

## **“Syllable-timing in Brazilian Portuguese”: uma crítica a Roy Major**

Plínio Almeida BARBOSA  
(LAFAPE/IEL/UNICAMP)  
plinio@iel.unicamp.br

*ABSTRACT: This paper reintroduces the discussion about stress-timing in Brazilian Portuguese (BP). It begins by surveying some phonetic and phonological issues raised by the syllable- vs stress-timed dichotomy which culminated with the emergence of the p-center notion. Strict considerations of timing of V-V units and stress groups are taken into account to analyze the long term coupling of two basic oscillators (vowel and stress flow). This coupling allows a two-parameter characterization of language rhythms (coupling strength and speech rate) revealing that BP utterances present a high-degree of syllable-timing. A comparison with other languages, including European Portuguese, is also presented. The results analyzed indicate that Major’s arguments for considering Portuguese (sic) as stress-timing are misleading.*

*KEY WORDS: Rhythm Typology, Coupled Oscillators, Duration, Brazilian Portuguese*

*RESUMO: A tese de R. Major, segundo a qual haveria evidências para se considerar o português brasileiro (PB) como “stress-timing” ou tendendo para tal, é rediscutida. As questões fonético-fonológicas suscitadas pela dicotomia de línguas “stress-timed” e “syllable-timed” e o suposto isocronismo absoluto são apresentadas sob um prisma estritamente prosódico-temporal. Um modelo empregando dois osciladores acoplados (acentual e silábico) possibilita a caracterização biparamétrica (taxa de elocução e força de acoplamento) de um conjunto arbitrário de frases de uma língua e permite mostrar que, em PB, há alto grau de “syllable-timing”. À luz de uma análise fonética mais cuidadosa dos fatores ligados ao ritmo, mostra-se que os argumentos apresentados por Major para justificar “stress-timing” em PB são completamente equivocados.*

*PALAVRAS-CHAVE: Tipologia Rítmica, Osciladores Acoplados, Duração, Português Brasileiro*

### **0. INTRODUÇÃO**

Já no século 20, na década de 40, ao comparar o inglês falado com outras línguas (como o espanhol), Lloyd James (1940) e Pike (1945) afirmaram que o ritmo do inglês é caracterizado por uma sucessão de alternância entre sílabas acentuadas e não acentuadas em que as primeiras parecem ocorrer a intervalos regulares de tempo (doravante isocronismo acentual). Em contraste, o espanhol seria caracterizado por uma sucessão de sílabas igualmente espaçadas no tempo (doravante isocronismo silábico). Para marcar essa distinção aparente, Lloyd James propôs os termos militares respectivos de “ritmo de código Morse” e “ritmo de metralhadora” (no original, *Morse-code* e *machine-gun rhythms*). Assinalando a mesma distinção, Pike (1945:34-35) propõe os termos lingüisticamente mais apropriados de línguas

“stress-timed” e “syllable-timed”. Adotaremos aqui as conhecidas traduções de línguas de ritmo acentual e línguas de ritmo silábico, respectivamente. Mas manteremos os termos de “stress-timing” e “syllable-timing” para nos referir às organizações temporais subjacentes.

Ao propor os termos acima, Pike teve o cuidado de não afirmar que uma língua com uma tendência rítmica específica (seja ela para “syllable-timing” ou para “stress-timing”) não pudesse manifestar a outra. Ao contrário, cita o próprio inglês como apresentando características de syllable-timing em situações específicas, como alguns estilos de elocução e de trechos de fala cantada (Pike;1945:71). Essa possibilidade também é confirmada por observações mais recentes, a partir de excertos de discursos políticos. O testemunho da fala de Martin Luther King é muito esclarecedor<sup>1</sup> nesse sentido. Em uma gravação de um de seus discursos, disponível na URL <ftp://gahu.ucd.ie/pub/fred/sounds/mlk-ring.wav>, o trecho de 12,4 a 13,8 s é um exemplo de “syllable-timing” e o trecho de 15,7 a 21,2 s, exemplo de “stress-timing”.

Mas a precaução tomada por Pike não surte mais efeito a partir da década de 60, quando Abercrombie enuncia que “as far as is known, every language in the world is spoken with one kind of rhythm or with the other.” (1967:97). E, dando exemplos dos dois tipos de língua, ele cita o inglês, o russo e o árabe como de ritmo acentual e o francês, o telugu e o iorubá como de ritmo silábico. Mesmo que alguns pesquisadores como Kelly (1993) procurem isentar Abercrombie de toda culpa ao propor uma divisão mutuamente exclusiva para os ritmos das línguas do mundo, afirmando que, na verdade, ele tivesse em mente os aspectos fonológicos dos ritmos das línguas em questão, a afirmação do foneticista inglês deu lugar a uma busca, no domínio da produção, pelo isocronismo absoluto ou de *sílabas acentuadas* ou de *sílabas, tout court*. (O termo de *isossilabismo* é normalmente usado para o segundo caso, para marcar uma distinção em relação ao uso exclusivo do termo de isocronismo – “isochrony” - para referir apenas às línguas com sílabas acentuadas isócronas. Lehiste (1977), por exemplo, faz exatamente esse uso para o inglês. No presente artigo, usamos o termo “isocronismo” para nos referir indistintamente à recorrência subjetivamente regular de sílabas acentuadas – “stressed syllables”, no original - ou à recorrência de sílabas cujo grau de acentuação não é relevante.)

---

<sup>1</sup> Este interessante exemplo foi lembrado por Fred Cummins, na lista eletrônica SPROSIG (Special Interest Group in Speech Prosody, da ISCA, International Speech Communication Association).

Procuraremos demonstrar nesse trabalho a tese de que ainda é possível falar de isocronismo em produção, desde que se conceba um modelo de produção do ritmo de natureza hierárquica, com pelo menos dois níveis: o acentual e o silábico. Mostraremos também que, com o devido cuidado metodológico, é possível mensurar esse isocronismo de uma forma contínua mas lingüisticamente pertinente (pois a maneira como se dá a influência de um nível rítmico sobre o outro varia de língua para língua) que questiona seriamente abordagens estritamente qualitativas, como a de Dauer (1983). Não por serem desprovidas de sentido, mas por serem aproximações de uma realidade bem mais complexa. Através da metodologia apresentada aqui concluiremos que o PB é de tipo misto: silábico e acentual.

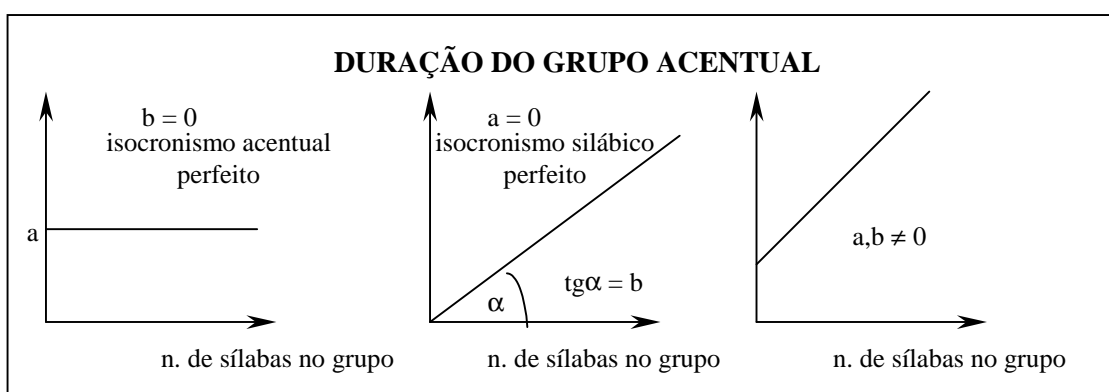
## 1. DA SENSACÃO DE ISOCRONISMO PARA A NOÇÃO DE *P-CENTER*

### 1.1. Da falha em se encontrar isocronismo absoluto no falante

Se tomado em sua versão mais radical, isocronismo em produção, o isocronismo acentual exclui *ipso facto* a possibilidade de isocronismo silábico. Para melhor entender esse fato, tomemos como exemplo um enunciado em uma língua hipotética com três acentos (“stresses”) seguidos delimitando, portanto, dois grupos acentuais com números de sílabas distintos. Digamos que o primeiro grupo tenha quatro sílabas e o segundo, oito. Se os dois grupos acentuais têm a mesma duração objetiva (isocronismo acentual), ou houve alongamento de uma ou mais sílabas do primeiro grupo para que a duração do mesmo fosse a mesma da do segundo ou encurtamento de uma ou mais sílabas do segundo grupo, para que sua duração se equiparasse à do primeiro. Fica claro que, se isso ocorre, as sílabas não têm mais a mesma duração e, portanto, não há isocronismo silábico. Se, inversamente, as sílabas dos dois grupos têm todas a mesma duração (isocronismo silábico), o isocronismo acentual é desfeito (porque o segundo grupo acentual de nosso exemplo teria o dobro da duração do primeiro). Essa lógica binária deve ter conduzido Abercrombie a formular da maneira citada acima, a célebre divisão entre os ritmos das línguas do mundo e, assim, suscitar a busca por isocronismo absoluto em produção.

O raciocínio acima pode ser colocado em forma gráfica (figura 1). Para tanto, é preciso traçar a relação existente entre as variáveis “número de sílabas no grupo acentual” e “duração do grupo acentual”. Pelo exposto acima para o isocronismo acentual, em que não há variação de duração entre grupos acentuais com diferente número de sílabas, a relação entre as variáveis é

correspondente ao primeiro padrão abaixo: uma reta paralela ao eixo das abscissas. Se, ao contrário, para o isocronismo silábico, a duração dos grupos é uma proporção direta do número de sílabas dos mesmos, a relação entre as variáveis corresponde ao segundo padrão da figura 1: uma reta passando pela origem. O terceiro padrão, que será explorado mais adiante, combina características dos dois primeiros (isocronismos acentual e silábico) e é a negação mesma de isocronismo absoluto, independente da unidade tomada (seja ela acento ou sílaba).



