental e um grande número de harmônicos. Essas frequências dos harmônicos são todas múltiplos exatos da frequência fundamental [...] quando as ondas sonoras passam através da garganta, da boca, e das cavidades nasais, as frequências próximas às frequências de ressonância dessas cavidades são irradiadas muito ampliadas, [...] Essas regiões de frequência reforçada determinam a qualidade vocálica. (FLETCHER, 1929 *apud* HONDA, 2002, p. 190)

Estudos subsequentes de Chiba e Kajiyama (1941) puseram fim a essa discussão ao afirmar que a natureza acústica das vogais é determinada pela forma do trato vocal no momento de produção. Os estudiosos também demonstraram que o espectro vocálico pode ser calculado através do conhecimento da área de transferência da cavidade oral. Para chegar a tais conclusões, Chiba e Kajiyama (1941) se utilizaram de medições do aparelho fonador humano obtidas através de imagens de raio-x e de cálculos das frequências de ressonância de ressonadores simples e duplos de tamanhos compatíveis (BORDEN; HARRIS, 1980)

Tendo por base esses estudos preliminares, Fant (1960) desenvolve e publica sua Teoria Acústica de Produção da Fala. Seu principal objetivo era contribuir para “o entendimento das relações entre os dados articulatórios da produção da fala e os dados físicos de descrição da onda sonora” (FANT, 1960, p. 5). Considerando o estudo de Fant, e recorrendo, sempre que necessário, a trabalhos que visam a explicitar os fundamentos da teoria em questão, tais como Borden e Harris (1980), Pickett (1980), Kent e Read (1992) e Marusso (2005), buscamos apresentar abaixo seus conceitos fundamentais, principalmente, no que diz respeito à produção dos sons vocálicos.

## 2. ANÁLISE DAS CARACTERÍSTICAS ACÚSTICO-ARTICULATÓRIAS DA FONTE SONORA

A teoria da fonte e filtro assume, em qualquer som produzido pelo aparelho fonador humano, a existência de uma fonte sonora, que provê a

entrada acústica no sistema, e de um filtro, que modula esta fonte com o objetivo de realizar os mais diversos sons vocálicos e consonantais recorrentes nas línguas naturais. A fonte de ruído pode ter sua origem associada a três fatores:

- a) vibração das pregas vocais (vozeamento);
- b) ruído turbulento (fricção) do ar originado por constrição;
- c) a combinação das duas fontes supracitadas.

O fechamento intermitente das pregas vocais provoca o aumento da pressão subglotal do fluxo de ar proveniente dos pulmões, o que origina a fonte sonora do sistema. Essa pressão precisa ser maior que a encontrada na cavidade oral para que seja produzido o vozeamento (STEVENS, 1998). Com a vibração das pregas vocais, uma série de ondas sonoras periódicas complexas é produzida. A determinação do número de ciclos por segundo dessas ondas nos possibilita determinar, em Hertz (Hz), o valor da frequência fundamental (FO) da fonte. Esse valor é comumente percebido pelo ouvinte como a altura tonal ou *pitch* de um determinado som.

Os valores de frequência fundamental (FO) não são dependentes apenas da pressão subglotal – fatores como tamanho, massa, densidade e tensão das pregas vocais também são importantes para a determinação de seu valor médio de vibração. Desse modo, se levarmos em conta as diferenças de tamanho, massa e densidade entre homens, mulheres e crianças, teremos diferentes FO para esses indivíduos.

Por apresentar a maior dimensão, a fonte sonora do homem produz a mais baixa frequência fundamental do grupo por necessitar de maior energia para a produção de ondas sonoras periódicas. Mulheres, de maneira geral possuidoras de pregas vocais de menor dimensão, apresentam um valor médio superior ao encontrado na fala dos homens. As crianças de ambos os sexos possuem os valores mais altos.

As características de tensão das pregas vocais são também importantes para determinar a FO da fonte sonora, por variarem de acordo com a tensão aplicada pelo falante sobre estas pregas no momento da fonação. Quanto maior a tensão, maior será a FO, ocorrendo o oposto com o relaxamento das pregas. As considerações apresentadas acima, tomando as pregas vocais como fonte sonora única, são pertinentes a todas as soantes, como as vogais. Na Figura 1 podemos observar o posicionamento das cavidades supraglóticas para a realização de um som vocálico oral.

Características da FO, apesar de importantes para a análise dos sons vocálicos, não são relevantes na produção de sons em que as pregas vocais não são utilizadas como fonte sonora, os sons não vocálicos. Estes, caracteristicamente possuidores de ruído fricativo, como podemos ver na Figura 2, possuem sua fonte de ruído localizada na região de estreitamen-

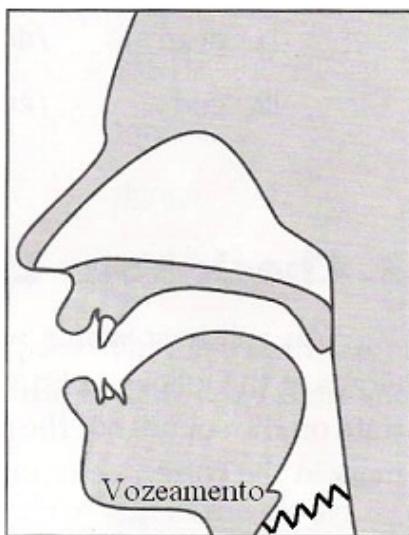


FIGURA 1 – POSIÇÃO DO TRATO VOCAL PARA A PRODUÇÃO DA VOGAL ORAL [ə] (adaptada de ORION, 1997)

to do trato vocal. Essa constrição gera um alto nível de pressão na região entre a glote e o ponto de estreitamento, produzindo ruído turbulento no momento em que o ar passa nessa região. Esse ruído é então filtrado pela cavidade supraglótica e suas ondas sonoras aperiódicas são irradiadas para fora do trato vocal.

