

ESTRATÉGIAS PROSÓDICAS DA LEITURA EM VOZ ALTA E DA FALA ESPONTÂNEA

Waldemar FERREIRA NETTO¹

Fernanda CONSONI²

- RESUMO: O presente trabalho tem como objeto a análise prosódica comparada entre leitura em voz alta e fala espontânea, tendo em vista a hipótese de que se os aspectos prosódicos característicos da fala não estão presentes em grande parte das manifestações materiais da linguagem, na leitura de textos em voz alta, o leitor tem de, necessariamente, criar a prosódia para o texto a partir de suas hipóteses, seguindo princípios específicos, especialmente no que diz respeito à variação de frequência. A partir de um procedimento estabelecido de análise automáticas, as análises realizadas apontaram que a variação da frequência nos textos lidos é mais aguda do que a da fala espontânea e que a relação entre a frequência média e a frequência final na leitura aproxima-se mais da relação entre tom dominante e tom fundamental da escala diatônica ocidental, estabelecendo uma cadência harmônica perfeita para a finalização das frases. Esses resultados apontam para a associação das intuições do leitor às suas formas prosódicas prototípicas definidas no período de aquisição da linguagem.
- PALAVRAS-CHAVE: Fonologia. Prosódia. Fala espontânea. Leitura.

Introdução

O presente trabalho tem como objeto a análise prosódica comparada entre leitura em voz alta e fala espontânea tendo em vista a hipótese de que a implementação da escrita alfabética e a expansão de seu uso pela imprensa acarretaram uma mudança na forma material da linguagem que deve ser considerada com muita atenção. Trata-se particularmente do fato de que a sonoridade da linguagem terminou por enfatizar a cadeia segmental, na medida em que os desenhos das letras são representações das variações espectrais da fala e não das variações da frequência ou de intensidade. Dessa maneira, os aspectos prosódicos característicos da fala não estão presentes em grande parte das manifestações materiais da linguagem hodierna. (FERREIRA NETTO, 2006) Dessa hipótese decorre que, na leitura de textos em voz alta, o leitor tem de, necessariamente, criar a prosódia para o texto, a partir de suas hipóteses. Assim, tendo em vista que essa criação deve seguir princípios prosódicos específicos, especialmente no que diz respeito à variação de frequência, é nosso propósito

¹ USP - Universidade de São Paulo. Faculdade de Filosofia Letras e Ciências Humanas - Departamento de Letras Clássicas e Vernáculas. São Paulo - SP - Brasil. 05508-900 - waldemar.ferreiranetto@gmail.com

² USP - Universidade de São Paulo. Faculdade de Filosofia Letras e Ciências Humanas - Departamento de Letras Clássicas e Vernáculas. São Paulo - SP - Brasil. 05508-900 - fernanda.consoni@gmail.com

estabelecer o cotejo entre a prosódia da fala espontânea e a prosódia da leitura em voz alta, entendendo que a prosódia da leitura em voz alta tem um caráter de criação *ad hoc*, que segue as intuições própria do falante sobre sua própria fala, enquanto que a prosódia da fala espontânea flui automaticamente, sem a interferência consciente do falante.

A fala espontânea envolve variações que provocam mudanças significativas na frequência fundamental média dos núcleos silábicos. Uma das possibilidades de interpretação desse fato é a definição de uma coleção de tons cuja harmonia se estabelece a partir de um tom fundamental (SHOENBERG, 2001). A seqüência melódica tem diferentes trajetórias que envolvem afastamentos do tom fundamental, mas tem, também, aproximações que criam sempre a expectativa de que a finalização dessa trajetória seja esse mesmo tom fundamental. A coleção de sons harmônicos estabelecida a partir do fundamental não é entretanto linear, isto é, não se estabelece pela proximidade de frequências; ao contrário disso, na medida em que frequências próximas criam sensações auditivas desagradáveis, a proximidade harmônica dá-se pela adaptação mútua dos movimentos moleculares que formam as ondas sonoras. Assim, os movimentos harmônicos cuja periodicidade estabeleça ciclos em que a superposição adeque-se aos movimentos moleculares do fundamental são mais harmônicos do que aqueles em que essa superposição seja menos adequada. Dessa maneira, entende-se que frequências dobradas, cuja superposição é ótima, são as que têm maior harmonia com o fundamental. Frequências três, quatro, cinco, seis, e assim sucessivamente, vezes a do fundamental estabelecem os diferentes graus de semelhança tonal. Assim, se um fundamental de 100 Hz ocorre, sons de 200 Hz, 300 Hz, 400 Hz serão mais harmônicos com eles do que sons de 101 Hz, 102 Hz ou 103 Hz. Neste caso, a superposição dos movimentos moleculares será bastante difícil, enquanto no outro caso parecerá quase espontânea.