**Figura 1: Os Três Padrões de Variação da Duração dos Grupos Acentuais em Relação ao Número de Sílabas Contidas Neles.**

Se representarmos a variável *duração do grupo acentual* por  $I$  e a variável *número de sílabas no grupo acentual* por  $n$ , qualquer um dos padrões da figura 1 corresponde à equação da reta no plano bidimensional:

$$I = a + b.n \quad (1)$$

O parâmetro  $a$  representa o **ponto de interseção** da reta no eixo das ordenadas. O parâmetro  $b$ , a **inclinação da reta** (expressa pela tangente do ângulo entre a reta e eixo das abscissas). Para o isocronismo acentual,  $b = 0$  (e, pela fórmula acima,  $I = a$ , ou seja, duração do grupo acentual constante). Para o isocronismo silábico,  $a = 0$  (pela fórmula acima teríamos,  $I = b.n$ ). Para o terceiro padrão,  $a$  e  $b$  são não nulos ( $I = a + b.n$ , com  $a, b \neq 0$ ).

Os estudos em línguas tão diversas como inglês, sueco, alemão, espanhol e francês (Roach;1982, Fant & Kruckenberg;1989, Kohler;1986, Pointon;1980; Miller;1984; Wenk & Wioland;1982) mostraram de fato que, em todas elas, a relação entre duração do grupo

acentual e número de sílabas<sup>2</sup> é sempre a do terceiro padrão, com *a* e *b* diferentes de zero. **Não haveria assim nenhuma língua com isocronismos acentual ou silábico perfeitos em termos de produção de fala:** “as suggested by Mitchell (1969), there is no language which is totally syllable-timed or totally stress-timed - all languages display both sorts of timing; languages will, however, differ in which type of timing predominates.” (Roach;1982:78).

Para o sueco, por exemplo, Fant & Kruckenberg (1989:27) calculam uma regressão linear entre as variáveis já citadas, a partir de dados provenientes da leitura de enunciados de apenas um sujeito e encontram um coeficiente de correlação elevado ( $R = 92\%$ ), que é indicativo da forte linearidade da relação entre as duas variáveis, tal como apresentado na figura 1. O resultado para a equação da reta é:

$$I = 207 + 118.n \quad (\text{em milissegundos}) \quad (2)$$

Para os autores que trabalharam com línguas tidas como de ritmo silábico, o isocronismo de sílabas em termos de produção nunca foi encontrado (veja a forte crítica de Wenk & Wioland;1982 contra o isossilabismo no francês). Nessas línguas, a duração é o principal ou um dos parâmetros acústicos principais para assinalar o acento, desviando portanto a sílaba acentuada da suposta constância de duração em relação às outras sílabas.

A impossibilidade de encontrar medidas de duração objetivas que justificassem os padrões rítmicos propostos, levou à busca de alternativas no domínio da percepção (Lehiste;1977) e nos aspectos fonológicos das línguas (Bertinetto;1977, Dauer;1983).

## 1.2. Da possibilidade de isocronismo do ouvinte: aspectos fonológicos e perceptuais

A proposta de Dauer (1983) segue de certa forma a sugestão apresentada pela já citada resenha de Mitchell (1969) ao texto de Abercrombie, isto é, todas as línguas exibem características de “stress-timing” e “syllable-timing” variando apenas o grau em que uma das características predomina. Por serem características mutuamente exclusivas, como esclarecido acima, uma língua de alto grau de ritmo acentual terá automaticamente baixo grau de ritmo silábico.

---

<sup>2</sup> É claro que esse tipo de investigação pressupõe a universalidade fonética de termos como grupo acentual (e, portanto, de *acento*) e *sílaba*. A noção de grupo acentual foi criticada por Nooteboom (1991), que recebeu forte oposição de Bell-Berti (1991) e Fant, Kruckenberg & Nord (1991). Nooteboom (1991) foi um dos mais ferrenhos opositores à noção de isocronismo em produção.

Dauer propõe uma escala, caracterizada pela dimensão “stress-based” (que se poderia chamar “stress-timing”. Bertinetto;1989 sugere o termo “compressibilidade”), em relação à qual, “(...) a language is more or less stress-based, depending on how large a role stress plays in that language, especially with respect to the three areas described above.” (Dauer;1983:59). As três áreas descritas por Dauer são estrutura silábica (predominância na língua em questão de sílabas fechadas ou abertas), grau de extensão do fenômeno de redução vocálica (reduzindo freqüentemente ou não as vogais à vogal neutra, como em inglês, ou mantendo relativamente a qualidade vocálica, como em italiano e espanhol) e a maneira como o acento se realiza foneticamente e seu papel no sistema fonológico da língua em questão. Ela coloca o inglês no extremo “stress-based” da escala, por manifestar essas três características em alto grau. Fundamentando-se em Major (1981), o português do Brasil (doravante PB) é colocado logo antes do inglês, exibindo assim, para esses autores, elevado componente de “stress-timing”.

É importante salientar que a proposta de Dauer, contrariamente ao estudo fonético-instrumental de Major (1981), é fundamentada em aspectos fonológicos e, portanto, simbólicos. Qualquer uma das características acima pode ser contabilizada na forma de presença vs ausência de determinados processos ou representações fonológicas genéricas (como presença/ausência de sílabas fechadas ou abertas) ou específicas (como presença/ausência de sílabas CV, CVC, VC, V, etc). Um método para computar dessa forma o grau de “stress-timing” de uma língua é proposto pela própria Dauer (1987) e usado, por exemplo, para determinar a tipologia rítmica do búlgaro (Dimitrova;1998).

A partir dessas considerações, não é difícil perceber que a escala de Dauer dá conta de um aspecto matemático importante, o da **ordem**: uma sucessão de pontos em correspondência biunívoca com o conjunto dos números naturais, ordenamento este obtido por uma projeção de variáveis discretas sobre o eixo *grau de “stress-timing”*. Porém, sua escala não dá conta da noção de gradiência, que é um contínuo entre dois extremos, pois esse só poderia ser obtido pela projeção de variáveis igualmente contínuas, como aquelas exibidas por parâmetros fonéticos (sejam eles acústicos, auditivos ou articulatórios), sobre o mesmo eixo.

Além da procura de respostas no componente fonológico da língua para explicar a sensação de isocronismo, uma grande parte dos cientistas estudou a percepção do ritmo, para buscar nesse domínio razões para a tendência do ouvinte a perceber durações como sendo mais isócronas do que realmente eram, em termos de produção. Fraisse (1974), Lehiste (1977),

Allen (1972, 1975) e Huggins (1972) fazem parte desse grupo. Trabalhando ora com material acústico distinto da fala, ora com a fala de laboratório, fenômenos de ritmização subjetiva são assinalados por esses autores como resultado dos experimentos: os ouvintes subestimam a duração de intervalos longos e superestimam a duração dos curtos (Allen;1975). Estudos realizados com outras línguas (Benguerel & D’Arcy;1986 e Scott, Isard & Boisson-Bardies;1985) também relatam fortes tendências à ritmização subjetiva e, portanto, em seu conjunto, **a sensação de isocronismo em percepção presentes nessas línguas aponta para a consideração de características universais do componente rítmico, independentemente de qualquer caracterização do mesmo como acentual ou silábico.**

### 1.3. O ponto de ancoragem universal para percepção de isocronismo: o p-center

Essa tendência universal à equalização de seqüências objetivamente anisócronas levou à emergência do conceito de *perceptual-center* ou simplesmente *p-center*. O p-center seria a ponto no sinal acústico em que o ouvinte se ancoraria para perceber a seqüência sonora (de sinal de fala) como ocorrendo a intervalos regulares: “The P-center of a word corresponds to its psychological moment of occurrence.” (Morton, Marcus & Frankish;1976).

Para o criador do termo, Marcus (1976), como para autores na linha de Pompino-Marschall (1989, 1991) o ponto de ancoragem é exclusivamente psicoacústico. Os resultados de experimentos com fala sintética realizados por Pompino-Marschall (1989), por Scott (1993) e os realizados com fala natural, por Janker (1995), confirmam essa tese, ao obterem alto grau de confiabilidade na predição da localização do p-center, sem recurso a informação articulatória (como proposto por Tuller & Fowler;1980). A literatura em torno do *p-center* não foi explorada no Brasil (ver porém tentativas em Barbosa;1995a,b), mas ela é crucial para entender o fenômeno do isocronismo da fala em sua integralidade. O protocolo experimental de Pompino-Marschall (1989) é bastante adequado para esse fim.

No que diz respeito à produção, a célebre experiência de Allen (1972), em que o sujeito é convidado a pronunciar uma seqüência de sílabas em sincronismo com a sucessão de toques do seu dedo indicador sobre uma mesa, é substituída pela tentativa do sujeito em colocar em sincronismo a produção de uma seqüência de sílabas CVC com as batidas de um metrônomo calibrado para 60 batidas/minuto. Os resultados de Pompino-Marschall são reproduzidos na figura 2A e indicam que o sujeito procura coincidir o *onset* da vogal com as batidas do

metrônomo (o zero da ordenada). Na figura, a abscissa indica o primeiro membro do par de sílabas e a legenda, o segundo.

