

IncurSIONando pelos domínios da eletroglotografia: proposta de um corpus EGG

Juliano S. de Mattos, Dirceu G. da Silva, José A. Apolinário Jr. e Edson Cataldo

Resumo—Este artigo apresenta um resumo teórico sobre o assunto eletroglotografia e uma proposta de um *corpus*, composto de sinais gravados em um eletroglotógrafo, sincronizado com os sinais de voz correspondentes. Trata-se de uma primeira ferramenta didática para o assunto, desenvolvida em conjunto com o IME, a UFF e a PUC-Rio, que visa disponibilizar sinais EGG masculinos e femininos para uso acadêmico.

Palavras-Chave—Eletroglotografia, base sinais, processamento de sinais de voz.

Abstract—This paper presents a theoretical review about the topic electroglottography and proposes a corpus with real signals recorded from an electroglottograph synchronized with the corresponding speech signals. This is a first didactical tool for this subject, jointly developed by IME, UFF, and PUC-Rio, aiming to make available EGG signals, both male and female, for academic use.

Keywords—Electroglottography, corpus, speech signal processing.

I. INTRODUÇÃO

A eletroglotografia é um método não invasivo introduzido em [1], que estima a variação da área de contato entre as cordas vocais (*vocal fold contact area - VFCA*) durante a produção da voz [2], [3], [4] e vem sendo utilizado há décadas para fins clínicos e de pesquisa [1]. Resumidamente, o eletroglotógrafo mede as variações da impedância elétrica causadas pela variação da área de contato entre as cordas vocais, uma vez que estas vibram. A impedância é medida através de um par de eletrodos, presos ao pescoço do locutor, que aplicam uma pequena corrente elétrica ao local.

Quando as cordas vocais estão fechadas, há uma baixa impedância à passagem da corrente elétrica. Quando as cordas vocais estão separadas, devido ao fluxo de ar que as atravessa, a impedância da laringe é alta. Portanto, a impedância da laringe varia de acordo com a área de contato das cordas vocais [2], [5].

Juliano S. de Mattos : Departamento de Engenharia de Telecomunicações, Universidade Federal Fluminense (UFF), Escola de Engenharia, Rua Passo da Pátria, 156 - São Domingos, 24.210-240, Niterói, RJ - Brasil. Edson Cataldo: Departamento de Matemática Aplicada, Programa de Pós-Graduação em Engenharia de Telecomunicações, Universidade Federal Fluminense (UFF), Escola de Engenharia, Rua Passo da Pátria, 156 - São Domingos, 24.210-240, Niterói, RJ - Brasil. Dirceu G. da Silva: CETUC/PUC-Rio, Rua Marquês de São Vicente, 225, Gávea, 22.453-900, Rio de Janeiro, RJ - Brasil. José A. Apolinário Jr.: Departamento de Engenharia Elétrica (SE/3), Instituto Militar de Engenharia (IME), Praça General Tibúrcio, 80, Urca, 22.290-270, Rio de Janeiro, RJ - Brasil. E-mails: juliano_santiago@yahoo.com.br, dirceu@cetuc.puc-rio.br, apolin@ime.ub.br e ecataldo@im.uff.br.

Diversos estudos comparativos foram realizados utilizando fotografia estroboscópica [22], [7], [8], [9], vídeo estroboscópico [10], [11], imagens de alta velocidade [12], [13], [14], fotoglottografia [15], [16], [17], [18], [19], [20], medidas de pressão subglotal [19] e filtragem inversa [21], [22], [23], [24]. Todos esses estudos confirmaram que o sinal do eletroglotógrafo (sinal EGG) está relacionado com a área de contato das cordas vocais: quanto maior o contato da superfície, maior a admitância medida [2].

Em adição à resistência em torno do pescoço, a medida da impedância também é influenciada pela reatância (capacitiva ou indutiva) da carga examinada. Uma capacitância variável pode hipoteticamente existir na glote quando as duas cordas vocais estão separadas por uma fina camada de ar, como suposto por [23]. Esta hipótese pode ser verificada alterando a frequência da corrente alternada usada para medição da impedância: a corrente permanecerá a mesma apenas se a carga for puramente resistiva. De acordo com [25], a impedância é essencialmente resistiva em uma ampla faixa de frequência [5].

