

## O PAPEL DAS VOGAIS REDUZIDAS PÓS-TÔNICAS NA CONSTRUÇÃO DE UM SISTEMA DE SÍNTESE CONCATENATIVA PARA O PORTUGUÊS DO BRASIL \*

Patrícia A. de AQUINO

**RESUMO** *Nesta dissertação, desenvolvida no LAFAPE (Laboratório de Fonética Acústica e Psicolinguística Experimental) foram feitas análises fonético-acústicas em que se compararam as vogais tônicas e pós-tônicas do português do Brasil. Essas análises se propõem, em primeiro lugar, a fornecer dados sobre o fenômeno da redução vocálica em PB, contribuindo para as discussões desse fenômeno, tanto no âmbito da Fonética, quanto no da Fonologia. Em segundo lugar, o presente estudo vem corroborar uma decisão prévia, tomada na elaboração de um corpus de unidades utilizado no sistema de síntese concatenativa do LAFAPE, de gravar unidades diferentes para concatenar vogais tônicas e pós-tônicas. Tal decisão é decorrente da existência de uma forte relação entre prosódia e segmentos. Sabe-se que ambientes prosódicos fracos favorecem a redução dos segmentos, enquanto os ambientes prosodicamente fortes a desfavorecem. Diferenciando as vogais tônicas (que, geralmente, ocorrem em ambientes fortes) das pós-tônicas (que, geralmente, ocorrem em ambientes fracos) consegue-se, já no nível segmental, introduzir algumas características prosódicas da fala natural. Foi corroborada também, na maioria dos casos, a opção por utilizar unidades demissilábicas e intersilábicas. Há, entretanto, um resultado que aponta para a necessidade de unidades maiores: percebeu-se uma significativa influência da consoante precedente sobre as vogais pós-tônicas. A fim de contemplar tal influência, talvez seja necessário gravar unidades maiores, evitando a segmentação no interior dessas vogais.*

**ABSTRACT** *In this thesis, developed at LAFAPE (Laboratório de Fonética Acústica e Psicolinguística Experimental), phonetic-acoustical analysis of stressed and post-stressed vowels of Brazilian Portuguese were made. The purposes of the analysis were: 1) to give information about the vowel reduction phenomenon in BP, and contribute to the discussion about this phenomenon in the domains of both Phonetics and Phonology. 2) to test a previously made decision to differentiate stressed from post-stressed vowels for recording the synthesis system unit dictionary. The reason for this decision is the existence of a strong relation between prosody and segments. Weak prosodic*

---

\* Texto resultante da Dissertação de Mestrado apresentada ao Curso de Linguística do Instituto de Estudos da Linguagem, da Unicamp, no dia 17 de dezembro de 1997, sob a orientação da Professora Dra. Eleonora Cavalcante Albano.

*environments favor segment reduction, and strong prosodic environments do not. By differentiating the stressed vowels (which, generally, occur in strong environments) from post-stressed vowels (which, generally, occur in weak environments) we have been able to introduce some prosodically conditioned characteristics of natural speech at the segmental level. In the majority of cases, the option to utilize demysyllabic and intersyllabic units was also supported. There is, however, one result that shows the need for longer units: there is a significant influence of the precedent consonant on the post-stressed vowels. To take this influence into account, it will probably be necessary to record longer units, avoiding segmentation inside those vowels.*

Esta dissertação é constituída, basicamente, de duas partes: a primeira delas contém, além de uma breve resenha sobre a história e o funcionamento de sistemas de síntese de fala, o detalhamento das etapas da elaboração do corpus de unidades do sistema e síntese de fala concatenativa que vem sendo desenvolvido no LAFAPE (Laboratório de Fonética Acústica e Psicolingüística Experimental) do IEL/UNICAMP.

