

# Taxa de elocução, grupo acentual, pausas e fonoestilística: temporalidade na prosa e na poesia com interpretação livre

Ana Cristina Fricke Matte<sup>1</sup>

<sup>1</sup>Faculdade de Letras– Universidade Federal de Minas Gerais (UFMG)  
a9fm@yahoo.com

**Abstract.** *The relationship between duration of speech and duration of silent pauses is here investigated in one methodological approach for application in the phonetic and the phonoestilistic fields. Speech rate and duration of VV units in stress groups are correlated with the absolute and normalized duration of silent pauses that delimitates the text analyzed.*

**Keywords.** *Acoustic phonetics; phonestilistics; prosodic modeling; speech rate; silent pause.*

**Resumo.** *Neste trabalho investigou-se a relação entre a duração da fala e das pausas silenciosas intra ou entre sentenças, numa discussão metodológica visando sua aplicação nos campos da fonética e, mais especificamente, da fonoestilística. Taxa de Elocução e duração de unidades vogal a vogal são correlacionadas com a duração absoluta e normalizada das pausas silenciosas que delimitam o segmento de texto estudado..*

**Palavras-chave.** *Fonética acústica; fonoestilística; modelamento prosódico, taxa de elocução; pausa silenciosa.*

## 1. Introdução

Este trabalho foi pensado no âmbito de uma discussão teórico-metodológica da taxa de elocução (speech rate) como variável para estudo de fenômenos de produção/percepção de fala. Em função deste objetivo, todo o processo para obtenção de cada resultado está descrito nos tópicos correspondentes a cada uma das etapas: a) segmentação, etiquetagem e obtenção dos dados; b) duração das pausas; c) relação entre a duração da pausa e o grupo acentual; d) taxa de elocução e pausas. O estudo do grupo acentual (doravante GA) foi incluído porque se trata de uma análise mais fina da variação da duração ao longo das sentenças, enquanto a taxa de elocução (doravante TE) provê valores médios para a sentença inteira. Em função de nosso interesse pelas pausas silenciosas, a TE foi calculada entre trechos de limitados por pausas silenciosas não incluídas para o cômputo. Por esse motivo, poderia ser também chamada de taxa de articulação.

O estudo da fala neutra, que é, há três anos, o foco de nossos estudos em fonoestilística (Matte, 2005), trouxe alguns elementos do que sejam padrões para estudos lingüísticos e fonoestilísticos. O termo fala neutra, recorrente na literatura, refere-se a uma fala, geralmente lida, que procure dissimular efeitos que extrapolem o limite do que seja exclusivamente lingüístico (sintaxe, semântica). Nossos testes indicam que existe possibilidade de controle dos resultados de tal gravação que permita excluir sentenças com padrão estatisticamente diferente da média das outras, com bases

em informação obtida durante o teste de fala neutra. Além disso, dentre os elementos a respeito dos padrões lingüísticos estudados, surge a diferença significativa de padrão rítmico entre prosa e poesia, que aponta para um maior comprometimento da prosa com o conteúdo e da poesia com a forma de exprimi-lo (Matte, 2005). Classificadas pelo teste de fala neutra e condições de gravação, a prosa neutra (sem afeto) e a poesia emotiva (livremente interpretada), com efeito, ocupam praticamente uma mesma posição no gradiente de mudança no padrão do grupo acentual, da menor à maior variância no locutor estudado.

Parte do *corpus* utilizado no presente trabalho focaliza exclusivamente a prosa emotiva, ou, estrito senso, a prosa cuja interpretação ficou a cargo do locutor sem qualquer interpelação outra que interpretar livremente. Esse tipo de leitura, segundo os estudos realizados até o presente momento, seria aquele em que se espera o fortalecimento das proeminências e ênfases e maior concentração no conteúdo do que na expressão, no sentido hjelmsleviano (Hjelmslev, 1968).

A outra parte do *corpus*, constituído de poesias emotivas, dadas as expectativas acima, resultaria de uma negociação entre a marcação silábica, ou periodicidade rítmica, e a variação imposta pela interpretação enfática e emotiva.

No que concerne às pausas silenciosas, não foi observada diferença de duração entre os dois *corpora*, como descrito no tópico correspondente.