O sinal eletroglotográfico (EGG) resultante, o eletroglotograma, mostra a variação, nas cordas vocais, da impedância em função do tempo. A impedância também varia consideravelmente com os tipos de pele e com os movimentos verticais da laringe. Filtros passa-altas são usados para eliminar as interferências de baixa frequência e extrair apenas as variações causadas pela vibração das cordas vocais. Ademais, um controle automático de ganho também foi construído, dentro do Eletroglotógrafo, para manter um nível de sinal apropriado a despeito de consideráveis variações de impedância entre locutores. Essas técnicas causam distorções de fase e amplitude que podem influenciar a forma de onda do sinal EGG [25]. Conseqüentemente, o sinal do eletroglotógrafo não deve ser considerado uma medida absoluta do contato entre as cordas vocais, e certos cuidados devem ser tomados quando da interpretação do sinal. A despeito dessas limitações, o sinal EGG é uma informação útil sobre a situação das cordas vocais durante a fonação. O estudo da derivada do sinal EGG também é importante, pois possibilita a determinação dos instantes de abertura e fechamento glotal, do quociente de abertura e uma estimativa precisa da frequência fundamental do sinal [2].

II. O ELETROGLOTÓGRAFO

O modelo EG2-PCX fabricado pela *Glottal Enterprises* foi o eletroglotógrafo cedido pelo Instituto Militar de Engenharia (IME) para ser utilizado na construção de uma base de dados (*corpus*) de sinais EGG.

O modelo EG2-PCX (Figura II) possui um simulador de laringe (LS-1) e um compensador de fase (C-1) que podem ser conectados ao equipamento. O simulador de laringe, como o próprio nome sugere, é usado no lugar dos eletrodos para simular as pequenas variações de resistência do pescoço medidas pelo o eletroglotógrafo durante a produção da voz. O LS-1 é um auxílio à medição e à verificação da performance dos equipamentos, uma vez que pode ser conectado em qualquer unidade do fabricante, possibilitando um comparação objetiva entre eletroglotógrafos. O C-1 foi desenvolvido para minimizar as distorções de fase inseridas pela filtragem passa-alta e pode ser utilizado em conjunto com o LS-1 quando a aquisição de dados for realizada pela placa de som de um computador. As Figuras 2 e 3 ilustram os referidos acessórios, simulador de laringe e compensador de fase, respectivamente.



Fig. 1. Eletroglotógrafo EG2-PCX fabricado pela Glottal Enterprises.

Existem outros modelos de eletroglotógrafo que estimam, além do VFCA, o IVFCA (*Inverse vocal fold contact area*), a derivada do sinal EGG na polarização VFCA, conhecida como DEGG (*Differentiated EGG*), e o EXT LF (*Extended Low Frequency response*), que permite a observação das componentes de baixa frequência dos movimentos da laringe.



Fig. 2. Simulador de laringe: simula as pequenas variações de resistência do pescoço medidas pelo o eletroglotógrafo.

Comparando a forma de onda do sinal EGG com imagens de alta velocidade, foi relatado em [21] que o fechamento glotal ocorre no instante no qual a derivada do sinal EGG (DEGG) possui seu máximo absoluto. Estes picos do sinal DEGG são claramente identificados, conforme Figura 4.

Esta abordagem foi corroborada por [2], que considerou os picos do DEGG como indicadores de abertura e fechamento glotal. As Figuras 4 e 5 ilustram esta relação. Entretanto, em alguns casos, podem ocorrer picos imprecisos, duplos e até a ausência de picos.

III. PARÂMETROS

Os parâmetros que podem ser obtidos do sinal do eletroglotógrafo são os instantes de fechamento e abertura glotal,



Fig. 3. Compensador de fase: desenvolvido para minimizar as distorções inseridas pelo filtro passa-altas.