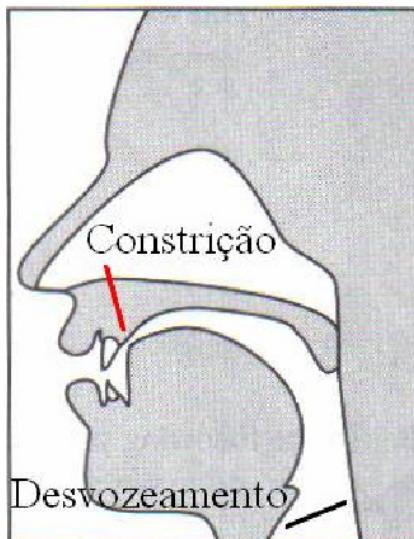


FIGURA 2 – POSIÇÃO DO TRATO VOCAL PARA A PRODUÇÃO DA FRICATIVA ALVEOLAR DESVOZEADA [s] (adaptada de ORION, 1997)

Por fim, tratamos sobre a combinação das duas fontes supracitadas: o vozeamento das pregas vocais e o ruído turbulento gerado pela constrição do trato vocal. Utilizamos como exemplo, na discussão desse tipo de fonte sonora composta, as fricativas vozeadas. Nesses sons, associado ao ruído turbulento descrito anteriormente, como em outras consoantes, podemos ter, de maneira concomitante, o vozeamento das pregas vocais, gerando duas fontes de energia acústica para o sistema (Figura 3).

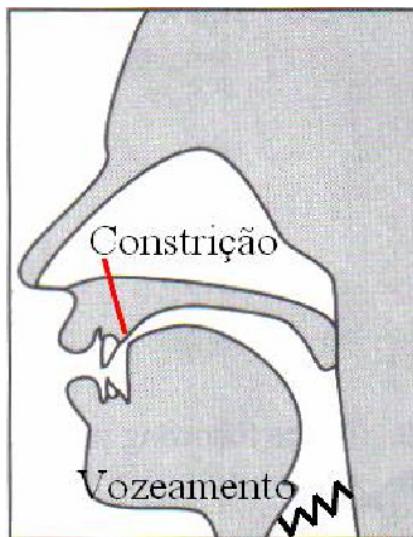


FIGURA 3 – POSIÇÃO DO TRATO VOCAL PARA A PRODUÇÃO DA FRICATIVA ALVEOLAR VOZEADA [z] (adaptada de ORION, 1997)

Podemos depreender, a partir dessa breve discussão acerca das fontes de energia acústica para a produção de sons vocálicos e não vocálicos, que a vibração das pregas vocais ocupa um lugar central nesta discussão. Buscamos, a seguir, aprofundar nosso conhecimento acerca de algumas características de nosso objeto de estudo.

A Figura 4 apresenta o aspecto das cordas vocais em posição aberta<sup>1</sup>. Essas pequenas pregas, para servirem de fonte de som, devem se encontrar fechadas, permitindo breves aberturas, devido à alta pressão abaixo da glote (o espaço em forma de V entre as pregas vocais), para a produção do som vozeado.

A vibração gerada pelas cordas vocais produz uma onda sonora e inicia-se com um aumento do pico da pressão até a obtenção de um valor de amplitude máximo, medido em decibéis (dB). Esse pico é seguido de uma

<sup>1</sup> Um vídeo das cordas vocais produzindo vozeamento pode ser acessado em: <[http://www.entusa.com/normal\\_larynx.htm](http://www.entusa.com/normal_larynx.htm)>.

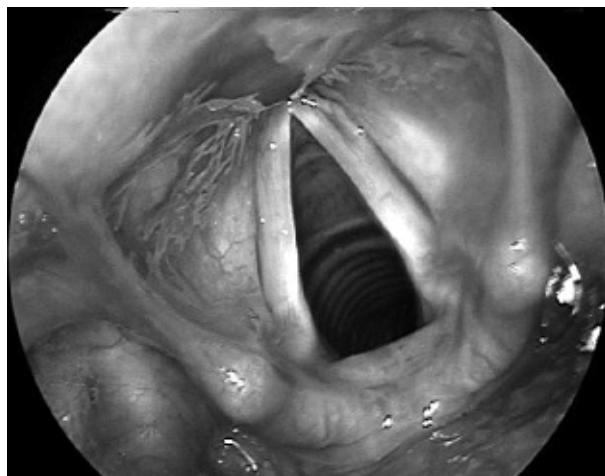


FIGURA 4 – PREGAS VOCais EM POSIÇÃO ABERTA (ENTUSA, 2007)

redução até um valor mínimo de pressão e sua volta, a um valor zero (LA-DEFOGED, 1962). Podemos observar este ciclo na Figura 5, cujo eixo vertical representa pressão positiva ou negativa em decibel e, seu eixo horizontal, a duração dos ciclos.

Esse tipo de onda sonora, produzida pelas pregas vocais, é chamada de ciclo e seu número de repetições por segundo é, como vimos, medido em Hz. Uma vez que as produções dos ciclos se originam da abertura e fechamento das pregas vocais, temos valores idênticos para as medições acústicas de ciclos e para as medições articulatórias de abertura e fechamento das pregas vocais.

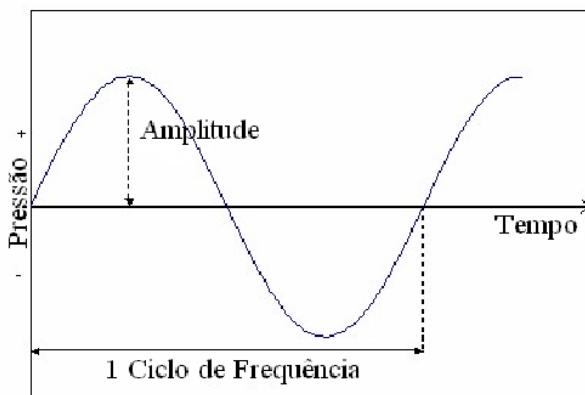


FIGURA 5 – UM CICLO DE FREQUÊNCIA DE UMA ONDA SONORA PERIÓDICA (adaptada de KOO, 2007)

Os ciclos produzidos pelas pregas vocais são periódicos por possuírem aproximadamente a mesma taxa de repetição. Um som aperiódico, por outro lado, não possui tais características de repetição sistemática por um dado período de tempo, sendo caracterizado como um som sem regularidade aparente quando comparado a sons com onda sonora characteristicamente periódica. Sons periódicos, portanto, são possuidores de vozeamento como sua fonte de energia acústica, podendo ainda vir a possuir fricção como segunda fonte. Sons aperiódicos envolvem apenas o ruído turbulento fricativo como fonte de energia sonora.

O som periódico produzido pelas pregas vocais, como vimos, determina o valor da FO da fonte acústica no caso das vogais. Esse som é constituído, além da primeira frequência, de diferentes harmônicos: simples múltiplos da frequência fundamental. Se a frequência fundamental de um homem se encontra na faixa dos 120 Hz (120 aberturas e fechamentos das pregas vocais por segundo) podemos calcular seus três primeiros harmônicos em 240, 360, e 480 Hz, respectivamente. Qualquer falante, como já mencionamos, pode alterar a taxa de frequência fundamental e consequentemente mudar os valores dos harmônicos constituintes da fonte sonora.