## 2. Corpus

O *corpus* compreende a leitura de uma história e quatro poesias por seu próprio autor, um contador de histórias profissional natural de Belo Horizonte e residente em São Paulo, capital, na faixa de quarenta anos. A amostra foi selecionada da quarta sessão de gravação da pesquisa Para uma investigação do padrão prosódico-emocional no português do Brasil<sup>1</sup>, escolhidos os textos nos quais o autor apresentou menor dificuldade de leitura. Este parâmetro foi observado como relevante, após a aplicação do teste de fala neutra (Matte, 2005), pois as sentenças com relato de dificuldade de pronúncia mostraram padrões de duração significativamente diferentes do padrão dito neutro mesmo nos *corpora* de fala emotiva, ou fala com interpretação livre.

Após análise do *corpus* selecionado para o presente trabalho, verificou-se que contém: a) 73 pausas em leitura livre na prosa e 50 na poesia, b) 464 unidades V-V<sup>2</sup> de prosa e 382 de poesia. Diferentes formas de segmentação e etiquetagem, discutidas no próximo tópico, foram utilizadas para fundamentar com sua análise a discussão metodológica aqui empreendida.

## 2. Segmentação, etiquetagem e obtenção dos dados

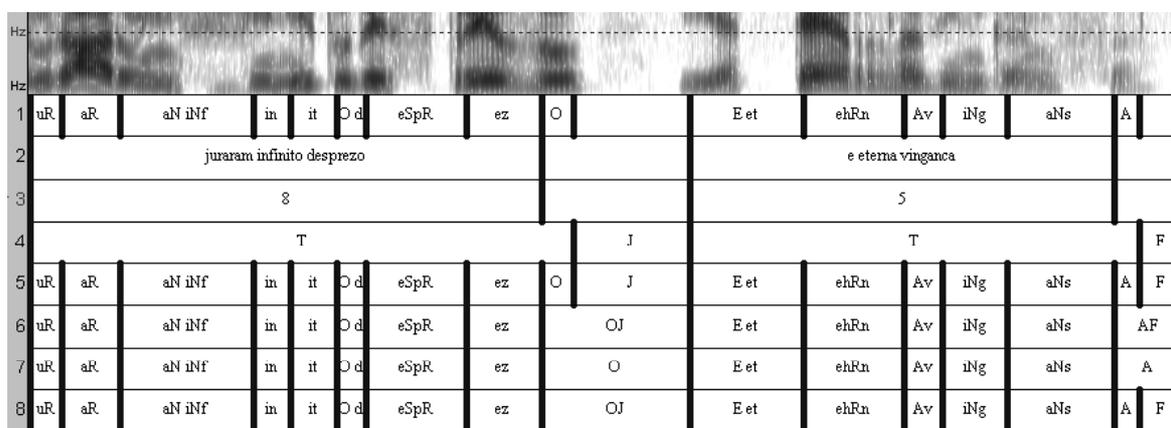
Cada um dos arquivos de som contendo os textos gravados passou pelo mecanismo de extração automática de início de vogal, implementado por Plínio Almeida Barbosa no programa BeatExtractor.psc na forma de scripts Praat. A

---

<sup>1</sup> A coleta de dados foi realizada com apoio da FAPESP, processo 03/11429-6, sob coordenação da autora do presente trabalho.

<sup>2</sup> A unidade V-V baseia-se no conceito de perceptual center, (Marcus, 1981, Barbosa, 2004) e compreende os segmentos fônicos no intervalo delimitado por dois onsets vocálicos consecutivos.

segmentação automática gera um conjunto de fronteiras delimitando início de vogal, e portanto delimitando unidades V-V, na forma de um arquivo do Praat com camadas de texto vinculadas a durações de arquivo de som, o TextGrid. No caso, a segmentação automática gerou a primeira camada aqui estudada, provida com etiquetas fonológicas relativas ao conteúdo fônico das unidades V-V, baseadas em Albano & Moreira (1996). Durante esse processo, a camada VowelOnsets teve sua segmentação revisada a fim de garantir a uniformidade do parâmetro de início de vogal – adotamos como critério o segundo pico de onda periódica da transição consoante-vogal – e a fim de corrigir eventuais erros da segmentação automática, presentes especialmente em grandes variações de f0, frequentes nesse *corpus* de fala emotiva ou interpretação livre. Ainda durante essa primeira fase da segmentação, criou-se uma segunda camada na qual foi delimitada a sentença do início da primeira unidade V-V ao início da última antes da pausa silenciosa (observar as camadas 1 e 2 da Figura 1). Essa nova camada, chamada Frase, recebeu etiquetas com uma transcrição ortográfica simplificada da sentença – sem acentos, pontuação e cedilha –. A camada Frase tem por propósito orientar as análises posteriores e facilitar a localização dos segmentos em eventuais buscas no TextGrid. Para esse fim é necessário que tenha seu início e fim exatamente nos mesmos limites demarcados na camada superior, VowelOnsets (observe, na Figura 1, o segundo trecho de texto, etiquetado na camada 2 como “e eterna vingança”, que começa no início do primeiro V-V entre as pausas silenciosas – /E et/ – e termina no início do último V-V – /A/.