Caso não haja a presença simultânea (harmonia) de um tom grave fundamental, a formação da melodia (harmonia + modulação) exige que a seqüência das notas (modulação) ocorra em tempo suficiente para que as ondas das notas seqüenciadas possam superpor-se (harmonia) e gerar, se for o caso, um fundamental, pelo processo do *missing fundamental* (MF). Caso não haja esse tom grave fundamental nem rapidez suficiente na seqüência das notas e/ou as notas não gerarem um fundamental por MF, a coesão se perde e não é possível ao ouvinte formar uma *gestalt* melódica que possa desencadear uma emoção qualquer. A presença do tom grave fundamental estabelece uma seqüência de sons semelhantes (8ª, 5ª e 3ª, ou suas oitavas) na medida em que a superposição desses tons redunde na recriação do próprio fundamental gerador, por MF.

O processo de recriação do fundamental por sons inteiros múltiplos do fundamental inexistente decorre do princípio básico de que entre todas as formas

imaginárias de ondas estacionárias, só são possíveis aquelas cujos nodos estejam exatamente nos mesmos pontos P e Q. Em outras palavras, são permitidas apenas aquelas ondas estacionárias senoidais que ‘caibam’ um número de inteiro de vezes entre P e Q (ROEDERER, 2002). Essa restrição permite o raciocínio inverso, na medida em que ondas cujos nodos estejam exatamente nos mesmos pontos permitem a dedução de um fundamental específico, e de outros sons mais graves, cujos nodos estejam também nesses mesmos pontos.

Dessa maneira, a superposição ou a seqüenciação de sons semelhantes é um processo harmônico semelhante ao da cobra que morde o próprio rabo: sons agudos com nodos em posições idênticas (consonantes) geram sons graves com nodos nessas posições e vice-versa; sons graves geram sons agudos com nodos nas mesmas posições do grave que foi seu gerador. A modulação, ou seja, o acréscimo de sons a um conjunto coeso pelo princípio dos nodos em posições idênticas, portanto, permite a fuga desse círculo vicioso, iludindo o ouvido, que, pela *gestalt* instaurada, suporia um fundamental, mas o que se lhe apresenta é diferente.

A idéia de um tom fundamental que garante a inércia esperada de um sistema, sem que haja variações significativas, vem acompanhada da idéia de que as variações significativas decorrem da quebra dessa inércia, gerando a necessidade de retorno ao tom com que se está familiarizado. A percepção do conjunto, quer o que promova o retorno, quer o que promova a mudança, dá-se tão somente por um processo previamente estabelecido de que a seqüência dos fatos tem de ser tomada como um fenômeno naturalmente encadeado e pronto a formar constantemente um todo coerente e coeso, que sirva como referência para as novas variações aleatórias, ou para as variações de suspense com retorno “triumfal” ao mesmo ponto, tal como propôs Shoenberg (2001)³.

Tendo em vista essa característica própria dos sons tomados para a elaboração de melodia, entendemos que o mesmo princípio possa aplicar-se à fala quanto às suas variações de frequência que compõem a prosódia. A idéia básica de uma coleção de sons harmônicos entre si, estabelecidos a partir de um tom fundamental, entretanto, não se aplicar à fala da mesma maneira como ocorre com a melodia, em que pese tratar-se do mesmo fenômeno físico. A elaboração e a percepção de frequências na fala sujeita-se a padrões diferentes, para os mais diferentes propósitos (FÓNAGY, 2003; VAISSIÈRE, [2002?]). t’Hart (1981) verificou que a percepção das variações de frequências percebidas na fala mantinha-se em torno de 1,5 tom ou 3 st (semitons). Trata-se de um padrão bastante diferente daquele das interpretações melódicas, pois neste padrão, podem-se perceber

³ Cf. CHENOWETH, 1986; LONGACRE; CHENOWETH, 1986.

variações mínimas de até 0,1% do som que se produz (ROEDERER, 2002; MENEZES, 2003).