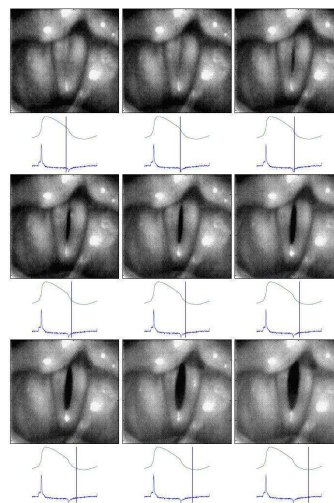


Fig. 4. Visualização da abertura por cinematografia ultra-rápida e eletroglotografia (locutor em fonação normal com frequência fundamental igual a 110Hz - sinais EGG e DEGG) [2]. A linha vertical nas formas de onda corresponde ao instante em que a imagem foi retratada.

encontrados a partir do sinal DEGG, e os instantes de máximo fechamento e máxima abertura encontrados a partir do sinal EGG. A Figura 6 (a),(b) e (c) apresentam um trecho de um sinal de voz de uma vogal /a/, o sinal EGG e o sinal DEGG, respectivamente.

A. Instante de abertura

O instante de abertura é definido como o instante em que o sinal EGG está próximo ao seu valor mínimo. Esse instante ocorre quando o sinal DEGG (derivada do sinal EGG) atinge seu menor valor absoluto, conforme Figuras 4 e 7. Fisiologicamente, corresponde ao instante no qual as cordas vocais iniciam sua separação (início da redução da área de contato entre as cordas vocais).

B. Instante de fechamento

O instante de fechamento é definido como o instante em que o sinal EGG está próximo ao seu valor máximo. Esse instante do sinal EGG ocorre quando o sinal DEGG atinge seu maior valor absoluto, conforme as Figuras 5 e 7. Fisiologicamente,

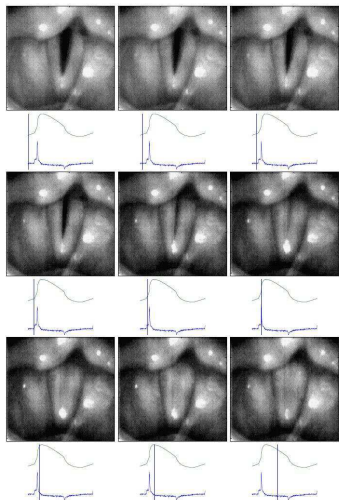


Fig. 5. Visualização do fechamento por cinematografia ultra-rápida e eletroglotografia (locutor em fonação normal com frequência fundamental igual a 110Hz - sinais EGG e DEGG) [2]. A linha vertical nas formas de onda corresponde ao instante em que a imagem foi retratada.

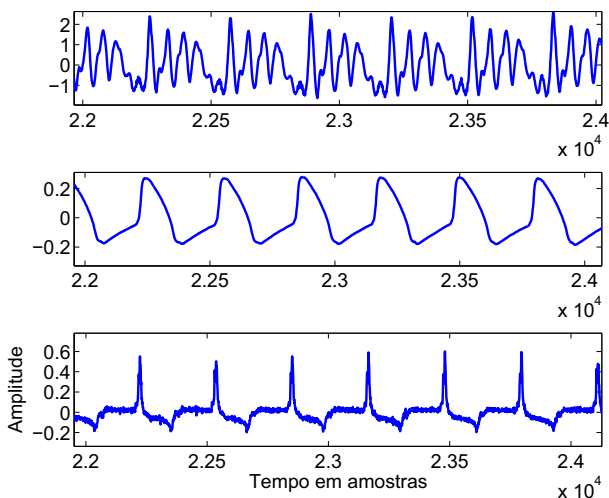


Fig. 6. Sinal de voz, sinal EGG e sinal DEGG.

corresponde ao instante que as cordas vocais se fecham (aumento da área de contato entre as cordas vocais).