Quanto à percepção, um protocolo elaborado por Pompino-Marschall (1989) propõe a escuta (via fones de ouvido) de seqüências de cinco sílabas sintéticas (S) em alternância com tons puros (T) em que o sujeito é convidado a variar, através de um potenciômetro, a distância relativa entre a seqüência silábica e a de tons até que perceba a seqüência composta como isócrona. As durações da consoante e da vogal que formam a sílaba (CV ou VC) foram manipuladas sinteticamente em passos de 40 ms com valores aleatorizados entre duas sessões de escuta sucessivas. O ponto médio entre dois tons sucessivos, projetado sobre a seqüência silábica, é então tomado pelo experimentador como estimativa da localização do p-center. As figuras 2B a 2D abaixo mostram os resultados obtidos para dois tipos de estrutura silábica (CV: 2B e 2C vs VC: 2D) e para vogais e consoantes distintas (/ma/: 2B vs /fi/: 2C vs /am/: 2D). A ordenada indica a localização do p-center em milissegundos, a partir do *onset* da sílaba. A legenda com os símbolos geométricos indica os valores de duração das vogais sintéticas também em milissegundos. A linha inclinada das figuras 2B e 2C indica a posição do *onset* da vogal. Na figura 2D, essa linha coincide com o eixo das abscissas. Note a proximidade entre essas linhas e as diversas localizações do p-center.

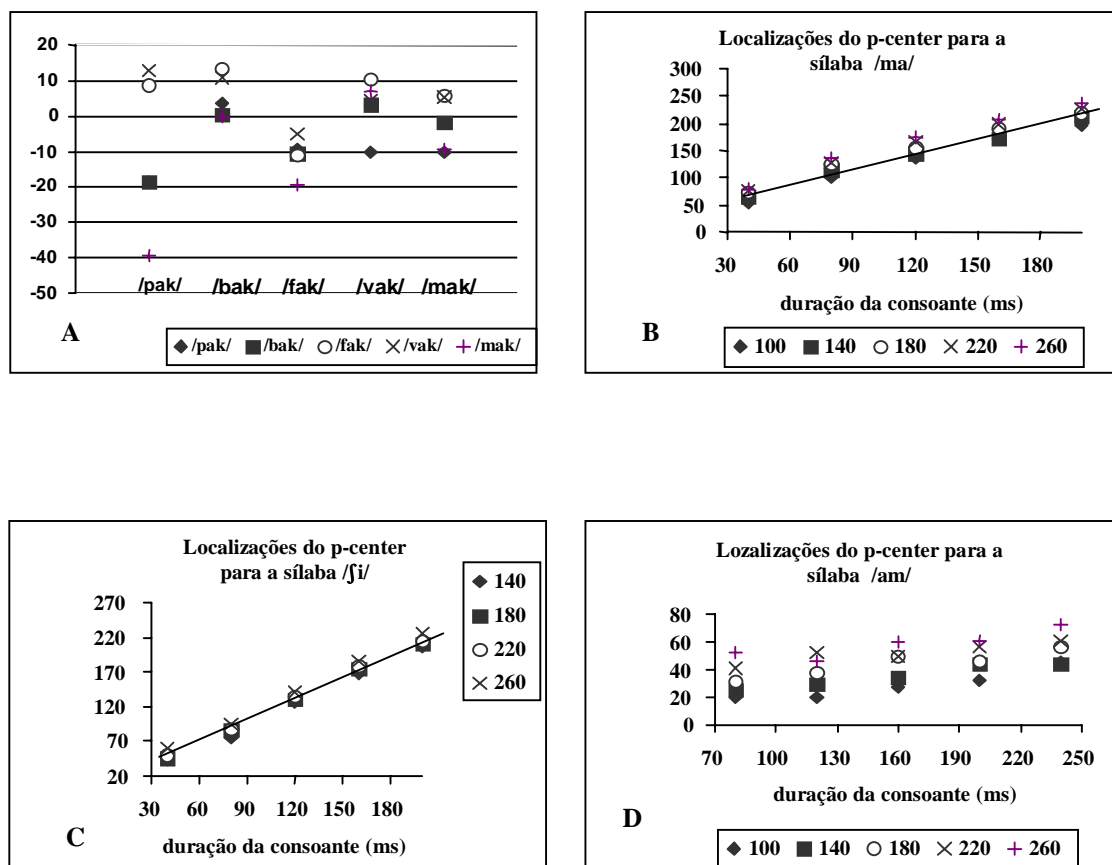
Além de confirmar a importância da transição CV para o processamento periférico auditivo (predominância de neurônios primários respondendo a eventos do tipo *onset*: Chistovich & Ogorodnikova;1982), o conceito de p-center constitui um argumento a favor dos modelos de produção de fala que consideram que essa mesma produção se dá pela superposição de dois mecanismos relativamente independentes: a produção contínua de vogais e a produção intermitente (via constrições sobrepostas ao movimento de portadora realizado pelas vogais) de consoantes (Fujimura;1995, Öhman;1966, Fowler;1983, Rhardisse & Abry;1995).

Em trabalho anterior (Barbosa & Bailly;1994), propusemos que o grupo entre dois *onsets* de vogal consecutivos (a que chamamos de grupo inter-p-center ou GIPC, pelas razões explicitadas acima) caracteriza melhor o ritmo do francês do que a sílaba.

Situação semelhante se dá em PB (Barbosa;1996). Nessa língua, mostramos que a evolução das durações dos GIPC sucessivos ao longo de frases isoladas indica pontos de culminância de crescendos duracionais que delimitam os grupos acentuais. Diferentemente do francês, os máximos de duração das sílabas indicam claramente todos os acentos lexicais dos



enunciados. As sílabas porém não revelam sistematicamente a estruturação prosódica nos mesmos enunciados.



**Figura 2: Localização do P-center em Função da Duração de Vogais e Consoantes de Sílabas Sintéticas.** (A) Posições do p-center para o experimento de produção relativas ao onset da vogal, para diversas sílabas CVC. (B a C) Posições do p-center, relativamente ao onset da sílaba, para o experimento de percepção e para diversas durações de consoante (abscissa) e de vogal, em milissegundos (legenda). Vide texto para explicação.

Vê-se assim que um conceito como o de p-center questiona a delimitação de grupos acentuais e, dessa forma, as análises fonético-acústicas sobre a organização temporal do PB (ou de qualquer outra língua) que não discutam a questão.

## 2. ORGANIZAÇÃO TEMPORAL EM PB

Dentro os trabalhos que resenharemos aqui destacamos aquele de Major (1981), conforme assinala o título de nosso artigo, pela influência que exerceu sobre a questão do ritmo do PB.

Essa influência não pode ser separada de dois fatores importantes da divulgação científica: o fato de ter sido escrito em inglês e para um periódico de renome.

## 2.1. O caráter limitado e parcial dos estudos de Roy Major

Através de medidas da duração de grupos acentuais, Major (1981) apresentou evidência em favor de ritmo acentual para o PB. Mas três questões relevantes, não tocadas pelo autor deveriam preceder suas análises: (a) como decidir qual palavra está acentuada? (b) a sílaba acentuada termina ou inicia o grupo? (c) quais as fronteiras do grupo acentual: sílabas ou vogais acentuadas? Essas perguntas serão exploradas quando da apresentação do *corpus* que analisaremos.

Após ter analisado grupos acentuais com dominância à esquerda (sem responder à questão b acima), as conclusões de Major, erroneamente generalizadas para o português (1981:350)<sup>3</sup>, apontam cinco razões para considerar o PB (correção nossa) como tendendo ao ritmo acentual: “(1) interstress durations are not directly proportional to the number of syllables; (2) many differences in interstress durations are not perceptible; (3) syllable duration is inversely proportional to the number of syllables in a word; (4) in casual speech unstressed syllables delete, which has the effect of equalizing the number of syllables in each stress group; e (5) shortening processes (of unstressed syllables), which reduce duration, have the effect of aiding stress-timing, i.e. ‘raising’, ‘monophthongization’, and ‘syllabicity shifts’”(1981:350).

Argumentamos abaixo que as quatro primeiras “razões” são, na verdade, características rítmicas universais (e portanto não justificam nenhum dos dois extremos da célebre dicotomia). A última delas provém de conhecimento parcial da fonética do PB. Verifiquemos uma a uma.

“Interstress durations are not directly proportional to the number of syllables”. Como vimos pela crítica que fizemos acima à busca do isocronismo absoluto em produção, as línguas exibem na verdade, o terceiro padrão da figura 1, ou seja, em **todas** elas (apontadas como de ritmo acentual ou silábico), não há proporcionalidade direta entre duração de grupo acentual e número de sílabas. **A afirmação é portanto uma característica universal do ritmo** e não uma defesa de um suposto ritmo acentual.

---

<sup>3</sup> Mesmo em PB diferentes tipos rítmicos são encontrados, dependendo do dialeto (Abaurre-Gnerre;1981, Cagliari & Abaurre;1986). Essa possibilidade não parece incomodar aqueles que, como Dauer (1983), consideram variedades distintas de línguas como o inglês, o espanhol, o grego e o italiano!

“Many differences in interstress durations are not perceptible”. **Essa conclusão também se refere a um fenômeno presente em qualquer língua.** Os autores que estudaram o isocronismo em percepção mostraram que duas durações que diferem por um valor abaixo de um limite conhecido como “just noticeable difference”, o chamado JND, não são percebidas como distintas. Em francês, por exemplo, classificada como “syllable-timed”, as sílabas acentuadas são muito variáveis e os sujeitos são menos sensíveis a modificações de duração ao final do grupo acentual que é o lugar da sílaba acentuada (Barbosa;1994).

“Syllable duration is inversely proportional to the number of syllables in a word”. **Uma outra característica universal:** quanto mais sílabas se acrescenta a uma palavra primitiva, mais curtas se tornam as sílabas que já estavam presentes nessa palavra. O exemplo é clássico em francês, em seqüências como *mal*, *malheur*, *malheureux*, *malheureusement* (O’Shaughnessy;1981)<sup>4</sup>.

“In casual speech unstressed syllables delete, which has the effect of equalizing the number of syllables in each stress group”. **Esse fato é um processo fonológico presente em qualquer língua do mundo** e é devido à hipoarticulação (Lindblom;1990).