A mudança no padrão perceptual envolve uma redefinição do objeto em questão. A coleção de sons estabelecidos pelo fundamental estará sujeita a uma variação grande demais para ser definida por um tom fixo. Dessa maneira, ter-se-á que se pensar em bandas de frequências estabelecidas a partir de uma tendência central. A própria variação de 3 st pode ser usada como critério para o estabelecimento dessa coleção.

Esse tipo de procedimento já fora objeto de discussão relativamente às canções infantis da Vestfália e à música javanesa (WEBER, 1995). Segundo Weber (1995, p.70)

[...] o mais provável é a interpretação da escala pentatônica como uma combinação de duas quartas diazêuticas, na qual originariamente as duas quartas só eram divididas por um intervalo que, de acordo com o movimento melódico (especialmente se para cima ou para baixo), era *móvel* e, eventualmente, irracional.

Weber (1995) ressaltou que a melodia, em grande parte, das músicas cerimoniais tem pouca variação de tons e maior estereotipagem em relação a momentos anteriores de sua história. É possível, também, retomar as experiências de Densmore (1909) e de Roberts (1932), que, analisando músicas não-ocidentais, encontraram outras possibilidades de formação de escala. Tendo em vista tratar o projeto de uma análise da prosódia segundo os princípios básicos da análise musical, tem-se a hipótese de que as idiosincrasias culturais manifestas na tradição musical de um povo refletem-se na prosódia (GLASER, 2000; SCHELLENBERG; TREHUB, 1999).

A se tomar a escala musical ocidental, teremos uma relação precisa de tônica e dominante (T e D ou I e V), isto é, um intervalo de quinta. Frases harmônicas musicais são encadeamentos de acordes, definidos por seus fundamentais graves. Tradicionalmente, chama-se encadeamento perfeito ou cadência perfeita a que produz conclusão satisfatória e é formada pelo acorde da tônica no tempo forte, precedido pelo acorde da dominante (COMPÔTE, 1977). No caso da fala, o tom médio, em geral com maior ocorrência, tem o valor precípua de um acorde dominante que, precedendo um tom com frequência de 1,41 vezes mais grave – que é o mesmo que um fundamental, quando se o considera de um ponto de vista musical – estabeleceria uma cadência perfeita e, portanto, seria percebido como uma conclusão satisfatória, se o fundamental ocorre em tempo forte.

Tendo em vista que na fala espontânea as variações excedem facilmente os limites das variações aceitas para as melodias, a definição do fundamental

que contém os dados das pastas de trabalho já abertas anteriormente e apaga todas elas em seguida à cópia feita. A sincronização da transcrição é feita automaticamente. A macro solicita ao usuário uma definição de padrão mínimo de intensidade, entre 500 e 2000, sendo que este último é o padrão no caso de nenhuma resposta, ou nos casos de resposta inválida. A unidade é definida pelo SFS.

Em seguida, a macro inicia suas operações, estabelecendo, primeiro, a média geral das frequências para os valores válidos. Valores válidos são definidos aprioristicamente, como:

Limiar inferior de frequência: 50hz

Limiar superior de frequência: 350Hz.

Limiar inferior de duração: 4 frames ou 20ms (1 frame = 2 ms)

Limiar inferior de duração: 10 frames ou 50 ms (1 frame = 2 ms)

Limiar de intensidade: entre 500 e 2000; o padrão é 2000. Opção do usuário.

Valores utilizados para a elaboração da escala de cinco tons:

Limite superior das frequências médias => valor médio * 1,09)

Limite inferior das frequências médias => valor médio / 1,09)

Distância entre cada média (3st = $3 * 1,0595 = 1,19$)

Valores utilizados para a elaboração da escala de intensidade:

Limite superior do valor médio de intensidade (1,50)

Limite inferior do valor médio de intensidade (0,5)

Limite entre cada valor de intensidade= 1,25 sup e 0,5 inf

Categorização da intensidade na escala (= 1 ou 3 ou 5)

A partir dessa frequência média, as frequências são categorizadas segundo uma escala padronizada, em que se definem 5 bandas – 1,19 vezes acima e abaixo do valor médio das frequências. As bandas de frequência, por sua vez, têm seu limite superior e inferior definidos por serem 1,09. Como se vê no Gráfico 2, as categorias definidas são bandas que permitem alguma variação entre seus limites, o que não permite uma avaliação precisa da tonalidade, em termos musicais, ou da frequência específica, mensurável em Hz, por exemplo.