A FO e seus respectivos harmônicos são muito importantes para a caracterização do espectro da fonte de energia apresentado na Figura 6. É importante lembrar que as alterações na taxa de frequência fundamental e a mudança nos valores dos harmônicos se aplicam apenas à fonte e não ao filtro. As alterações de ressonância decorrentes das características do filtro serão tratadas na próxima seção de nossa discussão.

Esse tipo de espectro é denominado linear por apresentar a distribuição de energia em forma de linhas verticais. Como podemos observar, a amplitude (eixo vertical) decresce a uma taxa constante à medida que a frequência aumenta (eixo horizontal). Isso nos mostra que a maior parte da energia em sons vozeados se encontra nas frequências mais baixas do espectro sonoro. A taxa de declínio é de 12 dB cada vez que a frequência dobra em cada harmônico.

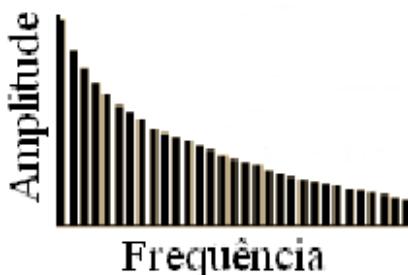


FIGURA 6 – ESPECTRO LINEAR DO SOM PRODUZIDO PELAS CORDAS VOCAIS (adaptada de SUNDBERG, 2005)

### 3. ANÁLISE DAS CARACTERÍSTICAS ACÚSTICO-ARTICULATÓRIAS DO FILTRO SONORO

Após termos caracterizado a fonte sonora de sons vozeados, passamos, a partir deste momento, a nos concentrar nas características do filtro sonoro que determinará diferentes ressonâncias na cavidade oral pela alteração do posicionamento dos diferentes órgãos da fala.

O modelo de filtro mais utilizado atualmente envolve a comparação do filtro com um tubo contendo um dos lados aberto e o outro fechado por uma membrana vibratória que é utilizada como a fonte acústica. O lado aberto do tubo representa a boca e o lado fechado pela membrana elástica tenta simular as pregas vocais. O tubo em si é o ressonador, representando o trato vocal. A Figura 7 (adaptada de KENT; READ, 1992) apresenta um tubo uniforme com um dos lados fechado.

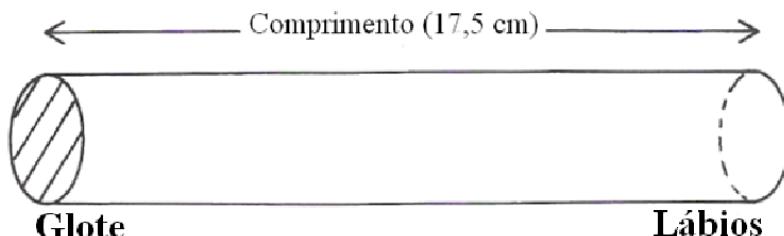


FIGURA 7 – TUBO SIMULANDO O TRATO VOCAL PARA A PRODUÇÃO DE VOGAL

Tubos desse tipo, por possuírem apenas uma saída para as ondas sonoras, podem ser utilizados na simulação de vogais orais. As primeiras frequências de ressonância de nosso tubo que, devido ao processo de diminuição da energia dos harmônicos, são as mais relevantes para a comunicação, podem ser facilmente calculadas. A fórmula para esse cálculo envolve o conhecimento de algumas variáveis, como o número da frequência de ressonância desejada ( $n$ ), sendo 1 a primeira frequência possível de ser calculada, e o comprimento do tubo ( $l = 17,5$  cm), assim como da constante “c” que corresponde à velocidade do som ( $c=35.000$  cm/seg.):

$$F_n = (2n - 1) c / 4l$$

A fórmula acima basicamente reza que o tubo ressoará com amplitude máxima, a característica da onda sonora que mais diretamente se relaciona com a intensidade, ou volume do som (TRASK, 1996), quando o comprimento da onda do som for quatro vezes o comprimento do tubo. Ao aplicar a fórmula acima ao tubo em questão, encontraremos as diferentes

frequências em que o espectro sonoro da fonte será enfatizado. A primeira frequência será calculada abaixo, à guisa de exemplo:

$$F1 = (2 \times 1 - 1) \times 35.000 \text{ (cm/seg.)} / 4 \times 17,5 \text{ (cm)}$$

$$F1 = (2 - 1) \times 35.000 / 70$$

$$F1 = 1 \times 500$$

$$F1 = 500 \text{ (Hz)}$$

A primeira frequência de ressonância do tubo (F1) se encontra, portanto, na faixa dos 500 Hz. Subsequentes frequências se encontram, ao aplicarmos a fórmula acima para encontrar os valores de F2, F3 e F4, respectivamente, em torno dos 1500 Hz, 2500 Hz, e 3500 Hz.

As frequências de ressonância acima não seriam relevantes se não tivéssemos levado em consideração um tamanho de tubo semelhante ao encontrado na média do comprimento do trato vocal de um homem adulto. Além disso, devemos enfatizar que não são encontradas diferenças significativas nas frequências de ressonância se o tubo for reto ou curvo (SONDHI, 1986). A configuração do trato vocal para a produção de uma vogal semelhante à produzida por nosso modelo é, dessa forma, bastante semelhante à configuração apresentada na Figura 1: o som [ə], o *schwa*.

Mudanças no comprimento do tubo, como podemos ver, alterarão também os valores das frequências de ressonância. Caso o tubo fosse mais comprido, as frequências seriam mais baixas. No caso de uma redução do comprimento, as frequências de ressonância apresentariam valores mais altos. Isso também nos ajuda a entender o porquê das diferenças encontradas na fala de uma criança, de uma mulher e de um homem adultos. De maneira semelhante à descrita em nossa discussão acerca das alterações do tamanho de nossa fonte sonora (as pregas vocais), o aumento no comprimento de nosso filtro (a cavidade oral) alterará as frequências de ressonância.

Antes de começar a discorrer acerca de diferentes posições do trato vocal, que causarão diferentes formas no tubo e, portanto, diferentes frequências de ressonância, é pertinente comentar a função da expressão ( $2n - 1$ ) na fórmula. Seu objetivo único é gerar um conjunto de números ímpares, uma vez que as ressonâncias ocorrem a múltiplos ímpares da primeira frequência de ressonância ( $F1 = 1c/4l$ ) explicitada acima. As frequências de F2, F3 e F4, serão expressas respectivamente, pelas fórmulas  $3c/4l$ ,  $5c/4l$ ,  $7c/4l$ .