**Figura 1 – Janela de edição do Praat, copiados somente parte do espectrograma e o TextGrid completo, com suas 8 camadas: 1) VowelOnsets ou VV, 2) Frase, 3) Nsil, 4) TextoPausa, 5) SegPausa, 6) VOpi, 7) VVp e 8) VopiF.**

A partir dessas duas camadas outras 5 foram criadas, com pequenas diferenças introduzidas para fundamentar nossa discussão:

a) Camada Nsil: cópia exata da camada frase, com apagamento do texto o qual é substituído pelo número de unidades V-V da sentença. Tem como propósito o cálculo automático da TE em unidades V-V por segundo por meio de script Praat implementado pela autora especialmente para o presente trabalho.

b) TextoPausa: esta camada também é obtida a partir da cópia da camada de frase com apagamento do texto, mas desta vez o limite final de cada sentença é modificado a fim de incluir totalmente a unidade V-V final no trecho de texto. A camada recebe as

etiquetas T = texto, F = pausa externa e J = pausa interna, sendo consideradas externas as pausas entre sentenças e internas as pausas intra-sentença (camada 4 na Figura 1). Estas etiquetas foram escolhidas porque J e F não fazem parte do conjunto de símbolos adotado para cálculo de z-score das unidades V-V na tabela acessória do Praat, originalmente criada por Plínio <sup>a</sup> Barbosa (2004) para uso com o script Praat SGDetector, de sua autoria.

c) SegPausa: cópia exata da camada VowelOnsets, esta camada recebe, além das etiquetas fonológicas já presentes, as etiquetas F ou J nos trechos de pausa silenciosa conforme sejam respectivamente entre ou intra-sentença (camada 5 na Figura 1).

d) VOpi: A sexta camada é obtida pela cópia da camada SegPausa e posterior exclusão dos limites entre a sílaba final de cada sentença e a pausa que a segue, incluindo, assim, nos VVs que antecedem pausas todas as pausas internas (J) e externas (F) (camada 6 na Figura 1).

e) VVp: cópia da camada inicial, VowelOnsets, com incorporação das pausas internas no segmento que as precede sem qualquer notação para pausas (camada 7 na Figura 1).

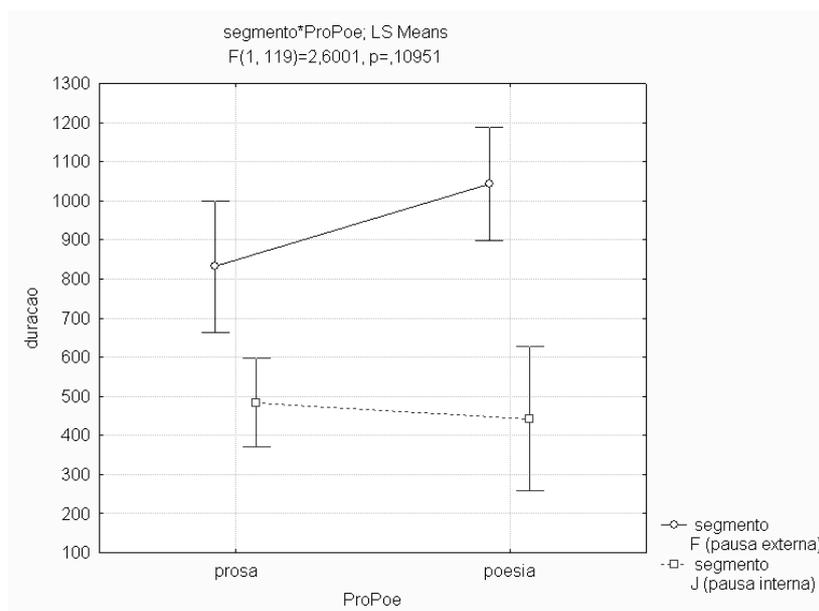
Somente após alguns testes foi buscada outra alternativa de segmentação (camada 8 da Figura 1), que será explicada no tópico “Relação entre a duração da pausa e o grupo acentual”.