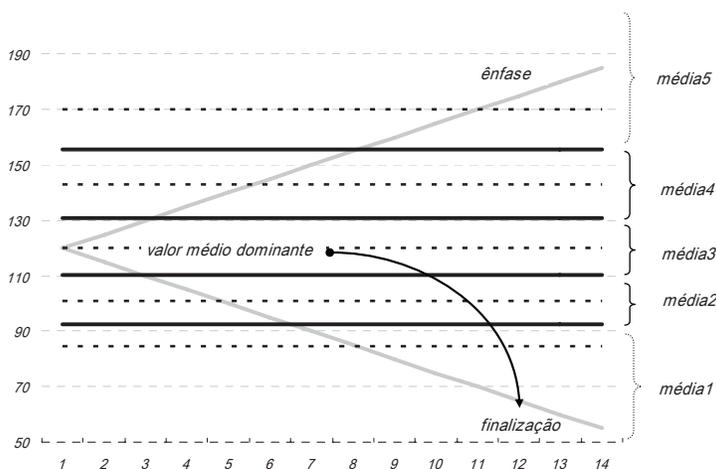


Gráfico 2 – Demonstração de bandas estabelecidas pelas médias de frequência, a partir de um valor médio dominante de 120Hz. As linhas pretas pontilhadas mostram os valores médios de cada banda. As linhas pretas contínuas mostram os limites de cada banda, com exceção das referentes às médias mais extremas, 1 e 5, que não definem limites inferior e superior, respectivamente. As linhas diagonais cinzas contínuas apontam as possibilidades de variação para maior, como marca de ênfase, e para menor, como marca de finalização. A linha curva mostra a relação entre o valor médio dominante e a finalização.

A rotina segue o estabelecimento das categorias das frequências e avalia as intensidades formando uma escala de intensidade com apenas três níveis: um médio e dois extremos, cujo distanciamento do médio é da ordem de 50% acima e abaixo. Em seguida, a rotina avalia a duração dos candidatos a pico silábico, estabelecendo uma duração mínima de 4 frames ou 20 ms e máxima de 20 frames ou 100 ms. O critério utilizado para a definição da duração é a manutenção da categoria de frequência. Assim, variações de frequência que ultrapassem a banda definida para a categoria serão tomadas como marcadores de novos picos silábicos. Para a elaboração do gráfico são tomados os valores obtidos e categorizados de frequência e de intensidade com duração suficiente (entre 20 ms e 100 ms). As transcrições mostradas têm somente função de orientação para as análises posteriores e não são observadas em nenhum momento da análise automática.

O objeto

Para o desenvolvimento de nossas análises, buscamos textos previamente gravados e disponíveis na internet⁵ em diversos momentos. A seleção das gravações foi feita mediante o critério de não terem sido filtradas em nenhuma frequência abaixo de 4000 Hz, com particular atenção quanto às frequências menores que 100 Hz, cuja filtragem parece ser muito freqüente nessas gravações. A análise foi feita tomando-se a locução de 20 falantes, sendo que 10 eram de leitura feita por apresentadores de rádiojornalismo e 10 eram de entrevistados. Foram selecionadas 5 frases de cada indivíduo. As entrevistas foram segmentadas, considerando-se somente os trechos que poderiam ser compreendidos como fala espontânea e desconsiderando os trechos de fala do jornalista entrevistador ou trechos de discursos. Assim, foi dada preferência a respostas às perguntas feitas pelo entrevistador. Todas as falas são de voz masculina e têm caráter assertivo. A segmentação das locuções em frases foi feita pelo próprio pesquisador, o que acrescenta um caráter de subjetividade à análise que ainda não conseguimos evitar. Com exceção daquela que foi apresentada como exemplo acima (Figura 1), as frases não foram transcritas.