As ondas sonoras de múltiplos ímpares se encontram no mesmo ponto e no mesmo sentido ao passarem do estágio de pressão positiva para o estágio de pressão negativa de seus ciclos, enfatizando umas as outras. As ondas sonoras de múltiplos pares, por outro lado, se encontram nesse ponto em trajetórias opostas, sendo, na maioria das vezes, suprimidas, como visto na Figura 8 (BORDEN; HARRIS, 1980).

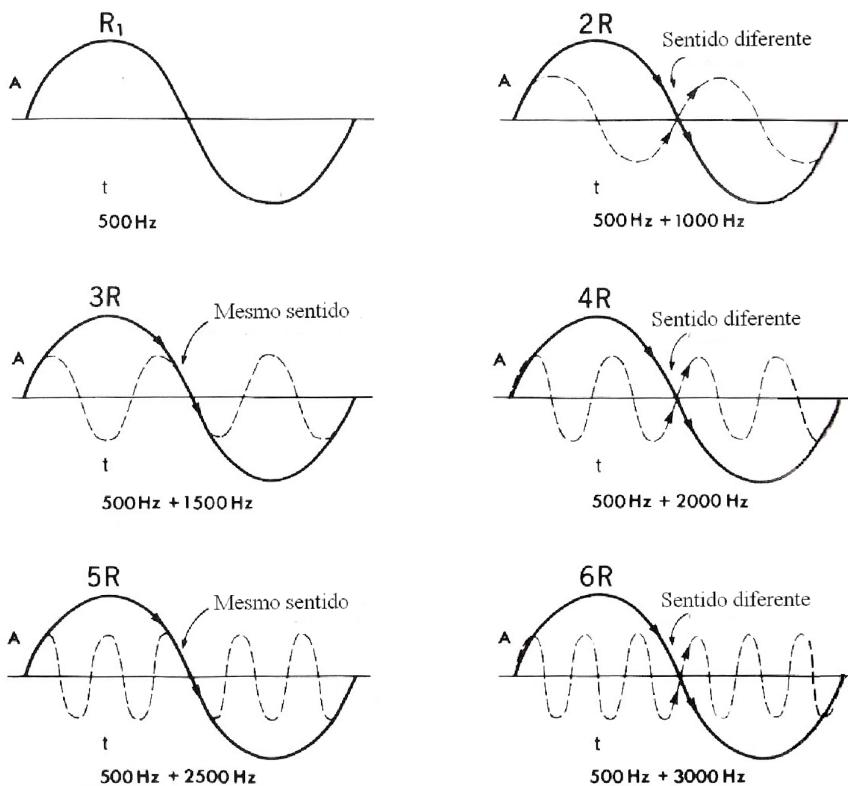


FIGURA 8 – RESSONÂNCIAS IMPARES SENDO ENFATIZADAS (3R, 5R)  
E PARES SENDO SUPRIMIDAS (2R, 4R, 6R) DEVIDO ÀS CARACTERÍSTICAS  
DE 1R (adaptada de BORDEN; HARRIS, 1980)

A primeira ressonância (R1), por possuir a frequência mais baixa (explicitada pela maior duração do ciclo no eixo horizontal tempo 't' comparada às outras ressonâncias) dentro do espectro, normalmente contará com uma amplitude maior (explicitada pelo maior deslocamento no eixo vertical de amplitude 'A' comparada as outras ressonâncias). Essa maior ênfase de R1 nos ajuda a entender sua capacidade, devido a sua maior pressão em dB, de suprimir outras frequências ressonantes mais elevadas (em Hz), mas de pressão inferior. Também devemos lembrar que o ouvido humano é mais suscetível às frequências baixas que às altas.

A Figura 9, a seguir, ilustra como diferentes configurações do trato vocal determinam diferentes posicionamentos das frequências a serem enfatizadas e suprimidas no espectro sonoro de saída.

Os diferentes pontos de constrição do trato vocal, na primeira coluna da figura acima, são convertidos em modelos de tubo, como vimos

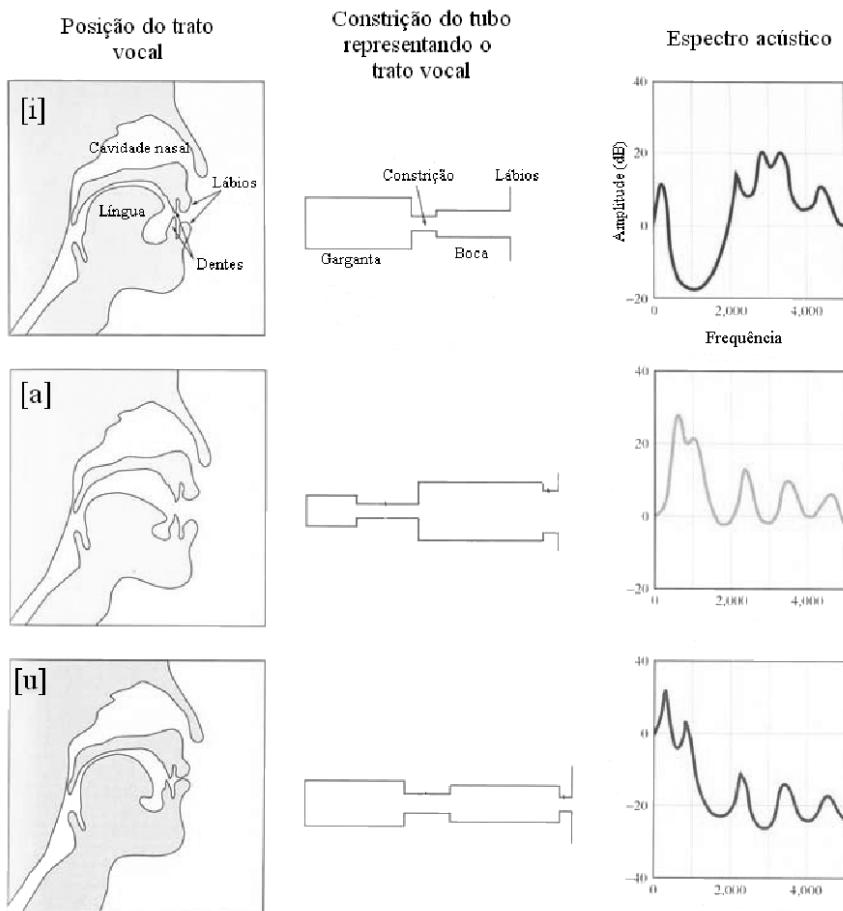


FIGURA 9 – POSIÇÃO DO TRATO VOCAL, DOS PONTOS DE CONSTRIÇÃO CORRESPONDENTES NUM TUBO, E DO ESPECTRO ACÚSTICO DAS VOGAIS [i, a, u] (adaptada de THIES, 2005a)