C. Instante de máximo fechamento e de máxima abertura

Define-se o instante de máximo fechamento como o momento de maior área de contato entre as cordas vocais, que ocorre após o instante de fechamento, quando o sinal EGG atinge seu maior valor absoluto como mostrado na Figura 7. Analogamente, o instante de máxima abertura é definido como o momento de menor área, que ocorre após o instante de abertura, quando o sinal EGG atinge seu menor valor absoluto.

IV. CUIDADOS DURANTE O PROCESSO DE GRAVAÇÃO

O sinal EGG é captado por dois eletrodos posicionados, apropriadamente, próximos às cordas vocais. O eletrodo que

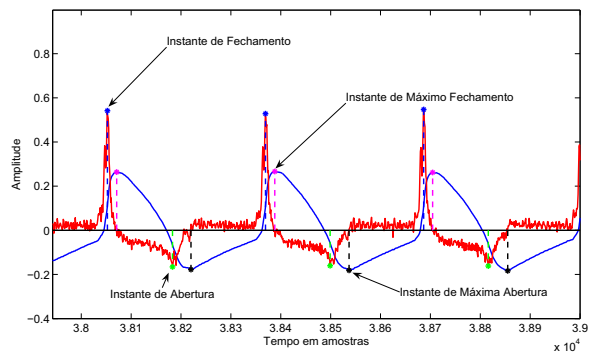


Fig. 7. Sinais EGG e DEGG com seus instantes de fechamento, abertura e os instantes de máximo fechamento e abertura.

deverá ser posicionado no lado esquerdo do pescoço do locutor possui dois fios na cor vermelha. Existem duas placas douradas em cada eletrodo, separadas por uma tarja de material isolante que em contato com o pescoço, deverá permanecer na posição horizontal (tarja paralela ao velcro que prende os eletrodos ao pescoço). O gel que acompanha o EG2-PCX deve ser colocado nas placas douradas para facilitar a passagem da corrente elétrica.

O indicador no painel frontal do eletroglotógrafo que auxilia no posicionamento dos eletrodos e prevê uma indicação quantitativa do movimento vertical da laringe está representado na Figura 8. O ideal é que durante a fala um ou mais *leds* verdes estejam acessos; caso contrário, os eletrodos deverão ser reposicionados. Este teste deverá ser realizado antes do início das gravações sob pena de comprometer a qualidade do sinal obtido.

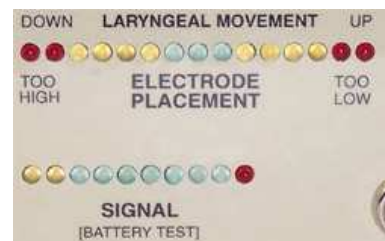


Fig. 8. Auxílio ao posicionamento dos eletrodos e indicação quantitativa do movimento vertical da laringe.

Após o correto posicionamento dos eletrodos, o microfone que captará o sinal de voz também deverá ser conectado ao eletroglotógrafo (parte traseira), que apresentará em sua saída o sinal EGG e o sinal da voz sincronizados, porém separados. A saída do eletroglotógrafo deverá ser conectada a um computador, via porta *USB* ou cabo estéreo, que através de um *software* de voz efetuará as gravações dos sinais no disco rígido ou em outra unidade previamente selecionada. Caso a conexão seja efetuada com o cabo estéreo, o canal esquerdo fornecerá o sinal de áudio (microfone) e o direito o sinal EGG.

Neste artigo o *software* de voz escolhido foi o *Audacity* e a conexão entre o eletroglotógrafo e o computador foi realizada

através da porta *USB*.

V. A BASE DE DADOS

Os sinais EGG e de voz obtidos foram gravados com uma taxa de amostragem de 44.100 Hz, com 16 bits de resolução e no formato PCM. As gravações foram realizadas em uma câmara acústica no Laboratório de Voz do Instituto Militar de Engenharia (IME). A Figura 9 ilustra o processo de gravação. A base de dados é formada por frases, sendo 10 balanceadas foneticamente para o português falado no Rio de Janeiro [26], 5 palavras e 28 frases de interesse para a perícia forense, elaboradas com a colaboração dos peritos do Instituto de Criminalística Carlos Éboli (ICCE, RJ) e 8 vogais sustentadas. Todas as frases, palavras e vogais sustentadas foram gravadas por 5 locutores do sexo masculino e 5 locutores do sexo feminino, com idades relacionadas na Tabela I. A referência [28] também disponibilizou uma base de sinais gravados com eletroglotógrafo, contendo 15 locutores, que pode ser encontrada no servidor <ftp://ftp.cs.keele.ac.uk>.