Os dados de duração foram coletados no TextGrid previamente segmentado com o auxílio do script SGDetector citado acima, modificado para uso no Praat pela autora especialmente para o tipo de transcrição utilizado e para obtenção das tabelas específicas deste trabalho.

### **3. Duração das pausas**

A observação das pausas tem como objetivo traçar para este corpus específico a média e o desvio padrão de pausas externas e internas para cálculo experimental de z-score (normalização dos dados de duração conforme tabela de durações intrínsecas para os fones do português brasileiro) e z-suavizado (suavização de 5 pontos aplicada à seqüência de dados de z-score). Os resultados obtidos apresentam alto poder de teste (100%).

Observamos num primeiro momento os dados de prosa e poesia em separado, pois nossos trabalhos anteriores sobre duração segmental indicam a necessidade dessa separação em função dos diferentes padrões encontrados para cada gênero (Matte, 2005). No entanto, a ANOVA FACTORIAL ( $F(1, 119)=2,6001$ ,  $p=,10951$ ) para duração das pausas não apresenta diferença significativa entre os dois gêneros, apenas sugerindo haver uma tendência a um aumento na diferença entre os tipos de pausa na poesia em relação à prosa, como pode ser observado no gráfico abaixo (Figura 2). O teste pos hoc Scheffé corrobora o resultado, mostrando diferença significativa entre a) pausa interna e externa da prosa e b) pausa interna e externa da poesia, mas nenhuma diferença significativa dos tipos de pausa se comparados entre os dois gêneros.



**Figura 2 – ANOVA com dois fatores (gênero e tipo de pausa) com os dados de duração em ms das pausas no total do *corpus* de prosa e poesia.**

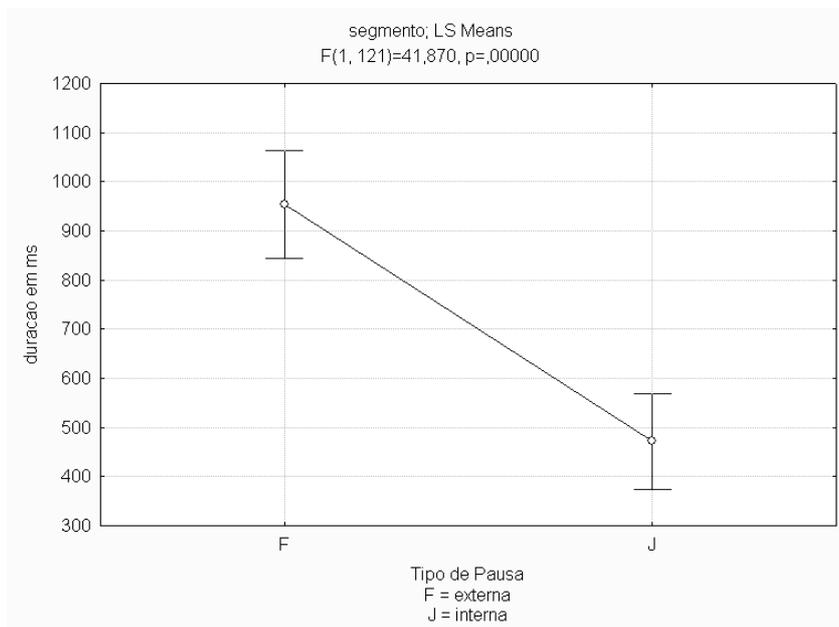
O script SGDetector, para cálculo de z-score e z-suavizado, utiliza uma tabela de referência com médias e desvios-padrão dos segmentos fônicos para o português brasileiro (Barbosa, 2004). Após a análise acima, para obtenção dos dados de duração das camadas SegPausa e VOp<sub>i</sub> acrescentou-se à tabela de referências duas linhas: F e J, com suas respectivas médias e desvio padrão. Embora o procedimento implique considerar a pausa silenciosa como um segmento fônico, decisão polêmica, ele permite trazer elementos para a própria discussão desse tratamento. Cabe notar que, caso a diferença entre prosa e poesia para as pausas tivesse sido significativa, deveríamos ter usado tabelas de referências diferentes para cada caso, o que não foi necessário no âmbito do presente estudo.

A observação da tabela de duração das pausas (Tabela 1) ilustra esses resultados, mostrando que a união dos dois corpora não aumenta o desvio padrão da amostra, nem para pausas internas (variância = 46.483) nem para pausas externas (variância = 323.306), de modo que podemos optar por utilizar em todo o *corpus* para fins de cálculo experimental do z-score das pausas a mesma média e o mesmo desvio padrão.