“Shortening processes (of unstressed syllables), which reduce duration, have the effect of aiding stress-timing, i.e. ‘raising’, ‘monophthongization’, and ‘syllabicity shifts’”. Em PB, há processos fonológicos vivos na direção oposta, isto é, que **favorecem o** syllable-timing. Redução de sílabas acentuadas como monotongação de “ou” e “ei” (/ou/ → [o] e /ej/ → [e]) em palavras freqüentes como *louco* [ˈlo.ku], *pouco* [ˈpo.ku] e *madeira* [ma.ˈde.rɐ], *bobeira* [bo.ˈbe.rɐ]. Epêntese de [ɪ] entre seqüências de consoantes (que nunca ocorrem em português europeu) como em *psicologia* [pi.si.ko.lo.ˈʒi.ɐ], *admitir* [a.di.mi.ˈtʃix], *optar* [o.pi.ˈtax]. A vogal epentética [ɪ] pode ser encontrada mesmo em estilos de elocução coloquiais. Em algumas variedades do PB, uma regra de abaixamento é encontrada em posição pré-tônica (Abaurre-Gnerre;1981)<sup>5</sup>.

Essas últimas linhas, referentes ao conhecimento parcial da fonética e fonologia do PB se relacionam com outras falhas de Major, como a do artigo na revista *Language* (1985:263), mostrando que o PB possui apenas duas vogais nasalizadas em posição pós-tônica quando na

---

<sup>4</sup> O experimento foi feito com o francês canadense, mas é igualmente válido para o europeu.

<sup>5</sup> Alguns exemplos são a produção de [pɛˈrɔbɐ], peroba e de [pɛrɛˈrɛkɐ], perereca no dialeto do Espírito Santo (Abaurre-Gnerre;1981:27).

verdade temos cinco: [ĩ], como em *ínterim*; [ẽ], como em *hífen*; [ê], como em *ímã*; [õ], como em *fóton* e [ũ], como em *álbum*.

Os trabalhos de autores brasileiros foram muito mais cuidadosos. Jamais afirmaram categoricamente o predomínio de um tipo rítmico em detrimento de outro.

## 2.2. Estudos fonético-descritivos ou fonológicos

Os trabalhos de Moraes & Leite (1992), Cagliari & Abaurre (1986) e Abaurre-Gnerre (1981) avaliaram a tipologia rítmica do PB. Os dois primeiros autores analisaram a duração de grupos acentuais (ou pés, na nomenclatura fonológica deles) com dominância à esquerda e concluíram (ou se pode concluir de suas análises) que o PB é de tipo misto, isto é, apresenta características de ritmos acentual e silábico, embora essa composição sofra da variabilidade intra e intersujeito presente em seus dados.

Moraes & Leite (1992) mostram que a duração silábica média em grupos acentuais de quatro e oito sílabas é estatisticamente indistinta (o que é um argumento a favor do ritmo silábico), mas que em 90 % dos grupos acentuais com mais de sete sílabas (veremos mais adiante que o aumento do número de sílabas no grupo favorece realmente o “stress-timing”) há o fenômeno de elisão de sílabas. Apesar de ser um argumento fonológico, na linha de Dauer (1983), ao examinar atentamente seus dados (Moraes & Leite:70, 73-74), podemos extrair um argumento estritamente temporal (e portanto, fonético), qual seja, um efeito de encurtamento típico de um ritmo acentual. De fato, a duração média de três repetições de um grupo acentual de 8 sílabas é de 1,521 segundos e a de um grupo de 9 sílabas, de 1,325 segundos, sendo que a diferença é significativa com  $p < 0,01$ .

Cagliari & Abaurre (1986) usaram medidas de dispersão da duração de grupos acentuais para avaliar o grau de stress-timing em PB (menos dispersão é esperada em relação a línguas de ritmo acantual. Um procedimento semelhante é usado por Roach;1982). Para dez adultos nativos de regiões geográficas distintas, os autores evidenciam um *crescendo* (de sujeito para sujeito) de “stress-timing” a “syllable-timing”.

Fundamentando-se em evidência fonológica, Abaurre-Gnerre (1981) argumenta que o PB exhibe processos (como epêntese) de línguas de ritmo silábico e que, impressionisticamente falando, os dialetos da Bahia e do Rio Grande do Sul seriam particularmente de ritmo silábico (1981:39). Ela também sugere que as noções de ritmo silábico e acentual sejam relacionadas a

estilo de fala e taxa de elocução: estilos formais e lentos tenderiam ao ritmo silábico e estilos coloquiais e rápidos, ao ritmo acentual.

### 2.3. Estudos a partir de uma visão dinâmica do ritmo

Outros trabalhos destacaram a importância da análise quantitativa para uma melhor caracterização fonético-lingüística do ritmo do PB (Barbosa;1996, 1997, Barbosa & Madureira;1999). O principal foco desses trabalhos foi apontar o papel de unidades do tamanho da sílaba (como o GIPC, e a própria sílaba fonética) em revelar a estruturação rítmica do PB ao nível lexical (sílaba) e supralexical (GIPC) e construir modelos de geração automática da duração que consideram a natureza cognitiva do ritmo como provavelmente conexionalista (especialmente Barbosa;1997). Os padrões duracionais exibidos nesses artigos indicam claramente o aumento monotônico da duração do GIPC, perturbado por modificações locais advindas da implementação do acento lexical. Os GIPCs pós-tônicos são as unidades mais reduzidas o que é uma característica de língua de ritmo acentual, segundo Dauer (1983). Essa redução é adquirida ao mesmo tempo em que a criança vai aprendendo a coordenar gestos vocálicos e consonantais entre si (Gama-Rossi;1999).

Uma visão dinâmica do ritmo das línguas é claramente expressa por uma teoria de osciladores acoplados.

### 3. OSCILADORES ACOPLADOS COMO ELEMENTOS FUNDAMENTAIS DE UM MODELO DE PRODUÇÃO DO RITMO DA FALA

Quando percebemos o ritmo de uma língua, estaríamos prestando atenção tipicamente no movimento de sucessão vocálica, que é um movimento alternante, em termos de picos de intensidade sonora ou de abertura mandibular (abertura crescente, aproximadamente a partir do *onset* da vogal e decrescente do grau de abertura máximo, *grosso modo* no meio da vogal corrente, até o *onset* da vogal seguinte).

Uma outra alternância pode ser considerada, essa do próprio fluxo vocálico: a alternância acentual. Além da oscilação correlata da produção de vogais, um deslocamento mais pronunciado da mandíbula de tempos em tempos é correlato articulatorio do mecanismo de acentuação. Ter-se-ia portanto uma oscilação rápida, a das vogais, caracterizando a silabidade e uma outra, mais lenta, caracterizando a acentuação.

Esses dois movimentos ou oscilações podem ser estudados através de modelos de osciladores acoplados, em que a influência de uma oscilação sobre a outra é o aspecto mais relevante e mais conseqüente do modelamento.

Osciladores acoplados são muito comuns em sistemas naturais e podem ser encontrados no cérebro (Parthasarathy;1999) nas células do coração e do pâncreas e até mesmo entre organismos, como o ruído estridente e uníssono dos grilos ou a pulsação síncrona luminosa das lucíolas (Strogatz & Stewart;1994).

Os diversos modos de oscilação que esses osciladores exibem podem explicar o controle de atividades tão complexas como a respiração e a mastigação (ibidem). Recentemente, alguns autores começaram a explorar a noção para explicar os desempenhos dos indivíduos nas atividades de produção e percepção de fala (McAuley;1995, Port, Cummins & Gasser;1995, Cummins & Port;1998; O'Dell & Nieminen;1999). A proposta dos últimos autores tem implicações interessantes para a discussão sobre línguas de ritmo acentual e silábico.

Michael O'Dell e Tommi Nieminen (1999) utilizam uma técnica matemática, a Teoria da Diferença de Fase Média (APD, no original: "Averaged Phase Difference") para obter descrições qualitativas de longo termo de um modelo de osciladores acoplados composto de um oscilador silábico e um oscilador acentual. Um modelo semelhante, mas procurando descrever a influência momentânea de um oscilador sobre o outro foi proposto no mesmo forum em Barbosa e Madureira (1999). Os dois modelos são capazes de explicar os padrões duracionais da pesquisa sobre isocronismo na fala considerando apenas variáveis temporais e a noção de **força de acoplamento** entre os dois osciladores. Nesse artigo exploramos o modelo de O'Dell e Nieminen por permitir uma visão mais simplificada e a longo termo da influência recíproca dos dois níveis oscilatórios.

Duas simplificações são feitas pela técnica APD: o ciclo de cada oscilador (representado pela função matemática seno) é descrito apenas em termos de fase e se considera apenas o efeito médio, para cada ciclo, da diferença de fase ( $\Phi$ ) entre os dois osciladores. Se representamos a fase instantânea do oscilador silábico por  $\theta_s$  e aquela do oscilador acentual por  $\theta_F$  (F de *foot*, pé) e o efeito do oscilador silábico sobre o acentual por  $H(\Phi)$ , podemos escrever que as derivadas das fases instantâneas podem ser expressas como:

$$\theta_F^I = \omega_F + H(\Phi) \quad (3a)$$

$$\theta_s^I = \omega_s - r.H(\Phi) \quad (3b)$$

Nas fórmulas acima,  $\omega_F$  e  $\omega_s$  são as frequências respectivas dos osciladores acentual e silábico e a expressão  $-r.H(\Phi)$  é o efeito do oscilador acentual sobre o silábico. O parâmetro  $r$  é definido como a força de acoplamento (entre os dois osciladores).