A segunda parte consiste em um estudo fonético-acústico comparativo entre as vogais pós-tônicas e tônicas da fala de um informante do português do Brasil (doravante PB), estudo esse que tinha como um de seus objetivos avaliar algumas decisões previamente tomadas durante a seleção das unidades concatenantes que fariam parte do corpus do sistema de síntese de fala do LAFAPE.

Em busca de uma fala sintética de boa qualidade, a principal decisão tomada durante a elaboração do corpus do sistema de síntese foi a de contemplar, na medida do possível, as principais interações existentes entre a prosódia e os segmentos.

Do ponto de vista dinâmico (em modelos como a Fonologia Articulatória Browman & Goldstein (1986)), tanto o ritmo como a informação segmental são produzidos por oscilações de maior ou menor força. Assim, é natural que a estrutura prosódica afete diretamente o conteúdo segmental. Segmentos plenos são produzidos em posições prosodicamente fortes, emitidas com maior força, e segmentos reduzidos são produzidos em posições prosodicamente fracas, emitidas com menor força.

Essa visão levou à escolha do método concatenativo (em oposição ao paramétrico) na medida em que algumas dessas interações podem ser captadas na escolha de unidades concatenantes.

Utilizaram-se, então, unidades que contemplam as diferenças fonéticas existentes entre os segmentos plenos e os reduzidos.

Quanto às consoantes, foram tratadas as de onset silábico como segmentos plenos e as de coda como segmentos reduzidos.

Quanto às vogais, as tônicas e pré-tônicas foram tratadas como segmentos plenos, enquanto as vogais pós-tônicas foram tratadas como segmentos reduzidos.

Acreditava-se que as vogais, por serem os núcleos silábicos e por serem extremamente afetadas pela prosódia, muito mais do que as consoantes, mereciam um estudo especial, que levasse em consideração a interação entre a prosódia e o segmento, para que se garantisse uma fala sintética de boa qualidade.



Um estudo tão específico sobre as vogais ainda não havia sido feito quando se tornaram viáveis as possibilidades de iniciar a gravação do corpus de unidades do sistema de síntese. No entanto, não se viu problema em dar início a tais gravações sem tal estudo prévio, pois acredita-se que a principal função de um sistema de síntese acadêmico é fornecer retorno às pesquisas lingüísticas de base, avaliando, através de seu resultado – a fala sintética – as decisões lingüísticas.

Na falta de tal estudo, fez-se um estudo-piloto, cujos resultados estão de acordo com a intuição de Câmara Jr (1970 e 1971), que propôs que as vogais tônicas e pré-tônicas não se diferenciam, enquanto há uma grande diferenciação das pós-tônicas em relação ao conjunto daquelas. Essas, as pós-tônicas, sofreriam a chamada “redução vocálica”.

Foram gravadas, então, unidades contendo vogais tônicas que seriam usadas para concatenar tanto as vogais tônicas quanto as pré-tônicas e unidades contendo vogais pós-tônicas, diferenciando as pós-tônicas das demais.

O corpus resultante, por ter sido cuidadosamente balanceado, pôde constituir o próprio corpus para o estudo fonético-acústico sobre as vogais que se pretendia realizar.

Esse estudo passou, então, a ser feito. Foram medidos, manualmente, através do algoritmo FFT<sup>1</sup>, no CSL4300 da Kay, os quatro primeiros formantes e a duração das trinta ocorrências de cada vogal. Essas ocorrências são devidas à variação da consoante precedente, pois cada vogal havia sido gravada com cada uma das consoantes e com cada um dos grupos consonantais de onset.

Acresceu-se ao corpus inicial o conjunto das vogais /a/s pós-tônicas não finais, como forma de garantir que esse corpus ficasse simétrico.

Os valores obtidos foram submetidos a análises estatísticas e, através dos resultados obtidos nessas análises, pôde-se não só avaliar as decisões tomadas em relação à seleção de unidades para o sistema de síntese do LAFAPE, como também, chegar a algumas conclusões sobre a redução vocálica.

A seguir vêm discutidas algumas teorias a respeito da redução vocálica.