Os aspectos variáveis que analisamos foram a frequência média da frase (fm), a frequência média dos 5 primeiros núcleos silábicos de cada frase e a frequência média final de cada frase.

Resultados

A comparação entre as médias das frequências médias (n=5 para cada sujeito) estabelecidas para cada sujeito teve uma diferença significativa, pelo teste de ANOVA ($F_0(4,41) < F_0(6,36)$; $P < 0,05$), organizadas pelo coeficiente de concordância de Kendall. De maneira geral, é possível verificar pela Figura 3 que as frequências médias das variáveis referentes à fala espontânea eram mais graves do que as de leitura. O mesmo fenômeno não se manifestou em relação à comparação das variáveis referentes à frequência média das 5 sílabas iniciais médias de cada sujeito e à comparação das variáveis referentes à frequência final média de cada sujeito.

⁵ <<http://www.radiojornalismo.com/midia.htm>>, em <<http://jovempan.uol.com.br/jpamnew>> e em <<http://www.senado.gov.br/radio/>>

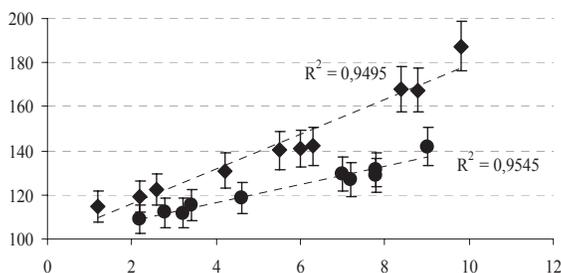


Gráfico 3 – Comparação entre as médias das freqüências médias de cada sujeito (n=5). Os marcadores com formato de losango representam os valores médios das variáveis de leituras e os marcadores circulares representam os valores médios das variáveis de fala espontânea. As linhas pontilhadas representam a tendência linear (R^2) dos conjuntos de valores organizados no eixo horizontal pelo coeficiente de concordância de Kendall

O fato de não haver diferença significativa na freqüência média das 5 sílabas iniciais bem como na freqüência final média, mas haver uma diferença significativa na média da freqüência média entre a leitura e a fala espontânea, aponta para a possibilidade de se interpretar os extremos de variação com ocorrência independente da freqüência média. Para isso verificamos a correlação entre essas variáveis.

A análise de correlação pelo teste de Spearman mostrou que é significativa a comparação das médias da freqüência média e a freqüência média das 5 sílabas iniciais ($R^2=0,73$ e $P<0,001$ para leitura; $R^2=0,63$ e $P<0,001$ para fala espontânea) e a comparação das médias da freqüência média e a freqüência final média ($R^2=0,47$ e $P<0,001$ para leitura; $R^2=0,50$ e $P<0,001$ para fala espontânea). Não foi significativa a correlação entre a freqüência média das 5 sílabas iniciais e a freqüência final média dos sujeitos. Esses resultados corroboram os resultados obtidos na análise anterior, apontando para a possível independência dos extremos entre si, mas não com a freqüência média. Por sua vez, a correlação que se obteve entre as variáveis referentes à freqüência média e a freqüência final é moderada em ambos os casos, enquanto que a correlação entre a freqüência média e a freqüência média das 5 sílabas iniciais está entre moderada e forte, apontando para uma maior vinculação entre essas variáveis do que entre as demais.

A análise isolada da correlação dos dados brutos relativos à leitura e à fala espontânea apresentou resultados semelhantes. Nos dados de leitura, somente a correlação entre a freqüência média e a freqüência média das 5 sílabas iniciais foi significativa ($R^2=0,73$ e $P<0,001$) bem como entre a freqüência média e a freqüência final ($R^2=0,47$ e $P<0,001$), não havendo correlação significativa da freqüência média das 5 sílabas iniciais e a freqüência final da leitura. Nos dados de fala espontânea, os resultados também se repetiram: houve correlação significativa

entre a frequência média e a frequência média das 5 sílabas iniciais ($R^2=0,63$ e $P<0,001$) bem como entre a frequência média e a frequência final ($R^2=0,50$ e $P<0,001$), igualmente sem haver correlação significativa da frequência média das 5 sílabas iniciais e a frequência final da leitura.