A diferença de fase  $\Phi$  é uma função do número de ciclos  $n$  do oscilador silábico contidos no ciclo momentâneo do oscilador acentual. Cada ciclo representa uma sílaba (oscilador silábico) ou um acento (oscilador acentual):

$$\Phi(n) = \theta_s - n\theta_F \quad (4)$$

A diferença de fase é assim escrita para que as fases dos dois osciladores possam ser expressas nas mesmas grandezas (a fase do oscilador silábico varia  $n$  vezes mais rapidamente do que a fase do oscilador acentual).

Após algumas manipulações<sup>6</sup>, a duração do grupo acentual,  $I$ , (um ciclo do oscilador acentual), que é o inverso da frequência final do oscilador acentual, pode ser expressa como:

$$I = \frac{1}{\omega_F + H(\Phi(n))} = \frac{r}{r \cdot \omega_F + \omega_s} + \frac{1}{r \cdot \omega_F + \omega_s} \cdot n \quad (5)$$

Se compararmos essa fórmula com a fórmula 1, tem-se:  $a = \frac{r}{r \cdot \omega_F + \omega_s}$  e  $b = \frac{1}{r \cdot \omega_F + \omega_s}$

Conclui-se portanto que

$$r = a/b \quad (6)$$

Se a relação entre número de sílabas no grupo acentual e duração do mesmo podem ser expressas como uma reta (uma regressão linear pode avaliar o grau de adequação dessa asserção através do valor  $R$  do coeficiente de correlação. Quanto mais próximo de 1 ele for, melhor o grau de linearidade entre as duas variáveis), a razão entre os parâmetros  $a$  e  $b$  nos dá, portanto, a força de acoplamento. Se esse valor for unitário, a influência do oscilador silábico sobre o acentual é equivalente. Se um oscilador domina sobre o outro, tem-se então *uma tendência* para “syllable-timing” ( $r < 1$ ) ou para “stress-timing” ( $r > 1$ ).

O'Dell e Nieminen apresentam as análises que um outro autor, Eriksson (1991), fez com as médias de duração do grupo acentual por número de sílabas no grupo, apresentados por

Dauer (1983) e encontra os seguintes resultados para a regressão linear (R é o coeficiente de correlação).

**Tabela 1: Regressões Lineares para Cinco Línguas.** Equações de regressão linear e coeficientes de correlação (R) para cinco línguas, usando os dados de Dauer (1983), a partir da análise de Eriksson (1991), apresentada por O'Dell & Nieminen (1999:1075). As forças de acoplamento (r) foram acrescentadas por mim. Foram considerados apenas os grupos acentuais com 4 ou menos sílabas.

inglês (britânico e americano)	$I = 201 + 102n$ (r = 1,97)	R = 0,996
thai	$I = 220 + 97n$ (r = 2,27)	R = 0,973
espanhol (cubano e europeu)	$I = 76 + 119n$ (r = 0,64)	R = 0,997
grego	$I = 107 + 104n$ (r = 1,03)	R = 1,000
italiano	$I = 110 + 105n$ (r = 1,05)	R = 1,000

Pode-se verificar pela tabela acima que o espanhol apresenta forte tendência ao ritmo silábico, com  $r = 0,64$ , conforme apontado por Abercrombie (1967) e que o thai e o inglês, forte tendência ao ritmo acentual, respectivamente com  $r = 2,27$  e  $r = 1,97$ . O italiano e o grego, com forças de acoplamento próximas à unidade, teriam composições semelhantes de ritmos silábico e acentual. Duas críticas podem ser entretanto feitas ao procedimento de Eriksson.

A primeira diz respeito ao cálculo da regressão linear a partir de médias (de onde os elevados valores dos coeficientes de correlação). Para serem mais realistas, essas regressões deveriam ser calculadas com os valores individuais dos pares “duração de grupo acentual” e “número de sílabas no grupo”. A segunda crítica diz respeito ao equívoco (já cometido pela própria Dauer) de se considerar para cálculo mais de uma variedade da língua (como inglês britânico e americano, espanhol cubano e europeu, sem contar as diferentes variedades de italiano e grego), variedades essas que podem ter características rítmicas distintas, como se verá mais adiante.

Um outro cuidado que se deve ter é o de não caracterizar o ritmo de uma língua apenas pelo valor da força de acoplamento, já que a **taxa de elocução**<sup>7</sup> desempenha um papel crucial em

<sup>6</sup> No equilíbrio, a diferença de fase  $\Phi(n)$  se estabiliza, ou seja,  $\Phi'(n) = \theta'_s - n\theta'_f = (\omega_s - n \cdot \omega_f) - (r+n)$ .  $H(\Phi(n)) = 0$ .

Decorre então que  $H(\Phi(n)) = \frac{\omega_s - n \cdot \omega_f}{(r+n)}$ .

<sup>7</sup> É o termo que propomos, em 1995, para traduzir *speech rate* (Barbosa;1999). É termo mais apropriado que “velocidade de fala”, pois a grandeza física geralmente usada para indicá-la, “sílabas por unidade de tempo”, não expressa a velocidade real de deslocamento dos articuladores da fala (como a mandíbula), que é obtida pelo uso de técnicas modernas de registro do movimento de produção de fala (como eletromiografia, microfones de raios X, entre outras).



acelerar ou desacelerar o oscilador silábico e assim modificar a relação entre esse último e o oscilador acentual.

Esses cuidados serão tomados em nossa análise da duração do PB baseada no cálculo das forças de acoplamento e inclinação da reta de regressão linear (parâmetro *b*). Para tanto, um *corpus* gravado em três taxas de elocução é proposto.

### 3.1. *Corpus* e metodologia para o estudo do ritmo do PB sob a perspectiva dos osciladores acoplados

A leitura de 36 frases por um locutor masculino (de Pernambuco, 35 anos na época da gravação) em três taxas de elocução foi registrada em um gravador digital e amostrada a 16 kHz para análise no analisador de fala CSL, modelo 4300 B, da Kay Elemetrics<sup>8</sup>. Das três taxas, duas foram controladas com um metrônomo luminoso para reproduzir uma fala natural em taxa rápida e em taxa lenta. A terceira é a taxa de elocução normalmente usada pelo locutor, rotulada normal. O *corpus* foi originalmente montado para o estudo das características rítmica e entoacional do PB. Para tanto variaram-se a posição da palavra na sentença, a estrutura sintática e o número de sílabas por grupo acentual. Para medir a duração no CSL, foram introduzidas marcas para delimitar os segmentos fonéticos. As marcas indicam o início de um segmento acústico<sup>9</sup>. Para a delimitação dos grupos acentuais a partir da delimitação dos segmentos, é necessário responder às três perguntas colocadas acima.

A determinação das palavras efetivamente acentuadas depende, evidentemente, de testes de percepção. Roach (1982), Dauer (1983) e Moraes & Leite (1992) relatam as dificuldades intrínsecas de uma tal decisão, que precisa ser complementada por medidas objetivas como duração e intensidade (ver Massini;1991 e Barbosa;1996). Em nosso *corpus*, toda palavra lexical (bem como algumas palavras gramaticais) foi ouvida como acentuada exceto em casos de choque acentual (*stress clash*), tais como “está lasso” e “está limpa”. Nesses casos, a primeira palavra foi considerada desacentuada (como se fosse na verdade uma só palavra: “estalasso” – homônima a “estalaço”, de “estalo” - ou “estalimpa”). Esse alto grau de saliência perceptual parece ser a regra em PB, como evidenciado no estudo de Moraes (1986). Nesse trabalho, testes perceptuais sugerem que proparoxítonos são ouvidos como tendo maior proeminência que paroxítonos e esses últimos, que os oxítonos.

---

<sup>8</sup> No LAFAPE, Laboratório de Fonética e Psicolinguística, do Instituto de Estudos da Linguagem, UNICAMP.

<sup>9</sup> Ao final da sentença é necessário também marcar o *offset* do último segmento.

Quanto ao aspecto da dominância à esquerda ou à direita, seguindo a proposta da Fonologia Prosódica (Nespor & Vogel;1986 e Bisol;1996, para aplicação ao PB) pode-se adotar a solução de que apenas ao nível do pé (binário) o PB tem dominância à esquerda e que, nos níveis superiores (palavra fonológica, grupo clítico, sintagmas fonológico e entoacional e frase fonológica), tem dominância à direita. Consideraremos assim que seria mais natural supor que o grupo acentual em PB *termina* com GIPC acentuado, e não *começa* por um. A razão para isso é que o grupo acentual mínimo, foneticamente falando, tem o tamanho de um grupo clítico (embora, devido à noção de GIPC, não tenha fronteiras coincidentes, como se verá mais adiante). Na frase “Pedro estuda na escola de João” há quatro grupos clíticos: [Pedro]<sub>C</sub> [estuda]<sub>C</sub> [na escolinha]<sub>C</sub> [de João]<sub>C</sub>. Grupos assim determinam grupos acentuais quase-isomórficos em uma fala normal, embora em uma fala mais rápida possa haver um grupo acentual com maior número de sílabas, representado pela junção dos dois últimos grupos clíticos. Não há razão alguma para se contar, no entanto, os três pés (o primeiro degenerado) do grupo “na escolinha” como definindo dois ou três grupos acentuais (desde que o pé permitido seja apenas o binário, é claro).

Em relação à delimitação dos grupos acentuais, depois das fortes evidências apresentadas aqui a favor do p-center, não é preciso explicar porque autores como Dauer (1983), Fant & Kruckenberg (1989), O’Dell & Nieminen (1999), Barbosa & Bailly (1994), Barbosa (1996), entre outros, delimitam os grupos acentuais nas línguas que estudaram entre dois *onsets* de vogal acentuada consecutivos. Major (1981) e Roach (1982), além de outros autores, preferem o uso conservador do *onset* da sílaba, mas a decisão é equivocada, porque o acento incide sobre a rima e inclui eventualmente as consoantes heterossilábicas seguintes. Mostramos (Barbosa & Madureira;1999) que o segmento [ʃ] nas frases abaixo é mais alongado no primeiro enunciado (com o oxítono), quando recebe o grau de acentuação da sílaba acentuada que o precede imediatamente. Os segmentos [b] das palavras “gaba’ e “gabá” não têm diferença de duração significativa.