A primeira teoria relevante é a de Câmara Jr. (1971), que mostra dois aspectos da redução vocálica, considerando-os dois aspectos do mesmo fenômeno:

- Redução do número de fonemas (Neutralização)
- diferente realização fonética

Vemos, no entanto, dois motivos para considerar esses aspectos como dois fenômenos distintos; o primeiro deles é o fato de a vogal /a/ não ser afetada pela neutralização, mas ser afetada pela redução propriamente dita. O segundo é o fato de a neutralização corresponder a um processo categórico, demonstrado pela redução do número de fonemas conforme a tonicidade no quadro a seguir:

---

<sup>1</sup> O método FFT (Fast Fourier Transform) utiliza a transformada de Fourier, que decompõe o sinal acústico em todas as suas componentes harmônicas. Os maiores picos correspondem às faixas de frequência em que os harmônicos passam sem sofrer atenuações pelo trato vocal (que atua como um filtro passa-faixas). Cada concentração desses harmônicos corresponde à região de um formante.

Vogais tônicas:

altas	i		u
médias (2º.grau)		e	o
médias (1º.grau)		ε	ɔ
baixa		a	

Vogais pré-tônicas:

altas	i		u
médias		E	O
baixa		a	

Vogais pós-tônicas penúltimas:

altas	i		U
média		E	
baixa		a	

Vogais pós-tônicas finais:

altas	I		U
baixa		a	

Figura I - Triângulos vocálicos conforme a tonicidade.

A redução, por outro lado, se apresenta como um processo gradiente, conforme se observa no gráfico a seguir.

F1 X F2 para todas as vogais

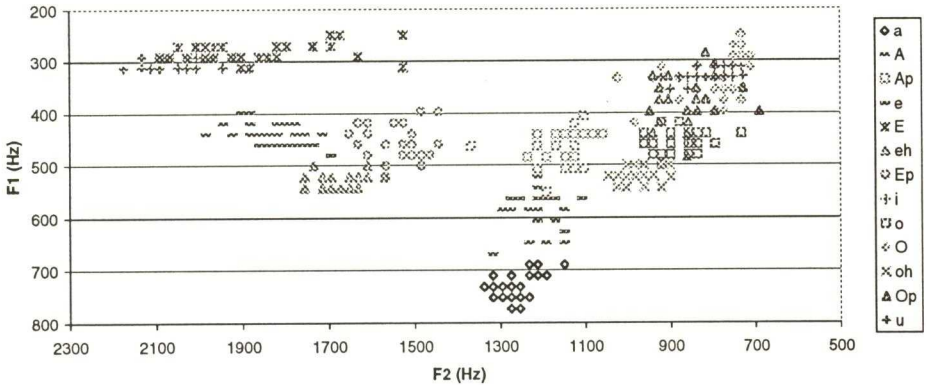


Figura II – F1 e F2 (em Hz) para todas as vogais.

A segunda é a de Lindblom (1963), que considera a existência de um alvo: a vogal que o falante tem intenção de realizar seria a tônica. Esse alvo seria mais facilmente atingido quanto maior a duração da vogal. Quanto menor a duração, mais dificilmente tal alvo seria atingido, ocorrendo o undershoot.

Infere-se dessas afirmações que quanto maior a duração de uma vogal menor a dispersão que suas realizações deveriam ter em relação ao alvo.

Poder-se-ia supor, conforme a teoria target undershoot de Lindblom, que as vogais tônicas, por serem significativamente mais longas do que as pós-tônicas (resultado comprovado pelos valores obtidos através das medidas efetuadas neste trabalho), levam maior tempo para atingirem o alvo, o que acarretaria uma menor dispersão das tônicas do que das pós-tônicas.

Essa teoria não explica, no entanto, o fato de a vogal /a/ ser, juntamente com a vogal /oh/, a mais longa, comparativamente às outras vogais e, mesmo assim, apresentar, conforme se vê na figura II acima, uma dispersão maior do que as demais vogais tônicas.