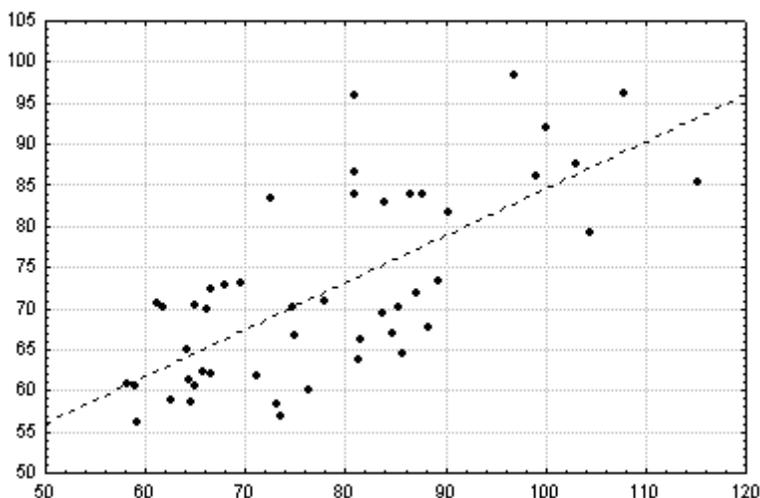
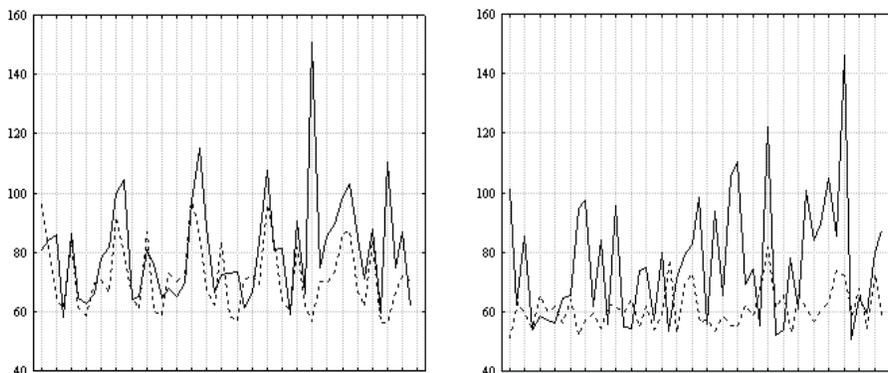


Gráfico 4 – Teste de correlação entre as frequências finais esperadas e as frequências finais obtidas para todas as frases dos sujeitos. A linha pontilhada mostra a tendência linear dessa correlação, com a exceção de duas frases que não apresentaram o mesmo comportamento.

Apesar de termos verificado haver correlação entre a frequência média e a frequência final tanto da leitura quanto da fala espontânea, será necessário, ainda verificar se a relação entre a frequência média e a frequência final estabelece uma cadência final perfeita, semelhante à da música. Para tanto, estabelecemos as frequências finais esperadas cuja razão com a frequência média é de 0,5 (ROEDERER, 2002) e verificamos se os valores obtidos poderiam estar correlacionados com esses (cf. Figura 4). No caso da leitura, houve correlação moderada entre os valores obtidos e o esperados ($R^2=0,48$, $P<0,001$), com a exceção de dois pontos (4%) que não apresentaram o mesmo comportamento; mas, no caso da fala espontânea, a mesma correlação não ocorreu ($R^2=-0,14$, $P=0,31$). O teste-t mostrou diferenças entre frequências esperadas e obtidas de leitura ($t = 2.95$; $p<0.001$) e de fala espontânea ($t = 4.85$; $p<0.001$). A comparação entre as frequências obtidas de leitura e fala espontânea não mostrou diferenças entre elas.



Gráficos 5 e 6 – Comparação entre leitura, à esquerda, e fala espontânea, à direita. As linhas contínuas mostram as frequências obtidas e as linhas pontilhadas mostram as frequências esperadas.