anteriormente. Essas diferentes características do trato vocal implicam diferentes frequências a serem enfatizadas (formantes) e suprimidas, produzindo um espectro acústico que difere em cada uma das posições articulatórias apresentadas. A vogal [i], por exemplo, apresenta características formânticas, visualizadas na terceira coluna, bastante diferentes das vogais [a] e [u]. O primeiro som, devido ao seu ponto de estreitamento ser diferente dos demais sons vocálicos apresentados, possui seus dois primeiros picos formânticos muito mais separados. Maiores considerações acerca da conexão entre configuração articulatória e características dos formantes podem ser encontradas quando tratamos diretamente desta relação, a seguir.

Devemos lembrar que as frequências de ressonância descritas acima apenas são determinadas de acordo com a forma e comprimento do tubo.

Essas frequências de ressonância, denominadas de formantes, não possuem energia em si. A configuração do trato vocal determinará as frequências de ressonância do filtro que poderão, ou não, vir a ser utilizadas por uma fonte sonora produzida pelas pregas vocais.

Os formantes são identificados por números, sendo apenas os primeiros, mais especificamente F1, F2, F3 e F4, de maior interesse para a caracterização vocalica devido ao baixo nível de energia existente em frequências mais altas. Os formantes são normalmente descritos tendo em vista sua frequência, medida em Hertz, e se analisados de maneira conjunta, constituem a função de transferência do trato vocal: devido às suas propriedades de enfatizar e suprimir certas regiões do espectro sonoro produzido pelas pregas vocais, os formantes determinam a nossa percepção dos sons vocalicos.

Devido ao fato de suas características variarem de maneira síncrona com o posicionamento do trato vocal, as características de frequência dos formantes podem ser previstas de maneira bastante segura se levarmos em conta o ponto de maior constrição da cavidade oral e as características das ondas de ressonância de uma vogal neutra.

Ao levarmos em conta o comprimento das primeiras ondas de ressonância do tubo, aplicando a fórmula de cálculo dos formantes apresentada anteriormente, observamos que a onda sonora da primeira frequência de ressonância (F1) terá um comprimento quatro vezes maior que o tubo. Em outras palavras, apenas  $\frac{1}{4}$  (25%) de um ciclo da onda pode energizar o ar dentro do tubo em um dado momento. A segunda ressonância (F2) possui um ciclo três vezes menor do que a primeira, o que implica que  $\frac{3}{4}$  (75%) de um ciclo da onda podem ser comportados dentro do tubo. As terceira (F3) e quarta (F4) ressonâncias possuirão, ambas, um ciclo de comprimento menor. Cada uma dessas ressonâncias possuirá respectivamente 1,  $\frac{1}{4}$  de ciclo (um ciclo completo e mais 25% do próximo ciclo) e 1,  $\frac{3}{4}$  de ciclo (um ciclo completo e mais 75% do próximo ciclo) dentro do tubo, conforme podemos visualizar na Figura 10.

Com o estabelecimento dos pontos de máxima amplitude (eixo vertical) de vibração das quatro frequências de ressonância do tubo, encontramos a localização dos pontos de nós das frequências em questão. Pontos antinós, por outro lado, são associados aos pontos de mínima amplitude da onda (localizados na linha central do tubo). O conhecimento dos locais onde esses pontos ocorrem dentro do trato vocal, associado ao conhecimento de como uma constrição próxima a um ponto de nó ou de antinó influencia no aumento ou na redução das frequências envolvidas, permite-nos prever o comportamento das diferentes frequências de ressonância frente a outras configurações dos órgãos da fala. A Figura 11 a seguir, adaptada de Chiba e Kajiyama (1941), mostra-nos o posicionamento dos pontos já discutidos.

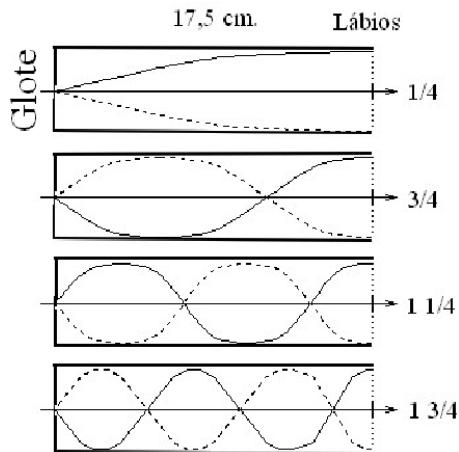


FIGURA 10 – COMPRIMENTO DE ONDA DAS QUATRO PRIMEIRAS FREQUÊNCIAS DE RESSONÂNCIA DE UM TUBO UNIFORME (adaptada de JACKSON, [200-J])

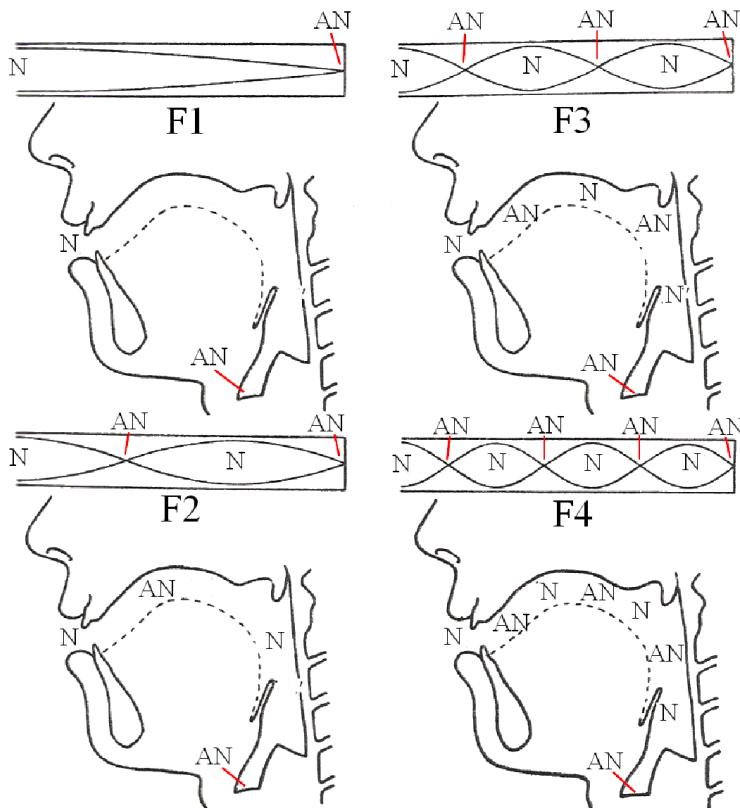


FIGURA 11 – PONTOS DE NÓS (N) E ANTINÓS (AN) NO TUBO E TRATO VOCAL (adaptada de CHIBA; KAJIYAMA, 1941)