Outro fator de que essa teoria não dá conta é o de não haver uma maior dispersão dos valores das pós-tônicas circundando o valor que corresponderia à vogal-alvo. Segundo a teoria, o que se esperaria obter, para cada vogal, seria uma figura semelhante à seguinte:

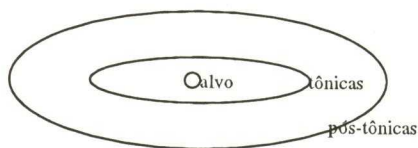


Figura III- Visualização do que ocorreria com as vogais tônicas e pós-tônicas na teoria target undershoot de Lindblom.

No gráfico da figura II ( $F1 \times F2$  para todas as vogais), as únicas vogais que apresentam dispersões semelhantes às da figura acima são as posteriores [u] (tônica), [O] (pós-tônica final) e [Op] (pós-tônica não-final). Admitindo que isso realmente ocorresse, haveria o problema de estabelecer o /u/, e não mais o /o/, como vogal-alvo para as reduzidas [Op] e [O]. Observando as vogais anteriores e, principalmente, as centrais, não parece haver uma variação de ocorrências ao redor do alvo, mas, que se trata, na verdade, de alvos diferentes.

Interessantemente, a única vogal tônica que apresenta grande dispersão é justamente a vogal que se encontra na posição mais isolada do gráfico. Não parece coincidência o fato de o /a/ ser produzido com uma constrição numa região do trato isolada em relação às de outras vogais tônicas - a região faríngea. Talvez seja exatamente a menor necessidade de distinguir essa vogal das demais que leve o falante a relaxar seus articuladores no momento de produzi-la.

Esse possível relaxamento dos articuladores acarretaria, segundo Stevens & House (1955), a produção de um *schwa*, o que pode ser comprovado pela trajetória das diferentes produções da vogal /a/, na qual se nota que [Ap] (a mais reduzida do conjunto das centrais) é bastante centralizada no gráfico, ocupando uma posição próxima da que seria ocupada pela vogal neutra. O mesmo fenômeno pode-se observar, em menor grau, para as anteriores, dentre as quais [Ep] ocupa a posição mais centralizada. Não se pode, no entanto, justificar essa teoria através dos valores das posteriores que, como se pode observar, não parecem mover-se para o centro do espaço vocálico: as vogais pós-tônicas, finais e não-finais ocupam praticamente o mesmo espaço ocupado pelo [u], com uma maior dispersão.

Diferentemente do esperado, observaram-se diferenças significativas entre as vogais [A] e [Ap], o que aponta para a necessidade de inserir as unidades formadas por Ap no corpus do sistema de síntese do LAFAPE.

Também diferentemente do esperado, percebeu-se uma grande e interessante influência da consoante precedente (CP) sobre as vogais pós-tônicas. Ela influencia todos os formantes das pós-tônicas enquanto, das tônicas, influencia apenas o F3.



Tabela 1. Resumo da Influência da Consoante Precedente

	F1	F2	F3	F4	DUR
Tônicas			X		X
Pós-tônicas	X	X	X	X	X

Parece, então, que uma grande coarticulação da vogal tônica com a CP é mais um fenômeno característico da redução vocálica.

É interessante, também, observar os valores de F<sup>2</sup> para F3 (tabela seguinte) que, além de ser o único formante da tônica afetado por CP, possui os maiores valores de F: 8,18 para as pós-tônicas e 5,17 para as tônicas, seguidos por 3,84 (F4 de uma pós-tônica).