Discussão dos resultados

Os resultados obtidos com essa investigação mostraram que a leitura promove uma intensificação na frequência média no tom da fala que induz a uma similaridade em relação a outras modalidades de fala. Cotejando as características do *maternalês* com a leitura, Kuhl et al. (2001) salientam que também na leitura os espaços vocálicos são normalmente muito maiores, o que apontaria para a hipótese de que o próprio *maternalês*, assim como a leitura, ou a fala em ambiente ruidoso, poderiam fazer parte de uma tendência biológica para falar com mais clareza, tanto quanto a situação o exija. Em relação à prosódia do *maternalês*, entretanto, a autora propõe que essa prosódia, além de tornar a fala exageradamente clara, atrai a atenção da criança e também demonstra atenção. O exagero na caracterização vocálica, por sua vez, acumularia uma função instrutiva.

Na medida em que verificamos que a leitura recebeu um tratamento prosódico diferenciado em relação à fala espontânea e que os textos lidos eram-no feitos por profissionais qualificados com propósitos de divulgação de informação, entendemos que as funções de modalização emocional possam ser descartadas. No entanto, a necessidade de atrair a atenção dos ouvintes para garantir-lhes a compreensão pode ser considerada uma possibilidade de interpretação desse fenômeno. Assim, uma frequência média mais aguda poderia ser tomada como um estímulo natural para a atenção dos interlocutores. Os resultados apresentados neste trabalho não permitem essa confirmação; para tanto, seria necessária uma comparação entre leituras feitas com propósitos diferentes desse da divulgação de informações. Por outro lado, também seria importante o cotejo com as caracterizações vocálicas na leitura e na fala espontânea de língua portuguesa.

Os trabalhos de Patricia Kuhl e seus colegas (KUHL; IVERSON, 1995; KUHL, 2000; KUHL et al., 2001), a partir de seu modelo de *perceptual magnet effect* (PME), pressupõem a existência de uma forma prototípica fixa, não-abstrata, que atue como parâmetro de comparação inicial para todas as demais formas que venham a ser percebidas. Essa proposta de análise traz à tona também o princípio de que os elementos iniciais da aquisição da linguagem serão tomados como formas prototípicas fixas e de que corresponderão a extremos dessas realizações. A adequação do modelo far-se-á no sentido de se estabelecer que o reconhecimento de tons prosódicos estão sujeitos também à formação de categorias ótimas, fundadas em um modelo específico, localizado nos extremos máximos das possibilidades de variação do fenômeno. Nesse caso, imaginar-se-ão que dois tons específicos graves e agudos serão as variações extremas de um tom médio, estabelecendo uma triangulação semelhante àquela do espaço vocálico, definidor das vogais cardinais /i, a, u/.

Por outro lado, é possível recuperar a hipótese de Rosenstock-Huessy (2002, p.40) de que “Lógica e historicamente a linguagem formal antecede a informal.” Dessa maneira, seria possível estabelecer uma diferença entre dois tipos de oralidade: a formal e a informal. A oralidade formal, segundo essa hipótese, antecederia a informal, de maneira que, a se tomar também a proposta de Kuhl e de seus colegas (KUHL; IVERSON, 1995; KUHL, 2000; KUHL et al., 2001), seria a oralidade geradora das formas prototípicas que atuariam como *perceptual magnet effect*. Na medida em que temos a hipótese de que a escrita não é portadora de prosódia entoacional, o falante terá de atribuir essa variação de frequência segundo suas próprias intuições, baseadas totalmente nas formas prototípicas que adquiriu. Nesse caso, a disponibilidade será a das variações de frequência da oralidade formal e não a das variações de frequência da oralidade informal. Essa hipótese justificaria as diferenças encontradas.

FERREIRA NETTO, W.; CONSONI, F. Oral reading and Spontaneous Speech Prosodic Strategies. *Alfa*, São Paulo, v.52, n.2, p.521-534, 2008.

- *ABSTRACT: The aim of this paper is to compare the prosodic analysis of oral reading and spontaneous speech, considering the hypothesis that the typical prosodic aspects of speech are not present in the real manifestations of language. When reading texts aloud, the reader must necessarily create the prosody for the text from his own hypotheses. These hypotheses are based on specific principles, especially on frequency variation. The automatic analysis showed that frequency variation in the texts that were read was higher than in spontaneous speech. Another finding is that the relation between the average and the final frequencies in reading approximates to the relation between dominant and basic tone of the diatonic occidental scale, establishing a perfect harmonic cadence to end the sentences.*
- *KEYWORDS: Phonology. Prosody. Spontaneous speech. Reading.*