Zé diz se gabá | chapado e baixo.      Zé diz se gaba | chapado e baixo.

É portanto o GIPC [a#ʃ] que é o local da culminância acentual e não a sílaba fonológica (ver também Barbosa;1996 e Vaissière;1983). A regra só parece ser quebrada em fronteira forte, como a da sequência “Zé diz se gabá. Chapado também.” (em que “Chapado” é nome próprio), como mostrado por Barbosa & Madureira (1999).

Essas considerações permitem entender o tipo de delimitação para o grupo acentual que fizemos antes de proceder às análises da duração, como exemplificado abaixo.

A sel(a do caval)(o fic)(a numa prateleir)(a de uma antig)(a cel)a

O material excluído corresponde à anacrusis (“A sel”) e ao *offset* da frase (última vogal “a”). temos portanto cinco grupos acentuais com respectivamente 4, 2, 6, 5 e 2 sílabas (elisão do “a” de “uma”).

De posse do número de sílabas por grupo e da duração dos mesmos, é possível computar as retas de regressão linear para todo o *corpus*, nas três taxas de elocução.

#### 4. UM FORTE COMPONENTE DE RITMO SILÁBICO NOS DADOS DO PB: COMPARAÇÃO COM OUTRAS LÍNGUAS

A regressão linear também considerou grupos acentuais com a sílaba acentuada iniciando o grupo, bem como grupos acentuais com um número de uma a quatro sílabas e grupos de uma a sete sílabas. A finalidade disso é verificar as conseqüências do domínio à esquerda ou à direita sobre os cálculos bem como verificar o que ocorre quando o grupo acentual se torna mais longo (em relação aos analisados por Eriksson;1991). A tabela 2 apresenta o resultado das diversas regressões lineares. Para avaliar as conseqüências da adoção de dominância à esquerda ou à direita sobre as equações de regressão linear considerou-se apenas a taxa de elocução mais confortável para o locutor, a que foi chamada “normal”.

Da tabela 2, é possível verificar na segunda coluna (por médias) que os coeficientes de correlação são tão elevados quanto os obtidos por Eriksson (tabela 1) para o grego e italiano, com os dados de Dauer (1983). A inclinação *b*, diminui com a variação da taxa de elocução de lenta a rápida (186, 142 e 97 ms/sílaba), indicando que pode ser usado como medida quantitativa da mesma.

A força de acoplamento varia em função da taxa de elocução: menor que 1,00 (0,81) na taxa normal e maior que 1,00, nas taxas lenta e rápida (1,44 e 1,67, respectivamente). Pode-se assim concluir que a força de acoplamento não pode, sozinha, indicar o grau de “stress-timing” de um conjunto de enunciados: é preciso informar o valor numérico da taxa de elocução. Que a taxa de elocução influencia a percepção do tipo rítmico já foi assinalado por Abaurre-Gnerre (1981), embora suas sugestões (de correlação direta entre coloquialidade a taxas rápidas ao ritmo acentual e de formalidade e taxas lentas ao ritmo silábico) não coincidam com o observado aqui. Apesar de uma possível tendência à silabação, os

fenômenos de alongamento final em taxas de elocução mais lentas, podem ser intensificados pela presença de pausas silenciosas que indicam a desaceleração do oscilador silábico, assinalando uma maior influência do oscilador acentual sobre o silábico, favorecendo assim o “stress-timing” (a própria duração da sílaba acentuada, normalmente bem maior que as demais, é a negação mesma do ritmo silábico, tal como definido no início deste artigo). A mesma situação pode ocorrer em taxas de elocução mais rápidas, de forma ainda mais acentuada, pela diminuição de tendência à silabação. As taxas lentas podem exibir “syllable-timing” quando a redução da taxa de elocução se torna extrema (mas nesse caso pode haver a possibilidade de perda de informação prosódica – como no caso dos ditados – pela perda de coesão e marcação de grau de hierarquização entre os diversos grupos acentuais). Esses dados constituem um contra-argumento à proposta de Abaurre-Gnarre (1981) em associar taxas rápidas com ritmo acentual e taxas lentas com ritmo silábico.

**Tabela 2: Regressões Lineares a partir do Corpus do PB.** Equações de regressão linear (R é o coeficiente de correlação) e força de acoplamento (r) para três taxas de elocução segundo o tamanho do grupo acentual (tam. GA), o método para cálculo da regressão linear (média calculada a priori para cada tamanho de GA) e a dominância do grupo acentual (apenas na taxa de elocução normal). O número de grupos acentuais (N) usado em cada regressão linear é também apresentado. A regressão linear feita a partir das médias considera apenas GA com número de sílabas menor que 5 (na verdade 2 a 4, visto que não há grupo de uma sílaba em nosso *corpus*). Ver texto para explicação.

taxa de elocução	por médias tam. GA < 5 σ	tam. GA < 5 σ dominância à direita	tam. GA < 5 σ dominância à esquerda	tam. GA < 8 σ dominância à direita	tam. GA < 8 σ dominância à esquerda
<b>lenta</b>	I = 268 + 186n R = 1,00 <b>r = 1,44</b> N = 108 (3 av.)	I = 267 + 186n R = 0,73 <b>r = 1,44</b> N = 108	não foi feita	I = 270 + 185n R = 0,84 <b>r = 1,45</b> N = 138	não foi feita
<b>normal</b>	I = 115 + 142n R = 1,00 <b>r = 0,81</b> N = 126 (3 av.)	I = 114 + 143n R = 0,80 <b>r = 0,80</b> N = 126	I = 108 + 144n R = 0,79 <b>r = 0,75</b> N = 126	I = 169 + 122n R = 0,82 <b>r = 1,39</b> N = 143	I = 151 + 127n R = 0,83 <b>r = 1,19</b> N = 143
<b>rápida</b>	I = 162 + 97n R = 1,00 <b>r = 1,67</b> N = 112 (3 av.)	I = 152 + 99n R = 0,78 <b>r = 1,53</b> N = 112	não foi feita	I = 172 + 92n R = 0,81 <b>r = 1,86</b> N = 134	não foi feita

Na terceira coluna, percebe-se que, quando se faz a regressão linear a partir de todos os pontos de análise (e não com médias prévias que reduzem o número de pontos em nosso caso

para três: a duração média dos grupos acentuais de 2 sílabas, a dos de 3 sílabas e a dos de 4 sílabas), o valor do coeficiente de correlação diminui (0,73 a 0,80). Mas as forças de acoplamento são semelhantes. Esse segundo método é mais adequado, por ser mais realista em termos da confiabilidade da regressão linear entre os diversos enunciados (se for muito baixa, a hipótese de linearidade da relação entre duração do grupo e o número de sílabas deve ser rejeitada).

Para a taxa de elocução normal, também se percebe que não há muita diferença se o grupo acentual tem dominância à esquerda ou à direita (0,80/0,75, nas terceira e quarta colunas, e 1,39/1,19, nas quinta e sexta colunas). Pode-se assim inferir que os resultados que seriam obtidos por uma reavaliação (com grupos dominados à direita) das análises dos autores que consideraram o grupo acentual como sendo dominado à esquerda provavelmente conduziriam às mesmas conclusões. Não nos parece assim que essa seja uma questão central.

A outra observação diz respeito ao tamanho do grupo acentual. Se as colunas 3 e 5 são comparadas, percebe-se que a inclusão dos grupos acentuais de 5, 6 e 7 sílabas fez com que todas as taxas de elocução passem a manifestar “stress-timing”. Assim sendo, quanto maior o tamanho do grupo, maior é a tendência para os fenômenos ligados ao ritmo acentual, como se viu na explicação acima para as taxas de elocução lentas.

As observações feitas até aqui são uma indicação do cuidado que se deve ter quando se trabalha com o ritmo e a duração, em que diversas variáveis, como taxa de elocução, tamanho do grupo acentual, estilo de elocução, entre outras, precisam ser controladas apropriadamente.

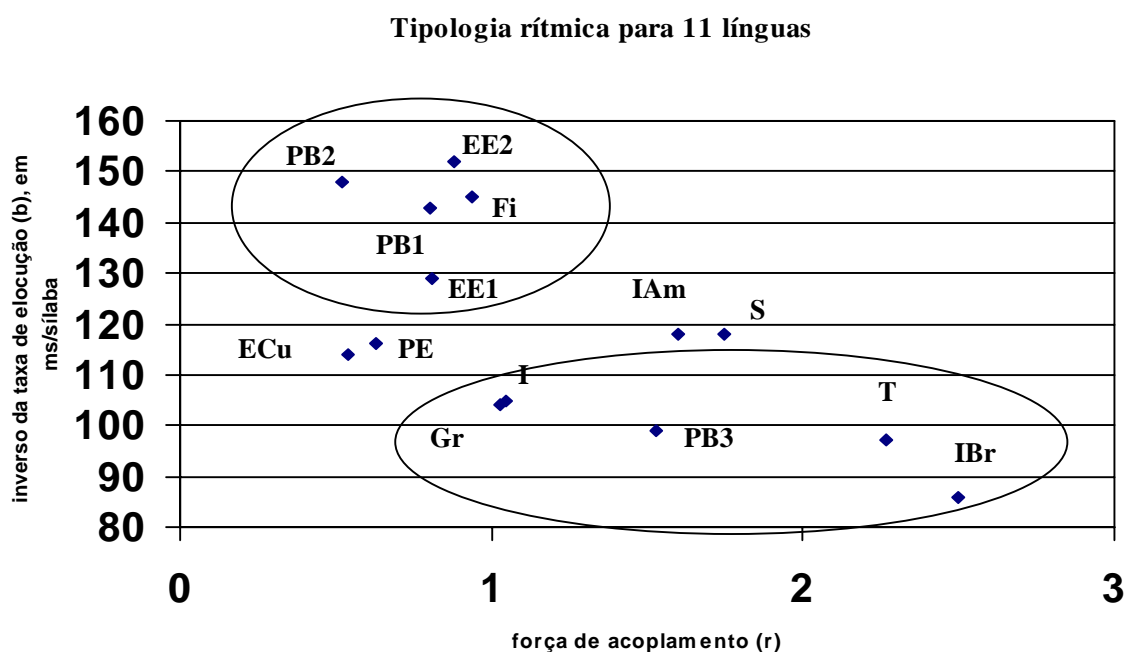
A partir dos dados de Dauer (1983), separando na medida do possível algumas variedades das línguas por ela apresentadas, de Fant & Kruckenberg (1989) e de O’Dell & Nieminen (1999), bem como de uma análise de dados do português europeu (PE)<sup>10</sup> e PB que apresentamos em outro lugar (Barbosa;em preparação), pode-se comparar as forças de acoplamento entre algumas línguas (figura 3), desde que se mantenha a comparação em torno do valor de uma taxa de elocução da língua que serve de base para a comparação (aqui, o PB). Referir-se à legenda da figura 3 para esclarecimento das abreviações seguintes.