**Tabela 1 – Média e Desvio Padrão das pausas internas e externas nos *corpus* de prosa, poesia e no conjunto formado por ambos.**

	Média Pausa Interna	DP-J	Média Pausa Externa	DJ-F
Poesia	442,6	146,7	1043,8	640,1
Prosa	483,4	236,9	832,0	439,5
Prosa+Poesia	472,2	215,6	953,6	568,6

A ANOVA de um fator (tipo de pausa) apresenta diferença significativa entre pausas externas e internas, como pode ser observado no gráfico abaixo (Figura 3).



**Figura 3 – ANOVA de um fator (tipo de pausa) no total do corpus, incluindo dados de prosa e dados de poesia.**

Finalmente, cabe observar que, embora a diferença entre as pausas internas e externas fosse esperada, chama a atenção o elevado valor médio das pausas internas, com 95% de probabilidade na faixa entre 372 e 513 ms.

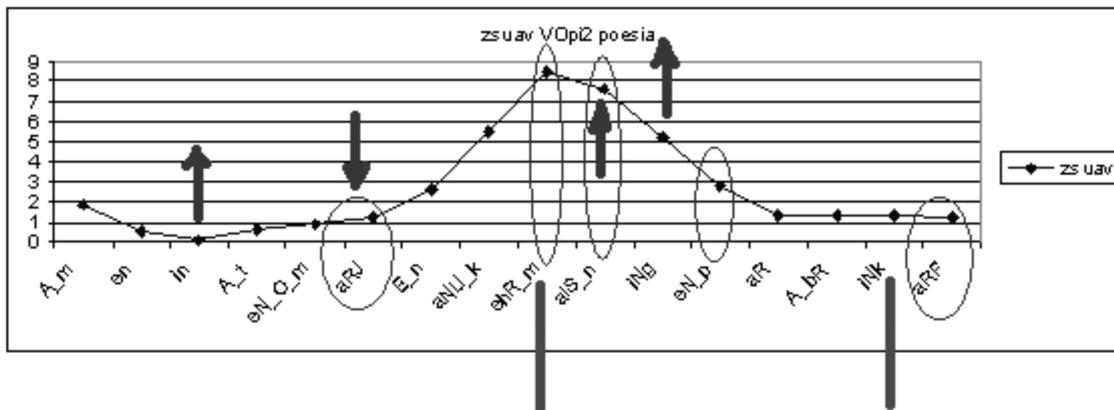
#### **4. Relação entre a duração da pausa e o grupo acentual**

Uma primeira observação do *corpus* foi realizada a título de experiência, com foco em duas sentenças escolhidas aleatoriamente. Os diferentes tipos de segmentação e cálculo de zscore resultaram em diferentes posições para a posição acentuada em cada sentença. Um primeiro teste de percepção com um informante, realizado com 4 escutas seguidas de cada sentença, indicou proeminências percebidas diferentes daquelas detectadas pelo programa, com maior proximidade no uso da segmentação VOp2 – pausas etiquetadas e incluídas no VV que as antecede e cálculo de z-score da pausa utilizando como referência o mesmo valor, das internas, para os dois tipos de pausa.

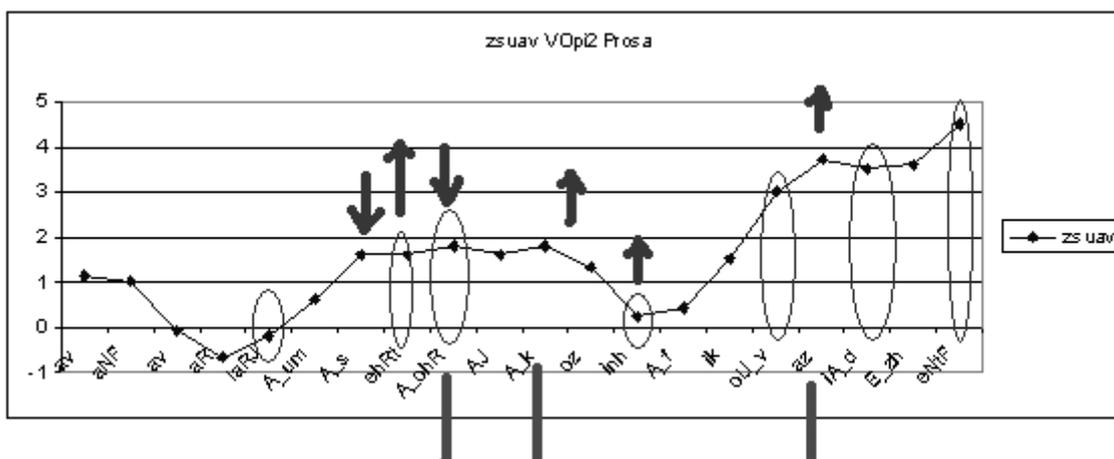
A pausa, portanto, parece ser fator preponderante para o bom funcionamento do programa que detecta os picos acentuais, quando utilizado em amostras de texto completo.