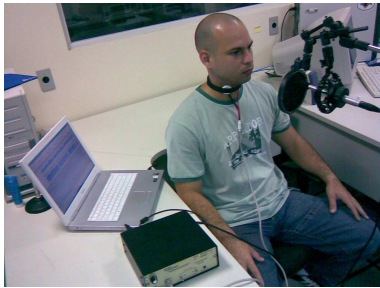


Fig. 9. Foto que simula o processo de gravação do sinal EGG (VFCA), realizado no Laboratório de Voz do IME.

A Tabela I representa a faixa etária dos locutores usados na formação da base de dados e duração total aproximada das gravações.

TABELA I

FAIXA ETÁRIA DOS GRUPOS DE LOCUTORES E A DURAÇÃO TOTAL APROXIMADA DA GRAVAÇÃO.

	Idade	Duração
Homens	20-38	75-105s
Mulheres	18-28	70-95s

As Tabelas II, III, IV e V, contém as frases, palavras e vogais sustentadas que compõem a base de dados. A base de dados deste trabalho está disponível em www.ime.br/~labvoz/.

VI. CÁLCULO DA FREQUÊNCIA FUNDAMENTAL USANDO A BASE DE DADOS

Os picos de maior valor absoluto do sinal DEGG foram utilizados para o cálculo da frequência fundamental (*pitch*) do sinal de voz. Os resultados foram comparados com a técnica de extração da *pitch* amplamente conhecido como *fxrapt* [27]. O sinal de voz usado na simulação foi uma vogal sustentada /a/. Os resultados das simulações mostraram que não houve diferenças significativas entre os valores de *pitch* calculados

TABELA II

FRASES FONETICAMENTE BALANCEADAS PARA O PORTUGUÊS FALADO NO RIO DE JANEIRO.

	Frases Balanceadas
1	Eu vi logo a Iôidô e o Léo.
2	Um homem não caminha sem um fim.
3	Vi Zé fazer essas viagens seis vezes.
4	O atabaque do Tito é coberto com pele de gato.
5	Ele lê no leito de palha.
6	Paira um ar de arara rara no rio Real.
7	Foi muito difícil entender a canção.
8	Depois do almoço te encontro.
9	Esses são nossos times.
10	Procuerei Maria na copa.

TABELA III

FRASES DE INTERESSE PARA PERÍCIA FORENSE.

	Frases - Perícia Forense
1	Aaa tá, amanhã eu ligo.
2	Aaa tá, tamo junto irmão.
3	Alô, quem fala?
4	Alô, alô, alô!
5	Alô, quer falar com quem?
6	Eu digo alô baixinho.
7	Cadê a parada?
8	Eu digo parada baixinho.
9	Esse bagulho é bom.
10	Cadê o bagulho?
11	Eu digo bagulho baixinho
12	Cara, cadê você?
13	Cara, amanhã faço a entrega.
14	Eu digo cara baixinho.
15	Café ou açúcar?
16	Eu quero dois gramas de açúcar.
17	Eu quero dez quilos de açúcar e dois quilos de café.
18	Eu digo café baixinho.
19	Eu digo açúcar baixinho.
20	Fala pra ele que to bolado com essa parada.
21	Tô bolado.
22	Eu digo bolado baixinho.
23	Eu digo copa baixinho.
24	Eu digo tudo baixinho.
25	Iiiiiii, isso não vai dar certo.
26	Tá tudo dominado no Turano
27	Eeeeeeeee, acho que sim.
28	Pede pra sair.

pelos *fxrapt* e pelos instantes de fechamento do sinal EGG (picos de maior valor absoluto do sinal DEGG). A Figura 10 ilustra os resultados encontrados.