Se tomarmos como referência as taxas de elocução em torno daquelas dos pontos da figura 3 correspondentes a PB1 e PB3 (obtidos da tabela 2, tomando-se os valores do par força de

---

<sup>10</sup> Agradeço especialmente a Charlotte Galves e Filomena Sandalo pela cessão do *corpus* de sinal de fala e medidas de duração no âmbito do projeto “Padrões rítmicos, fixação de parâmetros e mudança lingüística”. Frota e Vigário (1999) apresentaram as primeiras análises de duração desse *corpus*.

acoplamento,  $r$ , e inclinação da reta,  $b$ , para as taxas de elocução normal e rápida, respectivamente) com uma tolerância de 15 %, poderemos comparar esses enunciados do PB com o de outras línguas. Dessa comparação ressaí que há enunciados do PB tendendo mais para ritmo silábico (menor valor de força de acoplamento) para uma mesma faixa de variação da taxa de elocução (PB2 em relação a PB1). Essa diferença pode ser atribuída a uma variação de estilo de elocução (dado que locutor e conjunto de sentenças são distintos para os pontos PB1 e PB2). Em relação ao espanhol europeu e ao finlandês, o PB aparece então como semelhante em termos de tipologia rítmica (na fronteira entre “syllable-timing” e “stress-timing”, se usarmos a força de acoplamento  $r = 1,00$  como limiar) ou ligeiramente mais silábico.



**Figura 3: Caracterização Rítmica Biparamétrica de 11 Línguas.** Forças de acoplamento e inclinações da reta de regressão linear (uma medida quantitativa do inverso da taxa de elocução). As abreviaturas representam o inglês britânico (IBr), e o americano (IAm), o thai (T), o grego (Gr) e o italiano (I), espanhol europeu (EE1 e EE2) e cubano (ECu), a partir de regressões feitas ou refeitas por mim, com os dados de Dauer (1983). Os dados do sueco (S) provêm da fórmula de Fant & Kruckenberg (1989). Os dados do português europeu (PE) e brasileiro (PB2) são de um estudo de Barbosa (em preparação), a partir de um *corpus* de 20 frases apresentados em Frota & Vigário (1999). Os demais dados do PB (PB1 e PB3) são do estudo apresentado aqui, para grupos acentuais com menos de 5 sílabas. Os dados do finlandês (Fi) são de O’Dell & Nieminen (1999).

Para taxas de elocução mais rápidas (oval inferior), o PB é mais silábico do que o thai e o inglês britânico, porém mais acentual que o grego e o italiano. Para a faixa de variação de taxa de elocução restante, o PE aparece como intermediário entre o inglês americano ou sueco (mais acentuais que o PE) e o espanhol cubano (ligeiramente mais silábico que o conjunto de enunciados do PE aqui apresentado).

Como se percebe, com o devido cuidado metodológico, é possível extrair informações importantes quanto à tipologia rítmica de uma língua, no caso, o PB. Os resultados aqui apresentados variam em função da taxa de elocução (PB1 vs PB3, que são pontos obtidos da tabela 2) e do estilo de elocução (PB1 e PB2). Quanto à diferença entre variedades do português ou de outra língua, os dados aqui apresentados não permitem uma comparação direta, por terem sido pronunciados em taxas de elocução muito distintas. Os motivos já expostos nos permitem dizer com segurança que é temerário apresentar dados de tipologia rítmica de variedades distintas de uma mesma língua como um todo supostamente homogêneo.

## 5. À GUISA DE CONCLUSÃO

Há muito que a célebre dicotomia de línguas de ritmo acentual ou silábico<sup>11</sup> tem gerado discussões e publicações, muitas delas (Major;1981) apresentando análises parciais e conclusões precipitadas. Panoramas mais detalhados sobre o assunto podem ser lidos em Bertinetto (1989), Lehiste (1977) ou Barbosa (em preparação). Tendo sido abandonada como tal, a velha dicotomia sobrevive como rótulo ou como indicativo de tendências fonéticas ou fonológicas para as línguas estudadas.

A adoção de modelos de produção de ritmo de fala mais elaborados (como os osciladores acoplados apresentados aqui) permite uma observação mais acurada dos diversos padrões duracionais. O tempo é considerado nesse modelo como uma projeção de uma estrutura hierárquica que recoloca a questão das análises de caráter estritamente temporal em produção de fala. Eles também podem lançar luz sobre a relação entre os componentes rítmico e lexical à luz de modelos dinâmicos (Albano;no prelo, Barbosa;em preparação).

---

<sup>11</sup> Por simplicidade, mas sem nenhum prejuízo ao que foi apresentado aqui, excluímos as de ritmo moraico, cujo exemplo típico é o japonês.

Nosso estudo ressalta além disso o cuidado que se deve ter na consideração de uma medida quantitativa e no controle da taxa de elocução, variável que interfere diretamente sobre os padrões duracionais dos enunciados. Outras variáveis, como o estilo de elocução, ainda carecem de uma maneira precisa para serem quantificados.

O que é importante salientar é que os estudos de tipologia rítmica não devem ignorar aspectos metodológicos fundamentais como manifestos por variáveis fonéticas (taxa de elocução, tamanho do grupo acentual, limites do grupo: *p-center* ou sílaba), fonológicas (direção de dominância do grupo acentual) ou matemáticas (cálculo da regressão linear com todos os pontos, sem média *a priori*) tratadas aqui, sob o risco de, como Major, apresentar resultados altamente questionáveis que buscam reforçar idéias pré-concebidas baseadas em conhecimento parcial de nossa língua.

#### AGRADECIMENTOS

A Sandra Madureira e Eleonora Albano, pela leitura e sugestões em versões prévias deste artigo. Agradecemos também a Leda Bisol pelo entusiasmo e por ter possibilitado que o artigo constasse da seção Debate. Enfatizamos nossa gratidão à Fapesp, através do programa *Jovem Pesquisador em Centro Emergente* (nº 95/09708-6) e à Bolsa de Produtividade em Pesquisa do CNPq (nº 350382/98-0), vinculada ao projeto de nº 524110/96-4, que financiaram diretamente ou via equipamentos essas pesquisas. Não podemos deixar de agradecer ao nosso locutor, por sua voz e disponibilidade.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- ABAURRE-GNERRE, M. B. (1981) Processos fonológicos segmentais como índices de padrões prosódicos diversos nos estilos formal e casual do português do Brasil. *Caderno de Estudos Lingüísticos*, **2**: 23-34.
- ABERCROMBIE, D. (1967) *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- ALBANO, E. C. (no prelo) *O Gesto e suas Bordas: por uma fonologia acústico-articulatória*. Campinas: Mercado Aberto.
- ALLEN, G. D. (1972) The location of rhythmic stress beats in English I & II. *Language & Speech*, **15**: 72-100, 179-195.
- \_\_\_\_\_. (1975) Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics*, **3**: 75-86.
- BARBOSA, P.A. (1994) *Caractérisation et génération automatique de la structuration rythmique du français*. Thèse de 3<sup>e</sup> cycle inédita, INPG/ICP, Grenoble, França.