Tabela 2. Influência da Consoante Precedente

Vogal	Variável dependente	Valor de F	Valor de Pr>F
tônica	F1	1.22	0.2169
pós-tônica	F1	2.3	0.0007
tônica	F2	1.1	0.3430
pós-tônica	F2	2.56	0.0001
tônica	F3	5.17	0.0001
pós-tônica	F3	8.18	0.0001
tônica	F4	1.08	0.3698
pós-tônica	F4	3.84	0.0001
tônica	DUR	4.8	0.0001
pós-tônica	DUR	3.56	0.0001

Os dados indicam, portanto, que há, sim, influência da CP sobre F3, independentemente da tonicidade da vogal e sobre todos os formantes da pós-tônicas.

A influência de CP sobre a pós-tônica traz um dado que coloca um impasse para a síntese concatenativa. A fim de não aumentar demasiadamente o corpus de unidades, optou-se por, mesmo tendo em mente que o ideal seria não realizar cortes no interior de segmentos reduzidos, segmentar as vogais pós-tônicas pois, caso contrário, haveria um aumento da ordem de vinte mil unidades no corpus.

Dessa forma, uma vogal precedida por qualquer consoante será concatenada à segunda parte extraída de uma vogal precedida por /p/ (consoante que menos afeta a vogal precedente).

A alteração provocada por CP na primeira parte da vogal pode provocar grandes discontinuidades na concatenação com a segunda parte.

Para remediar esse problema, optou-se por utilizar unidades demissilábicas para a concatenação. As demissílabas (Fujimura & Lovins, 1978) se diferenciam dos difones na medida em que, para a obtenção desses, se realiza o corte no centro de sua fase estacionária (região que geralmente coincide com o centro da vogal). Para as demissílabas, realiza-se o corte no final da transição, o que impede que duas fases

<sup>2</sup> F é o valor de um teste que verifica se as variâncias das classes definidas pelas variáveis são iguais ou diferentes, e a probabilidade desse teste incorrer em erro é dada por Pr>F.

estacionárias diferentes sejam concatenadas, pois a fase estacionária será somente a da segunda parte da vogal.

A decisão de usar demissílabas só poderá ser avaliada quando o sistema de síntese estiver em pleno funcionamento, embora os resultados de testes preliminares sejam animadores.

Chega-se, agora, a um maior conhecimento do sistema vocálico do PB. Os resultados obtidos estão de acordo com a visão pragueana, que afirmava que o arquifonema poderia ter uma realização fonética diferente daquelas dos fonemas neutralizados (Trubetzkoy, 1949). A neutralização estabelece um alvo novo para as vogais, em torno do qual elas se dispersam, caracteristicamente. O novo alvo é o resultado da neutralização, enquanto a grande dispersão em torno dele decorre da redução. Acredita-se, portanto, que se trata de processos distintos.

Não se pode afirmar que as pós-tônicas /e/ e /o/ se realizam como um [i] e [u]. Pode-se ver, através dos resultados estatísticos que o [E] se diferencia tanto de /e/, quanto de /i/. Somente o [O] apresenta uma tendência de aproximar-se mais de /u/ do que de /o/, diferenciando-se daquele, ainda assim, quanto a F3 e à duração.

Essa assimetria entre os pares anterior (i/E) e posterior (u/O), sugerindo que as pós-tônicas posteriores são mais redutíveis do que as anteriores, parece refletir uma tendência do PB já observada por Câmara Jr. (1971) ao propor que não havia a posterior na posição da pós-tônica da posição medial, isto é, a penúltima das proparoxítonas (ver figura I).

Cabe ressaltar que a diferença entre as vogais [u] e [O] só foi depreendida graças ao fato de terem sido medidos e analisados os valores dos 4 primeiros formantes das vogais, levando-se em consideração a afirmação de Fant (1960) de que são os três primeiros formantes os principais determinantes da qualidade fonética de uma vogal.

Câmara Jr. (1971:23) já demonstrava consciência, também, do fato de que /E/ e /O/ são vogais diferentes das demais quando afirma que elas se realizam como um débil /i/ e um débil /u/, respectivamente, mostrando, assim, a percepção de uma diferença entre esses dois pares. Talvez o que faltasse para que esse pioneiro fosse mais radical e desvinculasse tais pares era um aparelho que permitisse a análise acústica dessas vogais, aparelho de que hoje se dispõe e que passa a fornecer dados importantes para embasar hipóteses fonológicas.