Os gráficos abaixo (Figura 4 e Figura 5) mostram a relação entre os picos e vales de f0 e as proeminências percebidas, contrapostas às marcas de pico de GA obtida pelo cálculo no padrão VOp2, descrito acima. Os dados das Figura 4 e Figura 5 referem-se a uma experiência piloto com um único teste de percepção das sentenças (um único sujeito).

Esse resultado, de forma alguma conclusivo, serve como sugestão de procedimento a ser realizado em futuros testes de percepção com maior número de informantes.



**Figura 4 – Poesia: curva de duração suavizada em uma sentença, com pausas incluídas na unidade VV e z-score de pausas internas e externas baseadas na média das primeiras. Os círculos mostram as proeminências percebidas pelo informante, as setas indicam picos ou vales de f0 e os traços verticais na base do gráfico indicam a posição acentuada conforme o resultado do detector de grupos acentuais. Texto: “A menina tem o mar e não quer mais ninguém para brincar”**



**Figura 5 - Prosa: curva de duração suavizada em uma sentença, com pausas incluídas na unidade VV e z-score de pausas internas e externas baseadas na média das primeiras. Os círculos mostram as proeminências percebidas pelo informante, as setas indicam picos ou vales de f0 e os traços verticais na base do gráfico indicam a posição acentuada conforme o resultado do detector de grupos acentuais. Texto: “Pra variar, a uma certa hora, a cozinha ficou vazia de gente”.**

Uma segunda experiência foi realizada durante o 53.o encontro do GEL, durante nossa apresentação. Foi solicitado aos presentes que, após ouvir três vezes, e confirmar na quarta escuta das sentenças representadas nos gráficos acima, marcassem “os lugares das ênfases de cada sentença”, sem definir previamente se tratava-se de marcar sílaba, vogal ou palavra, tendo sido preferida a marcação de sílabas pelos voluntários. Se considerarmos somente as proeminências detectadas por mais de 67% dos informantes, nenhuma correlação significativa é encontrada. Considerando-se apenas as

proeminências detectadas por 50% ou mais dos voluntários, somente obtivemos correlação significativa entre a duração absoluta e a proeminência tomada categoricamente a partir da escuta dos 6 participantes para a segmentação/análise VV na sentença de prosa ( $R = 0,56$ ,  $p < 0,05$ ). Na sentença de poesia também somente obtivemos correlação significativa para a duração em ms, desta vez nas abordagens VOpi ( $R = 0,72$ ,  $p < 0,05$ ), VOpi2 ( $R = 0,72$ ,  $p < 0,05$ ) e VVp ( $R = 0,72$ ,  $p < 0,05$ ). Este último foi o único caso em que também o z-score mostrou correlação positiva com as proeminências categoricamente percebidas ( $R = 0,62$ ,  $p < 0,05$ ). Em nenhum caso a posição no GA obteve correlação significativa com as proeminências percebidas.

Além disso, observou-se nessa pequena pesquisa piloto que, embora nenhuma das ANOVA de um fator (proeminência categórica) tenha apresentado diferença significativa para a variável dependente posição entre os segmentos percebidos como proeminência e os não identificados como tal, na prosa há uma tendência a aproximar as proeminências da posição acentuada, o mesmo não ocorrendo na poesia. Cabe também notar que há, nessa amostra, maior instabilidade (variância mais elevada) na curva de duração suavizada da prosa do que na poesia, corroborando a expectativa de que a prosa emotiva focaliza sobremaneira o conteúdo, dando lugar a ênfases mais destacadas e perdendo em termos de regularidade. Finalmente, observando-se o f0, nenhuma correlação significativa resultou nem do teste de Pearson nem do teste de Spearman entre os picos de f0 e as proeminências percebidas. Também não obteve-se resultado significativo na correlação entre a posição no grupo acentual e o f0.