- \_\_\_\_ (1995a) Estrutura rítmica da frase revelada por aspectos de produção e percepção de fala. Manuscrito referente a comunicação oral apresentada no *XLIII Seminário do GEL-SP*, 25 a 25 de maio.
- \_\_\_\_ (1995b) O grupo inter-perceptual-center: uma nova unidade de programação rítmica. Manuscrito referente a comunicação oral apresentada no *II Congresso Brasileiro de Neuropsicologia*, maio.
- \_\_\_\_ (1996) At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration: emphasis on segmental duration generation. *Cadernos de Estudos Lingüísticos*, **31**: 33-53.
- \_\_\_\_ (1997) A model of segment (and pause) duration generation for Brazilian Portuguese text-to-speech synthesis. *Proceedings of the Fifth Conference on Speech Communication and Technology*, **2**: 2655-2658.
- \_\_\_\_ (1999) Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: *Estudos de Prosódia*. Scarpa, E. (org.). Campinas: Editora da Unicamp, 21-52.
- \_\_\_\_ (em preparação) Revisiting speech isochrony under a dynamical perspective: the case of (Brazilian) Portuguese.
- \_\_\_\_ & G. BAILLY (1994) Characterisation of rhythmic patterns for text-to-speech synthesis. *Speech Communication*, **15**.1-2: 127-137.
- \_\_\_\_ & S. MADUREIRA (1999) Toward a hierarchical model of rhythm production: evidence from phrase stress domains in Brazilian Portuguese. *Proceedings of the XIV<sup>th</sup> International Congress of Phonetic Sciences*, **1**: 297-300.
- BELL-BERTI, F. (1991) Comments on “some observations on the organisation and rhythm of speech”. *Proceedings of the XII<sup>th</sup> International Congress of Phonetic Sciences*, **1**: 238-242.
- BENGUEREL, A.-P. & J. D’ARCY (1986) Time-warping and the perception of rhythm in speech. *Journal of Phonetics*, **14**: 231-246.
- BERTINETTO, P. M. (1977) “Syllable-blood”, ovvero l’italiano come lingua al isocronismo sillabico. *Studi di Grammatica Italiana*, **6**: 69-96.
- \_\_\_\_ (1989) Reflections on the dichotomy “stress-” vs “syllable-timing”. *Revue de Phonétique Appliquée*, **91-92-93**: 99-130.
- BISOL, L. (1996) Constituintes Prosódicos. In: L. Bisol (Org.) *Introdução a Estudos de Fonologia do Português Brasileiro*. Porto Alegre: Editora da PUCRS, 247-261.
- CAGLIARI, L. C. & M. B. M. ABAURRE (1986) Elementos para uma investigação instrumental das relações entre padrões rítmicos e processos fonológicos no português brasileiro. *Caderno de Estudos Lingüísticos*, **10**: 39-57.
- CHISTOVICH, L. A. & E. A. OGORODNIKOVA (1982) Temporal processing of spectral data in vowel perception. *Speech Communication*, **1**: 45-54.
- CUMMINS, F. & R. PORT (1998) Rhythmic constraints on “stress-timing” in English. *Journal of Phonetics*, **26**: 145-171.
- DAUER, R. M. (1983) Stress-timing and syllable-timing re-analysed, *Journal of Phonetics*, **11**: 51-62.
- \_\_\_\_ (1987) Phonetic and Phonological components of language rhythm. *Proceedings of the XI<sup>th</sup> International Congress of Phonetic Sciences*, **5**: 447-450.
- DIMITROVA, S. (1998) Bulgarian speech rhythm: stress-timed or syllable-timed? *Journal of the International Phonetic Association*, **27**: 27-33.
- ERIKSSON, A. (1991) *Aspects of Swedish Speech Rhythm*. Göteborg: University of Göteborg.

- FANT, G. & A. KRUCKENBERG (1989) Preliminaries to the study of Swedish prose reading and reading style. *Quarterly Progress and Status Report, Speech Transmission Laboratory, KTH*, **2**: 1-83.
- \_\_\_\_\_, A. KRUCKENBERG & L. NORD (1991) Temporal organization and rhythm in Swedish. *Proceedings of the XII<sup>th</sup> International Congress of Phonetic Sciences*, **1**: 251-256.
- FOWLER, C. A. (1983) Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology*, **112**.3: 386-412.
- FRAISSE, P. (1974) *La Psychologie du Rythme*. Paris: Presses Universitaires de France.
- FROTA, S. & M. VIGÁRIO (1999) Aspectos de Prosódia Comparada: Ritmo e Entoação no PE e no PB. Manuscrito correspondente a trabalho apresentado no Congresso da Associação Portuguesa de Lingüística, em setembro.
- FUJIMURA, O. (1995) Prosodic organization of speech based on syllables: the C/D model. *Proceedings of the XIII<sup>th</sup> International Congress of Phonetic Sciences*, **3**: 10-17.
- GAMA-ROSSI, A. J. A. (1999) Relações entre desenvolvimento lingüístico e neuromotor: a aquisição da duração no português brasileiro. Tese de Doutorado inédita. Campinas: Universidade Estadual de Campinas.
- HUGGINS, A. W. F. (1972) On the Perception of Temporal Phenomena in Speech. *The Journal of the Acoustical Society of America*, **51**.4/2: 1279-1290.
- JANKER, P. (1995) On the influence of the internal structure of a syllable on the P-center perception. *Proceedings of the XIII<sup>th</sup> International Congress of Phonetic Sciences*, **2**: 510-513.
- KELLY, J. (1993) David Abercrombie (Obituary). *Phonetica*, **50**: 68-71.
- KOHLER, K. J. (1986) Invariance and variability in speech timing: from utterance to segment in German. In: J. Perkell & D. Klatt (Eds.) *Invariance and variability in speech processes*. Hillsdale, New Jersey: Lawrence Erlbaum, 268-289.
- LEHISTE, I. (1977) Isochrony reconsidered. *Journal of Phonetics*, **5**: 253-263.
- LINDBLOM, B. (1990) Explaining phonetic variation: a sketch of the H & H theory. In: H.J. Hardcastle & A. Marchal (Eds.) *Speech Production and Speech Modelling*. Dordrecht: Kluwer, 403-440.
- LLOYD-JAMES, A. (1940) *Speech signals in Telephony*. London.
- MAJOR, R. C. (1981) Stress-timing in Brazilian Portuguese. *Journal of Phonetics*, **9**: 343-351.
- \_\_\_\_\_. (1985) Stress and rhythm in Brazilian Portuguese. *Language*, **61**.2: 259-282.
- MARCUS, S. M. (1976) *Perceptual-centres*. Tese de Doutorado Inédita. Cambridge: Cambridge University.
- MASSINI, G. (1991) *A duração no estudo do acento e do ritmo em português*. Dissertação de Mestrado inédita. Campinas: Universidade de Campinas
- MCAULEY, J.D. (1995) *Perception of time as phase: toward an adaptive-oscillator model of rhythmic pattern processing*. Tese de Doutorado Inédita. Indiana, EUA: Indiana University.
- MILLER, M. (1984) On the perception of rhythm. *Journal of Phonetics*, **12**: 75-83.
- MITCHELL, T. F. (1969) Review of David Abercrombie (1967). *Journal of Linguistics*, **5**.1: 153-164.
- MORAES, J.A. (1986) Acentuação lexical e acentuação frasal em português. Um estudo acústico-perceptivo. Conferência apresentada no II Encontro Nacional de Fonética e Fonologia, Brasília.

- \_\_\_\_ & Y. F. LEITE (1992) Ritmo e velocidade da fala na estratégia do discurso: uma proposta de trabalho. In: R. Ilari (Org.) *Gramática do Português falado*, v. II: *níveis de análise lingüística*. Campinas: Editora da Unicamp.
- MORTON, J., S. MARCUS & C. FRANKISH (1976) Perceptual centers (p-centers). *Psychological revue*, **83.5**: 405-408.
- NESPOR, M. & I. VOGEL (1986) *Prosodic Phonology*. Dordrecht: Foris Publications.
- NOOTEBOOM, S. G. (1991) Some observations on the temporal organisation and rhythm of speech. *Proceedings of the XII<sup>th</sup> International Congress of Phonetic Sciences*, **1**: 228-237.
- O'DELL, M. & T. NIEMINEN (1999) Coupled oscillator model of speech rhythm. *Proceedings of the XIV<sup>th</sup> International Congress of Phonetic Sciences*, **2**: 1075-1078.
- ÖHMAN, S. (1966) Coarticulation in VCV utterances: spectrographic measurements. *J. Acoustic. Soc. Am.*, **39**: 151-168.
- O'SHAUGHNESSY, D. (1981) A study of French vowel and consonant durations. *Journal of Phonetics*, **9**: 385-406.
- PARTHASARATHY, H. (1999) Mind rhythms. *New Scientist*, 30 October, 28-31.
- PIKE, K. (1945) *The Intonation of American English*. Ann Arbor: University of Michigan Press.
- POINTON, G. E. (1980) Is Spanish really syllable-timing? *Journal of Phonetics*, **8**: 293-304.
- POMPINO-MARSCHALL, B. (1989) On the psychoacoustic nature of the p-center phenomenon. *Journal of Phonetics*, **17**: 175-192.
- \_\_\_\_ (1991) The syllable as a prosodic unit and the so-called p-centre effect. *Forschungsberichte der Institut für Phonetik und Sprachliche Kommunikation der Universität München*, **29**: 66-124.
- PORT, R., F. CUMMINS & M. GASSER (1995) A Dynamic Approach to Rhythm in Language: Toward a Temporal Phonology. *Proceedings of the Chicago Linguistics Society*, Luka, B. & Needs., B. (Eds.), 375-397.
- RHARDISSE, N. & C. ABRY (1995) Mandible as syllable organizer. *Proceedings of the XIII<sup>th</sup> International Congress of Phonetic Sciences*, **3**: 556-559.
- ROACH, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. In: D. Crystal (Ed.) *Linguistic controversies. Essays in linguistic theory and practice in honour of F. R. Palmer*. London: Edward Arnold, 73-79.
- SCOTT, D. R., S. D. ISARD & B. BOYSSON-BARDIES (1985) Perceptual isochrony in English and in French. *Journal of Phonetics*, **13**: 155-162.
- SCOTT, S. K. (1993) *Perceptual centres in speech: an acoustic analysis*. Tese de Doutorado Inédita. London: University College London.
- STROGATZ, S. & I. STEWART (1994) Oscillateurs couplés et synchronisation biologique. *Pour la Science*, **196**: 40-46.
- TULLER, B. & C. A. FOWLER (1980) Some articulatory correlates of perceptual isochrony. *Perception & Psychophysics*, **27** .4: 277-283.
- VAISSIÈRE, J. (1983) Language-independent prosodic features. In: Cutler, A. & Ladd, D.R. (Eds.) *Prosody: models and measurements*. Berlin: Springer-Verlag, 53-66.
- WENK, B. J. & F. WIOLAND (1982) Is French really syllable-timed? *Journal of Phonetics*, **10**: 193-216.