Com base no estudo comparativo das vogais tônicas e pós-tônicas pôde-se concluir que nem todas as decisões tomadas para a elaboração do corpus do sistema de síntese foram acertadas.

Em primeiro lugar, o fato de haver uma forte interferência da consoante precedente sobre a vogal coloca, em princípio, em risco a decisão de não contemplar a transição entre elas nas unidades do corpus. Contemplar tais transições, no entanto, levaria a um aumento exponencial do número de unidades.

O uso das demissílabas como unidades pode, por outro lado, contornar o problema acima, mas somente a qualidade auditiva do sistema de síntese poderá confirmar essa hipótese. É interessante apostar na tentativa de contornar esse problema a fim de viabilizar um sistema que possa ser confeccionado com os recursos materiais e humanos do LAFAPE. Aproximadamente 700 horas de trabalho seriam necessárias para gravar,



segmentar e extrair as 20000 unidades estimadas caso não houvesse nenhum corte em segmentos reduzidos.

Um fator que aponta para a possibilidade de o problema ser contornado consiste na pequena influência que as consoantes precedentes têm sobre as vogais tônicas. Levando em conta que tanto as pré-tônicas quanto as tônicas serão concatenadas a partir de unidades extraídas de posição tônica, as pós-tônicas serão minoria. Assim, o problema poderá ser ocultado, ao menos em parte.

Caso as demissílabas não sejam suficientes para contornar o problema da influência da consoante precedente sobre a vogal, haverá a necessidade de um estudo específico sobre essa influência. Faz-se necessário ressaltar que tal influência só foi apontada por esse estudo, mas não foi (nem era esse o objeto do trabalho) estudada em profundidade. Para que isso seja possível, são necessárias várias repetições da mesma seqüência consoante-vogal, de forma a determinar quais são, exatamente, tais influências: quais as consoantes (ou grupos de consoantes) que exercem essas influências, etc. Resultados desse tipo podem indicar que algumas unidades longas são indispensáveis ao sistema, e outras, menos necessárias.

Quanto às vogais, tomou-se a decisão correta ao diferenciar as tônicas e pré-tônicas das pós-tônicas. Conforme já observado, as vogais pós-tônicas são completamente diferentes das tônicas, e utilizar a mesma unidade para concatenar uma vogal nas posições tônica e pós-tônica teria causado erros de acentuação sistemáticos.

Depois de todo o corpus do sintetizador gravado e digitalizado, além de o estudo aqui relatado ser realizado, foram feitos testes de concatenação. Esses testes vieram confirmar que, diferenciando-se vogais tônicas e pós-tônicas, a qualidade da fala sintetizada já é razoável, pois há 100% de acerto na colocação do acento lexical

Esses testes mostraram, também, que não havia, realmente, necessidade de diferenciar as tônicas das pré-tônicas no corpus do sintetizador, embora, algumas vezes, a concatenação de uma pré-tônica a partir de uma tônica, cause a impressão de uma fala hiperarticulada.

Atentando para essa impressão, mesmo em fala clara, e considerando o fato de que, quando se aumenta a taxa de elocução, vários segmentos tendem a ser mais reduzidos do que em fala clara e, percebendo que esse fenômeno é comumente observado em alguns monossílabos átonos e também nas pré-tônicas, tomou-se consciência da necessidade de um estudo sobre esse conjunto de vogais.

Sabe-se que não seria interessante gravar todas as pré-tônicas, da mesma forma como se fez com as pós-tônicas, pois, em fala clara, pausada, que é a mais aconselhável para um sintetizador, grande parte das pré-tônicas não se diferencia das tônicas: há, como afirmado acima, algumas pré-tônicas que se diferenciam e, por isso, percebe-se certa artificialidade na concatenação.