Outra observação interessante é relativa às primeiras posições dos grupos acentuais (as mais distantes da posição acentuada), que muitas vezes tem valor mais alto do que a própria posição acentuada ou, em outras palavras, em média duram mais que o próprio acento frasal. Observando-se caso a caso as posições em que isso ocorreu verificou-se que sempre eram posições adjacentes a posições acentuadas de grupo imediatamente anterior cuja duração estava bem mais alongada do que a média das posições acentuadas, ou seja, a queda de uma posição acentuada para o início do próximo GA não foi tão brusca como acontece na fala neutra em geral e também especificamente nesse locutor. A escuta dessas sentenças mostra que o efeito de ênfase pode sobrepor-se não a uma, mas a duas ou até três unidades do tamanho da sílaba da mesma palavra ou não, usando-se, para produzir tal efeito, tanto a curva de f0 quanto o alongamento das durações.

O efeito produzido é, sem dúvida, diferente do efeito produzido pela ênfase localizada em uma única unidade V-V, o que deve ser investigado em trabalhos futuros correlacionando f0 e duração normalizada no grupo acentual.

A diferença de comportamento da duração relativa no GA com e sem pausas pode ser observada com a análise de variância com o fator posição do segmento do GA, aqui somente realizadas para a prosa. O resultado de análises de variância, tanto das realizadas com apenas um fator (ora tamanho do grupo, ora posição do segmento no GA) quanto das realizadas com dois fatores (ANOVA factorial, posição no GA e tamanho do grupo), mostram diferenças de comportamento da duração dependendo da segmentação e da média utilizada para a normalização dos dados de pausa. No escopo deste trabalho foram investigadas sete possíveis abordagens da duração segundo estes parâmetros, sempre com dados de z-suavizado, ou seja, duração normalizada (z-score, que considera a duração intrínseca dos elementos fonológicos que compõem cada

unidade V-V) e suavizada (suavização de 5 pontos sobre os dados de z-score, trazendo para a normalização da duração informações sobre a seqüência na qual encontra-se o segmento) (Cf. Barbosa, 2004). As sete abordagens são: a) *VV* (Figura 1, camada 1): pausas não incluídas na análise. Essa segmentação exclui, inclusive, as consoantes de ataque que sucedem as pausas; b) *VVp* (Figura 1, camada 7): igual à anterior exceto pela inclusão das pausas internas no segmento que as precede, sem qualquer tipo de notação que especifique a pausa; c) *SegPausa* (Figura 1, camada 5): mantém a segmentação *VV* incluindo informação de pausa (a notação insere F para pausas externas e J para pausas internas), portanto com as pausas consideradas como segmentos a mais na seqüência, com valor de unidade V-V no cálculo da normalização. A tabela de referência diferencia a média das pausas internas da média das pausas externas (table 1); d) *SegPausa2* (Figura 1, camada 5): segmentação igual à *SegPausa*, usa a tabela de referência onde a referência para os dois tipos de pausa é a média e o desvio padrão das pausas internas (table 2); e) *VOpi* (Figura 1, camada 6): mantém a notação de *SegPausa* incluindo na segmentação as pausas internas e externas no segmento precedente. Usa table 1 para normalização das pausas; f) *VOpi2* (Figura 1, camada 6): segmentação e notação igual à abordagem *VOpi*, com uso da table 2 para normalização das pausas; g) *VopiF* (Figura 1, camada 8): somente inclui nos segmentos as pausas internas subsequentes. A pausa externa, a título de investigação, foi normalizada com o valor máximo (a 95%) e 50% do desvio padrão das pausas externas.

Cabe notar que a investigação dessas sete diferentes abordagens encontra-se em estágio preliminar. *VOpiF* foi uma tentativa de adequar a segmentação e normalização a arquivos com textos integrais (e não frases isoladas), ainda não satisfatório, em primeiro lugar pela arbitrariedade das informações que baseiam a normalização das pausas externas e, em segundo lugar, porque, assim como para todas as outras abordagens, testes rigorosos de percepção se fazem necessários antes de qualquer conclusão.

As análises de média indicam uma proximidade peculiar do valor da primeira posição do grupo acentual (pós-tônica ou não acentuada) e da última (posição acentuada ou tônica). Os GAs com apenas duas unidades em todas as abordagens não apresentam diferença na duração da unidade átona e da unidade acentuada que os compõem, embora esta última apresente-se discretamente mais alongada que a primeira.

No caso de GAs com três posições (da primeira para a última: -2, -1 e 0), somente a abordagem *VOpi* apresenta o mesmo comportamento das átonas que a *VV*, com a posição [-2] mais alongada que a [-1]. Nas outras abordagens, as posições átonas apresentam-se indiferenciadas.