VII. CONCLUSÕES

Neste artigo foi apresentada uma pequena introdução à eletroglotografia e um *corpus*, composto de sinais gravados em um eletroglotógrafo, sincronizado com os sinais de voz correspondentes. As aplicações práticas dos sinais EGG e DEGG que podem ser destacadas são: auxílio à detecção de patologias nas cordas vocais, modelagem do sinal de voz através dos parâmetros extraídos do sinal EGG, algoritmos para fins acadêmicos, determinação de parâmetros para reconhecimento de locutor. Os valores de frequência fundamental encontrados usando o sinal DEGG foram similares aos calculados pelo *fxrapt*, ratificando o valor da base e suas aplicações.

TABELA IV
PALAVRAS DE INTERESSE PARA PERÍCIA FORENSE.

Palavras - Perícia Forense	
1	[pataká]
2	[peteké]
3	[pitikí]
4	[potokó]
5	[putukú]

TABELA V
VOGAIS SUSTENTADAS.

Vogais Sustentadas	
1	a
2	ã
3	é
4	ê
5	i
6	ó
7	ô
8	u

REFERÊNCIAS

[1] Fabre, P., "Sphygmographie par simple contact d'électrodes cutanées, introduisant dans l'artère de faibles courants de haute fréquence détecteurs de ses variations volumétriques", *Comptes Rendus Soc Biol*, vol. 133, pp. 639-641, 1940.

[2] Nathalie Henrich, "Etude de la source glottique en voix parlée et chantée : modélisation et estimation, mesures acoustiques et électroglottographiques, perception", Thèse de doctorat de l'Université Paris 6, PhD thesis, pp. 87-96, 2001.

[3] Colton, R. H. e Conture, E. G., "Problems and pitfalls of electroglottography," *Journal of Voice*, vol. 4, no. 1, pp. 10-24, 1990.

[4] Baken, R. J., "Electroglottography," *Journal of Voice*, vol. 6, no. 2, pp. 98-110, 1992.

[5] Pulakka H., "Analysis of Human Voice Production Using Inverse Filtering, High-Speed Imaging, and Electroglottography," Helsinki University of Technology, Dept. of Computer Science and Engineering, 2005.

[6] Fourcin, A. J., "Laryngographic examination of vocal fold vibration, in Ventilatory and Phonatory Function," edited by B. Wyke Oxford University Press, London, pp. 315-326, 1974.

[7] Lecluse, F. L. E., Brocaar, M. P., and Verschuure, J., "The electroglottography and its relation to glottal activity," *Folia Phoniatr.* 27, pp. 215-224, 1975.

[8] Pedersen, M. F., "Electroglottography compared with synchronized stroboscopy in normal persons," *Folia Phoniatr.* 29, pp. 191-199, 1977.

[9] Teaney, D., and Fourcin, A. J., "The electrolaryngography as a clinical tool for the observation and analysis of vocal fold vibration," *The Voice Foundation*, 1980.

[10] Anastaplo, S., and Karnell, M. P., "Synchronized videostroboscopic and electroglottographic examination of glottal opening," *J. Acoust. Soc. Am.* 83, pp. 1883-1890, 1988.

[11] Karnell, M. P., "Synchronized videostroboscopy and electroglottography," *J. Voice* 3, pp. 68-75, 1989.

[12] Childers, D. G., Hicks, D. M., Moore, G. P., Eskenazi, L., and Lalwani, A.L., "Electroglottography and vocal fold physiology," *J. Speech, Hear. Res* 33, pp. 245-254, 1990.

[13] Childers, D. G., and Krishnamurthy, A. K., "A critical review of electroglottography," *CRC Crit. Rev. Biomed. Eng.* 12, pp. 131-161, 1985.

[14] Childers, D. G., and Larar, J. N., "Electroglottography for laryngeal function assessment and speech analysis," *IEEE Trans. Biomed. Eng.* BME-31, pp. 807-817, 1984.