Referências

- CHENOWETH, V. Music as discourse. *Word*, New York, v.37, n. 1-2, p.125-134, 1986.
- COMPÔTE, E. *Teoria e prática musical*. Porto Alegre: Escola Superior de Teologia São Lourenço de Brindes; Caxias do Sul: Sociedade Literária São Boaventura: Universidade Caxias do Sul, 1977.
- DENSMORE, F. Scale formation in primitive music. *American Anthropologist*, Lancaster, v.11, n.1, p.1-12, 1909.
- FERREIRA NETTO, W. *Variação de frequência e constituição da prosódia na língua portuguesa*. 2006. 89f. Tese (Livre-Docência)-Faculdade de Filosofia Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2006.
- FÓNAGY, I. Des fonctions de l'intonation : essai de synthèse. *Flambeau*, Tokyo, v.29, p.1-20, 2003.
- GLASER, S. The missing link: conections between musical and linguistic prosody. *Contemporary Music Review*, London, v.19, n.3, p.131-154, 2000.
- KUHL, P. K.; IVERSON, P. Linguistic experience and the "perceptual magnet effect". In: STRANGE, W. (Ed.). *Speech perception and linguistic experience: issues in cross-language research*. Baltimore: York Press, 1995. p.121-154.
- KUHL, P. K.; TSAO, F.-M.; LIU, H.-L.; ZAHNG, Y.; BOER, B. de Language/Culture/Mind/Brain: progress at the margins between disciplines. *Annals of the New York Academy of Sciences*, New York, v.935, n.1, p. 136-74, 2001.
- KUHL, P. K. A new view of language acquisition. *Proceedings of the National Academy of Sciences*, Washington, v.97, n.22, p.1850-1857, 2000.
- LONGACRE, R. E.; CHENOWETH, V. Discourse as music. *Word*, New York, v.37 n.1-2, p.125-134, 1986.
- MENEZES, F. *A acústica musical em palavras e sons*. São Paulo: Ateliê Editorial: Fapesp, 2003.
- ROBERTS, H. H. Melodic composition and scale foundations in primitive music. *American Anthropologist*, Lancaster, v.34, n.1, p.79-107, 1932.
- ROEDERER, J.G. *Introdução à física e à psicofísica da música*. Tradução de Alberto Luis da Cunha. São Paulo: Edusp, 2002.
- ROSENSTOCK-HUESSY, E. *A origem da linguagem*. Tradução Pedro Sette Câmara et. al. São Paulo: Record, 2002.

SCHELLENBERG, E. G.; TREHUB, S. E. Culture-general and culture-specific factors in the discrimination of melodies. *Journal of Experimental Child Psychology*, New York, v.74, n.2, p.107-127, 1999.

SCHOENBERG, A. *Harmonia*. Prefácio, tradução e notas de Marden Maluf. São Paulo: Ed. UNESP, 2001.

T'HART, J. Differential sensitivity to pitch distance, particularly in speech. *Journal of Acoustical Society of the America*, New York, v.69, n.3, p.811-821, 1981.

t'HART, J.; COLLIER, R.; COHEN, A. *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press, 1990.

VAISSIÈRE, J. Les universaux de substance prosodiques In: WAUQUIER, S. (Ed.). *Les universaux sonores*. Rennes: Presses Universitaires de Rennes, [2002?]. Sous presse. Suite à la conférence Sound Universals, Nantes, France. Disponível em: <<http://www.personnels.univ-paris3.fr/users/vaissier/pub/ARTICLES/>>. Acesso em: 9 mar. 2007.

WEBER, M. *Os fundamentos racionais e sociológicos da música*. Tradução, introdução e notas de Leopoldo Waizbort e prefácio Gabriel Cohn. São Paulo: Edusp, 1995.

WERTHEIMER, M. Laws of organization in perceptual forms. In: Ellis, W. D. (Ed.). *A Source Book of Gestalt Psychology*. London: Routledge & Kegan Paul, 1938. p. 71-88. Versão on-line desenvolvida por Christopher D. Green. Disponível em: <<http://psychclassics.yorku.ca/Wertheimer/Forms/forms.htm>>. Acesso em: 9 mar. 2007.

Recebido em dezembro de 2007

Aprovado em junho de 2008