Já nos grupos com 4 posições (-3, -2, -1 e 0, sendo 0 sempre a posição acentuada), a abordagem *VV* é a única que não diferencia estatisticamente as posições iniciais [-3] e [-2], como seria esperado segundo Barbosa (2004). A curva do *VOpiF* é extremamente suave e condiz, tal como *SegPausa*, *SegPausa2* e *VOpi*, com o comportamento esperado para o z-suavizado ( $-3 > -2 > -1 < 0$  e  $0 > -3$ ). Cabe observar que nenhuma das ANOVAs realizadas com os tamanhos 2 a 4 de GAs tem  $p < 0,05$ , portanto há necessidade de coletar um maior número de dados para qualquer tipo de conclusão.

A ANOVA factorial para os tamanhos de 5 a 8 também não é significativa ( $p > 0,05$  em todas as abordagens). Todas as curvas mostram-se mais suaves que nos

grupos menores. Nota-se que para a maioria das abordagens a primeira posição (não acentuada e mais distante do acento) possui z-suavizado semelhante à posição acentuada [0], o que só não ocorre com VOpi2.

Observando-se somente o tamanho do GA, o resultado esperado é a horizontalidade em tamanhos de 4 a 9 e z-suavizado aumentado nos tamanhos 2, 3 e maiores que 9. As abordagens que apresentaram tal comportamento foram VV ( $p = 1,00$ ), SegPausa ( $p = 0,92$ ), VOpi ( $p = 0,64$ ), VOpi2 ( $p = 0,94$ ) e VOpiF ( $p = 0,34$ ). Todas as abordagens, na ANOVA de um fator [posição], apresentam padrões semelhantes nas posições distando até 6 unidades (posição -6) da posição acentuada e significância estatística.

A análise destes dados de média na prosa indica a abordagem VOpi como a melhor segundo esses parâmetros, embora deva-se enfatizar a necessidade de testes de percepção e maior número de análises locais (sentença por sentença) a fim de orientar a opção de segmentação e normalização em *corpora* de textos completos. Em segundo lugar, com bons resultados, temos as abordagens VV, SegPausa, VOpi2 e VOpiF. O VVp não apresentou bons resultados em nenhuma das ANOVA realizadas, o que serve como alerta, embora de caráter preliminar, quanto ao uso indiscriminado da inclusão da pausa interna na sílaba fonética precedente.

## 5. Taxa de Elocução e Pausas

Concluindo, algumas observações se fazem necessárias acerca da TE. A taxa de elocução nesse *corpus* é bastante estável, pois há uma alta correlação entre o número de sílabas e a duração da sentença (para a poesia:  $R = 91\%$ ,  $p < 0,05$  e para a prosa:  $R = 96\%$ ,  $p < 0,05$ ). Além disso, a TE depende do gênero (se é prosa ou poesia). Segundo ANOVA de um fator (gênero) para a TE, a média para a prosa é de 4,1 VV/s e para a poesia é de 4,6 VV/s, com diferença significativa entre elas ( $p < 0,04$ ). O teste pos hoc Scheffé corrobora a diferença indicada pela ANOVA.

A correlação entre a duração absoluta das pausas subsequentes ao trecho de texto referente à TE e o valor da própria TE é praticamente nula (-5%) e não significativa tanto na prosa quanto na poesia. O mesmo é observado se a relação é feita com as pausas que antecedem o trecho.

## Referências

- ALBANO, E. C. and Moreira, A. A. Archisegment-based letter-to-phone conversion for concatenative speech synthesis in Portuguese. Proceedings ICSLP'96, v.3, 1996, pp. 1708-1711.
- BARBOSA, Plinio Almeida. "Elementos para uma tipologia do ritmo (lingüístico) da fala à luz de um modelo de osciladores acoplados", In *Cógnito – Cadernos Romênicos em Ciência Cognitiva* 2(1), 2004, pp. 31-58.
- HJELMSLEV, Loius. *Prolégomènes à une theorie du langage*. Paris: Minuit, 1968.
- MATTE, Ana Cristina Fricke. Seria, a poesia, fala emotiva por excelência. In: 1.o Simpósio Internacional de Cognição e Artes Musicais, 2005, Curitiba. Anais do 1.o Simpósio de Cognição e Artes Musicais. Curitiba: Deartes/UFPR, 2005. V.1, p. 71-78.