Se analisarmos de maneira conjunta a figura contendo os órgãos da fala na produção de uma vogal e o tubo apresentado logo acima, poderemos associar seus pontos de nó e antinó. Os pontos de nó são apresentados no tubo no momento em que a onda sonora apresenta amplitude máxima, quer positiva ou negativa, se aproximando dos seus limites. Pontos de antinó, por outro lado, se encontram onde as ondas sonoras cruzam-se em seu ponto de amplitude mínimo no centro do tubo. Previsões acerca do comportamento das frequências de ressonância ao inserirmos um ponto de constrição próximo a um nó ou antinó, a partir de uma posição neutra do trato vocalico, são apresentadas por Kent e Read (1992):

- a) uma constrição próxima de um ponto de nó baixa a frequência dos formantes;
- b) uma constrição próxima a um ponto de antinó aumenta a frequência dos formantes.

Regras gerais, tendo em vista conhecimentos acerca do posicionamento articulatório do trato vocal para a produção de vogais, juntamente com essas novas informações, permitem-nos fazer as seguintes previsões acerca do comportamento das frequências de ressonância dos dois primeiros formantes (KENT; READ, 1992):

- a) a frequência de F1 é mais baixa com uma constrição na parte mais anterior do trato. Exemplo disso ocorre com as vogais [i] e [u]. A primeira envolve o posicionamento alto e anterior da língua e a segunda, o arredondamento dos lábios. A frequência de F1 é elevada caso ocorra uma constrição na região da faringe ou laringe, como no som [a];
- b) a frequência de F2 é mais baixa caso exista uma constrição nos lábios ou na parte mais posterior acima da laringe, e.g. [u]. O aumento da frequência de F2 é obtido com constrição na parte anterior do trato vocal, na região da vogal [i].

As características da onda sonora descritas acima podem ser observadas em um espectro como os já mostrados na Figura 9. Podemos, ainda, observar mudanças na composição das frequências de ressonância de uma onda sonora através de uma imagem muito ampliada.

Vogais, portanto, são ondas sonoras complexas constituídas por concentrações e reduções da energia presente no espectro sonoro de saída. Uma onda sonora complexa nada mais é do que a sobreposição de duas ou mais ondas sinusoidais. Uma onda sinusoidal pode ser descrita como uma onda sonora com os mesmos valores de amplitude se repetindo em períodos de tempo semelhantes (FRY, 1979). Ondas sonoras como as apresentadas na Figura 10, em nossa discussão acerca dos ciclos da onda sonora, podem ser descritas como sinusoidais. A adição de sons de diferentes frequências resulta numa onda sonora complexa como mostrado na Figura 12 a seguir.

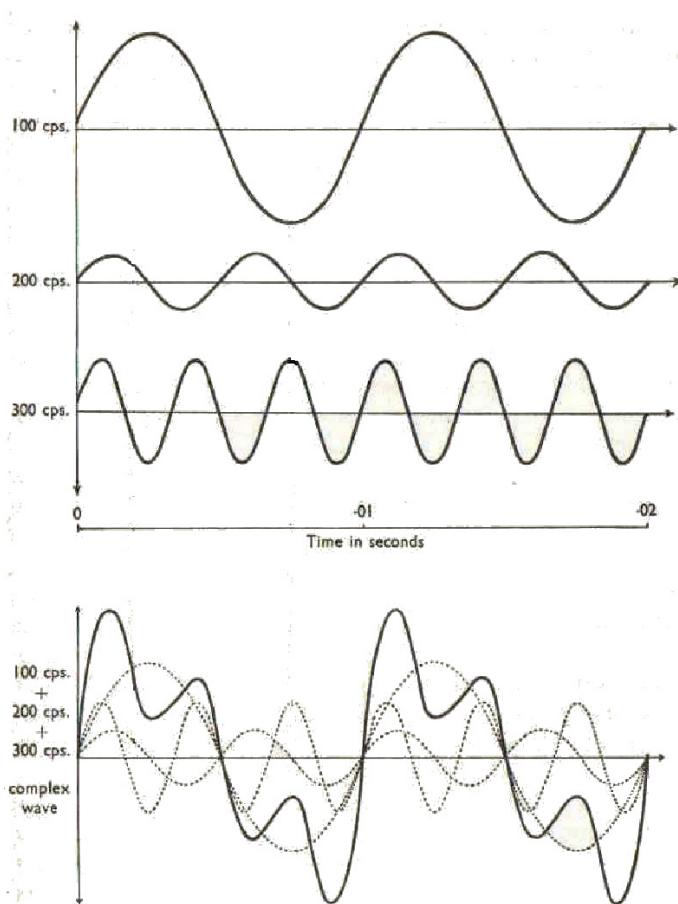


FIGURA 12 – TRÊS ONDAS SINUSOIDAIS DE FREQUÊNCIAS E AMPLITUDES DIFERENTES UNIDAS NUMA ONDA COMPLEXA (adaptada de LADEFOGED, 1962)

Descrições acústicas de vogais que permitem o estudo dos formantes se baseiam na utilização de algoritmos matemáticos que decompõem a onda sonora complexa das vogais em seus diferentes constituintes. Atualmente dois métodos são utilizados para a decomposição de ondas sonoras complexas: o FFT (*Fast Fourier Transform*) e o LPC (*Linear Predictive Coding*).

Além dos gráficos apresentados até este momento, outra forma de representar o sinal acústico de uma vogal é por meio do espectro acústico. Este pode ser descrito como a representação gráfica dos picos de amplitude de um som (eixo vertical), se estendendo por diversas frequências (eixo horizontal), em um momento particular de análise, como vimos na Figura 9. No entanto, em muitos momentos, é importante visualizar as características de amplitude sonora em uma figura contendo também a variante tempo. Esse

tipo de análise é normalmente obtido por meio da ajuda do espectrograma, que se constitui da representação de frequências (eixo vertical), tempo (eixo horizontal), e amplitude (escurecimento de qualquer ponto do gráfico). A Figura 13 mostra espectrogramas de alguns monotongos do Inglês com seus formantes indicados por setas.