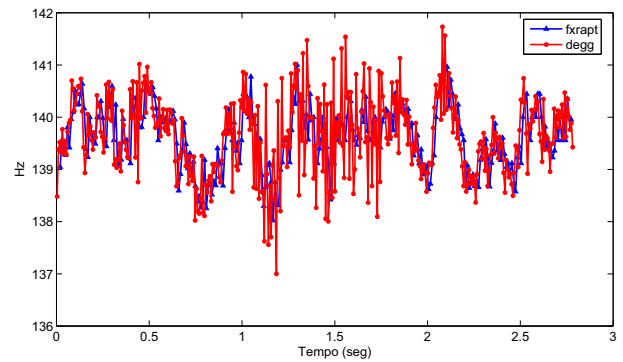


Fig. 10. Freqüência fundamental extraída pelo DEGG e pelo algoritmo do fxrapt.

[15] Baer, T., Lofqvist, A., and McGarr, N. S., "Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques," *J. Acoust. Soc. Am.* 73, pp. 1304-1308, 1983.

[16] Berke, G. S., Moore, D. M., Hantke, D. R., Hanson, D. G., Gerratt, B. R., and Burstein, F., "Laryngeal modeling: Theoretical, in vitro, in vivo," *Laryngoscope* 97, pp. 871-881, 1987.

[17] Dejonckere, P., "Comparison of two methods of photoglottography in relation to electroglottography," *Folia Phoniatr.* 33, pp. 338, 1981.

[18] Gerratt, B. R., Hanson, D. G., and Berke, G. S., "Laryngeal configuration associated with glottography," *Am. J. Otolaryngol.* 9, pp. 173-17, 1988.

[19] Kitzing, P., "Photo- and electroglottographical recording of the laryngeal vibratory pattern during different registers," *Folia Phoniatr.* 34, pp. 234-241, 1982.

[20] Titze, I. R., Baer, T., Cooper, D., and Scherer, R., "Automatic extraction of glottographic waveform parameters and regression to acoustic and physiologic variables," in *Vocal Fold Physiology: Contemporary Research Clinical Issues*, edited by A. J. Bless DM College Hill, San Diego, pp. 146-154, 1984.

[21] Childers, D. G., Naik, J. M., Larar, J. N., Krishnamurthy, A. K., and Moore, G. P., "Electroglottography, speech and ultra-high speed cinematography," in *Vocal Fold Physiology and Biophysics of Voice*, edited by I. Titze and R. Scherer Denver Center for the Performing Arts, Denver, pp. 202-220, 1983.

[22] Fourcin, A. J., "Laryngographic assessment of phonatory function," *ASHA Rep.* 11, pp. 116-124, 1981.

[23] Rothenberg, M., "Some relations between glottal air flow and vocal fold contact area," *ASHA Rep.* 11, pp. 88-96, 1981.

[24] Rothenberg, M., and Mahshie, J. J., "Monitoring vocal fold abduction-through vocal fold contact area," *J. Speech Hear. Res.* 31, pp. 338-351, 1988.

[25] Scherer, R. C., Druker, D. G. and Titze, I. R., "Electroglottography and direct measurement of vocal fold contact area," in O. Fujimora, ed., "Vocal Physiology: Voice Production, Mechanisms and Functions," Raven Press, New York, pp. 279-291, 1988.

[26] Alcaim, A., Solewicz, J. A. e Moraes, J. A. Freqüência de Ocorrência dos Fones e Lista de Frases Foneticamente Balanceadas no Português Falado no Rio de Janeiro. *Revista da Sociedade Brasileira de Telecomunicações*, v. 7, n. 1, dez 1992.

[27] Talkin, "A Robust Algorithm for Pitch Tracking (RAPT) in Speech Coding and Synthesis," W. B. Kleijn, K. K. Paliwal eds., Elsevier ISBN 0444821694, 1995.

[28] F. Plante, "A Pitch Extraction Reference Database," ESCA EUROSPEECH 95, 4th European Conf. on Speech Communication and Technology, 1995.