Pretende-se dar prosseguimento ao trabalho de aprimoramento do dicionário de unidades, analisando, agora, as pré-tônicas e os monossílabos átonos, e definindo qual seria o número mínimo de unidades contendo pré-tônicas a ser gravado; e, principalmente, determinando, no módulo do ORTOFON (Albano et al.), que vogais, além das pós-tônicas, sofrem reduções ou, ainda, processos de sandhi, para que possam ser concatenadas a partir das pós-tônicas e não a partir de tônicas.

De acordo com a safda atual do ORTOFON, a conjunção e, por exemplo, deve ser concatenada a partir do /e/ tônico, o que, em alguns ambientes prosódicos, deve, realmente, ocorrer, por exemplo, em casos como: "... e, conseqüentemente, ...". Essa mesma conjunção é, entretanto, em outros ambientes prosódicos, reduzida, por exemplo, em: "Traga uma caneta azul e um lápis preto".

Pretende-se, a partir de agora, fazer um levantamento de quais são os contextos que levarão as vogais que vêm sendo tratadas como plenas a serem tratadas como reduzidas.

Ao cabo da análise das pré-tônicas, haverá condições de fornecer um quadro, solidamente embasado, contendo dados sobre as vogais tônicas, pré-tônicas e pós-tônicas do português do Brasil, tal como manifestado na fala do informante analisado.

Retomando as conclusões acima, pode-se obter outra: fazer síntese de fala com a intenção de utilizar um número muito pequeno de unidades, desconsiderando informações lingüísticas, é perda de esforços. Por outro lado, tentar utilizar um grande número de unidades, também desconsiderando estudos lingüísticos, pode significar esforço infrutífero. Ou seja, é necessário que se considere a Lingüística como uma das ciências fundamentais para a elaboração de sistemas de síntese de fala.

Seria interessante, também, que os lingüistas passassem a lutar mais por ocupar o seu espaço na elaboração de qualquer sistema de síntese de fala, pois, além de serem os únicos cientistas capazes de fornecer um importante conjunto de informações necessárias à otimização desses sistemas, esses, como pôde ser visto, constituem ótimos testes das pesquisas básicas desenvolvidas em Lingüística.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- ALBANO, E. C.; Moreira, A. A.; Silva, A. H. P.; Aquino, P. A. e Kakinohana, R. 1995. "Um conversor fônico-métrico para síntese de fala em língua portuguesa". Trabalho apresentado no I CelSul, Florianópolis, novembro, a ser publicado como "Um conversor ortográfico-fônico e uma notação prosódica mínima para síntese de fala em língua portuguesa", in: Scarpa, E.M. (org) *Estudos de Prosódia no Brasil*, submetido à Editora da UNICAMP.
- BROWMAN, C. P. & Goldstein. 1986. "Towards an articulatory phonology" *Phonology Yearbook* (3), 219-252, Great Britain.
- CAMARA JR., J. M. 1970. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes
- \_\_\_\_\_. 1971. *Problemas de Lingüística Descritiva*. Petrópolis: Vozes
- FANT, G. 1960. *Acoustic Theory of Speech Production*. 'S-Gravenhage; Mouton.
- FUJIMURA, O. & Lovins, J. 1978. "Syllables as Concatenative Phonetic Elements", in A. Bell & J. Hooper (eds.) *Syllables and Segments*. Amsterdam: North Holland, 107-120.
- LINDBLOM, B. 1963. "Spectrographic Study of Vowel Reduction." *The Journal of the Acoustical Society of America*, 35, (1), 1773 -1781.
- STEVENS K. N. & House, A. S. 1955. "Development of a Quantitative Description of Vowel Articulation", *The Journal of the Acoustical Society of America*, 27, (3), 484-493.
- TRUBETZKOY, N. 1949. *Principes de Phonologie*, Paris, Klincksieck.