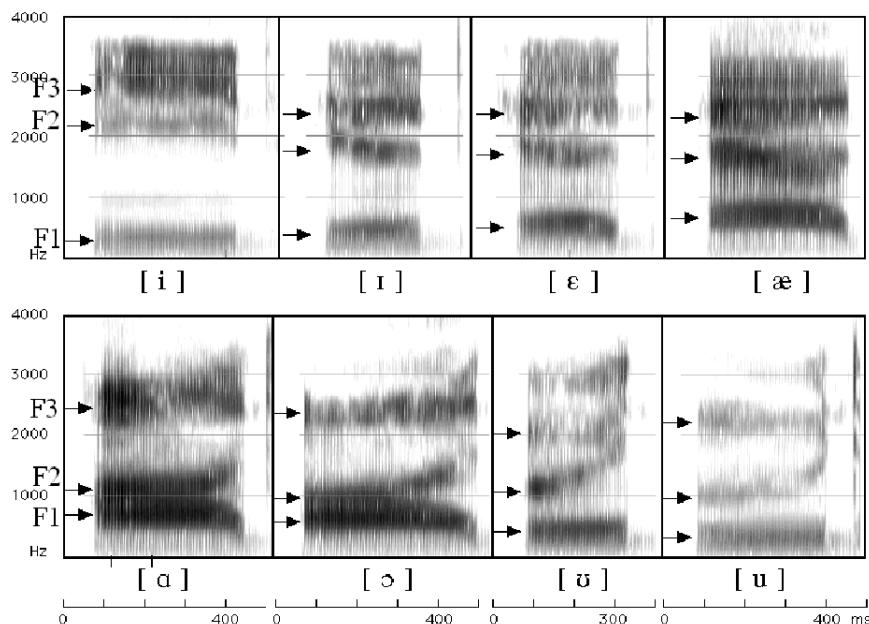


FIGURA 13 – ESPECTROGRAMA DE OITO SONS VOCÁLICOS DO INGLÊS EM POSIÇÃO [h]V[d] (adaptada de LADEFOGED, 2005)

As características descritas acima, pertinentes à fonte (pregas vocais) e ao filtro (a cavidade oral), são fundamentais para a descrição acústica das vogais orais. É importante, neste momento, lembrar algumas características de seu funcionamento.

Fonte e filtro, apesar de constituírem um conjunto em nossa análise, são em larga escala independentes. Isso explica a facilidade com que conseguimos produzir uma mesma vogal com alturas melódicas ou tessituras (*pitch*) completamente diferentes umas das outras. Essas diferenças de *pitch* podem ser atribuídas unicamente às diferenças da onda sonora de entrada acústica no filtro no caso das vogais. Por outro lado, uma mesma frequência fundamental produzida pela fonte, ao encontrar diferentes formas no filtro, apresentará qualidades acústicas totalmente diferentes a depender unicamente dos pontos de constrição encontrados, como observamos anteriormente.

#### 4. ANÁLISE ACÚSTICA DA RADIAÇÃO CARACTERÍSTICA

O último fator importante para o entendimento da Teoria Acústica de Produção da Fala envolve conhecimentos acerca da radiação característica. Esse fator envolve o efeito de filtragem que ocorre quando um som sai da boca e é irradiado se espalhando em todas as direções. Funcionando como um filtro que reduz a energia nas frequências mais baixas e possui menor efeito nas altas, a radiação característica possui comportamento inverso à frequência fundamental. Enquanto a radiação característica aumenta a uma taxa de 6 dB por harmônico, a frequência fundamental cai a uma taxa de 12 dB, totalizando uma queda de 6 dB por harmônico na onda sonora irradiada. Uma vez que a radiação característica é uma constante física do som, sobre a qual não temos controle, optamos pelo não aprofundamento deste aspecto da teoria. A Figura 14 abaixo (KENT; READ, 1992), mostra-nos o processo como um todo.

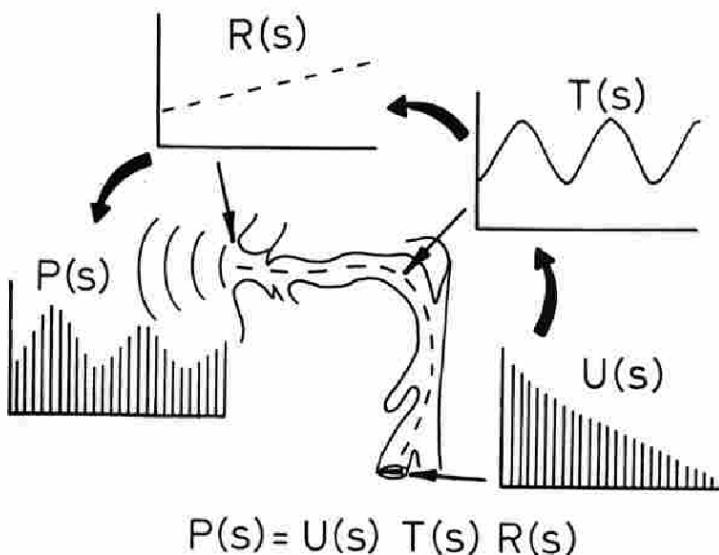


FIGURA 14 – CONCEITO DE PRODUÇÃO DE VOGAIS BASEADO NA TEORIA ACÚSTICA DE PRODUÇÃO DA FALA (KENT; READ, 1992)

A equação diz que a pressão espectral ( $P$ ) irradiada no espaço é o resultado do espectro produzido pelas pregas vocais ( $U$ ), modificada pela função de transferência da configuração do trato vocal ( $T$ ) e pela radiação característica ( $R$ ), com ( $s$ ) indicando frequências diferentes enfatizadas ou reduzidas durante o processo.

## 5. CONCLUSÃO

O uso de dados provenientes da Teoria Acústica de Produção da Fala possibilita a utilização de parâmetros quantitativos onde antes só se podia usar informações qualitativas, de base articulatória, normalmente associadas aos preceitos da Teoria das Vogais Cardeais (JONES, 1976), para a realização das descrições vocálicas. As duas teorias supracitadas podem ser usadas em conjunto para que, dessa maneira, os pontos fortes de cada uma possam ser unidos em uma representação mais esclarecedora no estudo de produções vocálicas das línguas naturais.

Como ponto forte de uma análise vocalica baseada em preceitos da Fonética Acústica, temos que ressaltar o seu caráter quantitativo que, aliado ao rigor metodológico nas aferições e da aplicação de cálculos estatísticos, pode aumentar a validade de um trabalho de caracterização vocalica de uma dada variante. No entanto, a visualização desses dados em tabelas, imagens e gráficos, sem nenhuma dimensão qualitativa, não é simples de ser assimilada.

Ao buscarmos uma maneira de escapar dessa limitação, no que concerne aos estudos sobre vogais, devemos, portanto, aplicar a precisão quantitativa dos estudos acústicos às descrições vocalicas características da Fonética Articulatória. Passamos, dessa forma, a utilizar um sistema que une o melhor das duas teorias em questão, o que facilitará a investigação e a descrição da produção da fala em estudos dessa natureza.

## RESUMO

O objetivo principal do presente artigo é discutir os preceitos fundamentais da Teoria Acústica de Produção da Fala, especialmente no que diz respeito às características acústicas de produção dos sons vocalicos. Discute-se, principalmente, os papéis acústico-articulatórios da fonte e do filtro do sistema acústico, bem como são apresentadas brevemente as características acústicas da radiação. A fonte do sistema acústico – as pregas vocais no caso dos sons vocalicos – produz o ruído periódico necessário à realização de vogais. O filtro do sistema acústico (as cavidades oral e nasal) enfatiza ou suprime determinadas regiões do espectro acústico do som produzido pela fonte do sistema sonoro. As variações do espectro acústico são determinadas pelo nosso controle direto da forma das cavidades supracitadas, o controle das peculiaridades acústicas da radiação característica, por sua vez, não nos é facultado, uma vez que esta ocorre a partir do momento em que a onda sonora deixa o nosso trato vocal.

Palavras-chave: *fonética; acústica; vogais*.

## ABSTRACT

This article's main aim is to discuss the fundamental principles of the Acoustic Theory of Speech production, focusing on the description of the acoustic characteristics of vowel sounds's production. The acoustic-articulatory roles of the source and filter of the acoustic system are discussed, as well as an overview of the acoustic peculiarities of the characteristic radiation. The source of the acoustic system – the vocal folds as regard the vowel sounds – produces the periodic noise necessary to vowels' production. The filter of the acoustic system (the oral and nasal cavities) emphasizes or suppresses certain regions of the acoustic spectrum of the sound produced by the source of the acoustic system. Such variation is determined by our direct control of the aforementioned cavities' shape, control over acoustic peculiarities of the characteristic radiation, on the other hand, is not under our control, once it takes place only after the sound wave leaves the vocal tract.

Keywords: *phonetics; acoustics; vowels.*

## REFERÊNCIAS

- BORDEN, Gloria J.; HARRIS, Katherine S. *Speech science primer: physiology, acoustics, and perception of speech*. Baltimore: Waverly Press, 1980.
- CHIBA, T.; KAJIYAMA, M. *The vowel: its nature and structure*. Tokyo: Tokyo-Kaiseikan Pub. Co., 1941.
- ENTUSA.Ear, Nose, & Throat - USA. *Normal larynx*. 1 fotografia, color. Disponível em: <[http://www.entusa.com/larynx\\_photo.htm](http://www.entusa.com/larynx_photo.htm)>. Acesso em: 04/ 05/2007.
- FANT, Gunnar. *Acoustic theory of speech production: with calculations based on x-ray studies of Russian articulators*. The Hague: Mouton, 1960.
- FRY, Dennis Butler. *The physics of speech*. Cambridge: Cambridge University Press, 1979.
- HONDA, Kiyoshi. Evolution of vowel production studies and observation techniques. *Acoustic Science and Technology*, v. 23, n. 4, p. 189-194, 2002.
- JACKSON, Philip. *Resonances of a closed tube*). 1 figura, p&b. [200-]. Disponível em: <[www.ee.surrey.ac.uk/Teaching/Courses/eem.ssr/slides06.ppt](http://www.ee.surrey.ac.uk/Teaching/Courses/eem.ssr/slides06.ppt)>. Acesso em: 01/05/ 2007.
- JONES, Daniel. *An outline of English phonetics*. 9. ed. Cambridge: Cambridge University Press, 1976.
- KENT, Ray D.; READ, Charles. *The acoustic analysis of speech*. San Diego: Singular, 1992.
- KOO, Owen. *Graph of a sinusoidal sound wave*. 1 figura, p&b. Disponível em: <[http://resources.emb.gov.hk/envir-ed/text/hkissue/e\\_m1\\_4\\_2.htm](http://resources.emb.gov.hk/envir-ed/text/hkissue/e_m1_4_2.htm)>. Acesso em: 04/ 05/ 2007.
- LADEFOGED, Peter. *Elements of acoustic phonetics*. Chicago: The University of Chicago Press, 1962.

\_\_\_\_\_. *A spectrogram of the words heed, hid, head, had, hod, hawed, hood, who'd as spoken by a male speaker of American English*. 1 figura, p&b. 2005. Disponível em: <<http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/course/chapter8/8.3.htm>>. Acesso em: 01/05/ 2007.

MARUSSO, Adriana S. Princípios básicos da teoria acústica de produção da fala. *Revista de Estudos da Linguagem*, v. 13, n. 1, p. 20-43, 2005.

ORION, Gertrude F. *Pronouncing American English: sounds, stress and intonation*. 2. ed. New York: Heinle & Heinle, 1997.

PICKETT, J. M. *The sounds of speech communication: a primer of acoustic phonetics and speech perception*. Baltimore: University Park Press, 1980.

STEVENS, Kenneth N. *Acoustic phonetics*. Cambridge, MA: the MIT Press, 1998.

SUNDBERG, Johan. *Harmonic spectrum of sound from the vocal folds*. 1 figura, p&b. 2005. Disponível em: <<http://hyperphysics.phy-astr.gsu.edu/hbase/music/vocres.html>>. Acesso em: 04/05/2007.

SONDHI, Man Mohan. Resonances of a bent vocal tract. *Journal of the Acoustical Society of America*, v. 79, n. 4, p. 1113-1116, 1986.

THIES, Alexandra. [Função de transferência do trato vocal]. 1 figura, color. 2005a. Disponível em: <[http://www.spectrum.uni-bielefeld.de/~thies/HTHS\\_WiSe2005-06/session\\_05.html](http://www.spectrum.uni-bielefeld.de/~thies/HTHS_WiSe2005-06/session_05.html)>. Acesso em: 30/04/2007.

TRASK, R. L. *A Dictionary of Phonetics and Phonology*. London: Routledge, 1996.

Submetido em: 30/09/2009

Aceito em: 29/04/2010