

SEGMENTAÇÃO AUTOMÁTICA E MANUAL DE NARRATIVAS ESPONTÂNEAS METODOLOGIA DE UM ESTUDO DE PERCEPÇÃO DA PROSÓDIA

*Ebson Wilkerson R. da Silva¹
Miguel Oliveira, Jr.²*

1. INTRODUÇÃO

Este capítulo apresenta os procedimentos metodológicos utilizados em uma pesquisa experimental sobre a percepção de fronteiras prosódicas. O objetivo da pesquisa foi verificar em que medida ouvintes reconhecem essas fronteiras e se as fronteiras reconhecidas coincidem com fronteiras prosódicas detectadas semiautomaticamente através de *scripts* do aplicativo *Praat*.

O capítulo pode ser lido como um guia para experimentos que envolvem percepção de fronteiras prosódicas da fala. Aqui apresentamos os passos que foram seguidos para a coleta dos dados e para a construção da pesquisa. Esses passos podem ser adaptados de acordo com as características dos estudos a serem desenvolvidos.

Organizamos o conteúdo deste capítulo da seguinte maneira: primeiro, apresentamos a descrição dos materiais de pesquisa e do experimento com os participantes, bem como tratamos do *corpus* da pesquisa que serviu como

¹ UFAL – Universidade Federal de Alagoas.

² UFAL – Universidade Federal de Alagoas.

base para a coleta do experimento de percepção; em seguida, descrevemos o experimento de percepção e o procedimento para seleção dos informantes, assim como o processo de coleta de dados de percepção; na seção seguinte, apresentamos os métodos utilizados para se obter a segmentação semiautomática, baseada em elementos prosódicos, como a duração e grupos acentuais, e também apresentamos conceitos de grupos acentuais e de unidades VV.

2. NARRATIVAS ESPONTÂNEAS

Quando os participantes de uma conversa decidem contar uma história durante uma conversa, sem que sejam solicitados a fazê-la, deixam transparecer o julgamento que fazem acerca do que acreditam ser suficientemente interessante para justificar o longo turno conversacional que ocuparão. Por essa razão, narrativas espontâneas constituem um excelente material de estudo da fala em contexto naturalístico, uma vez que há nesse tipo de produção uma motivação pragmática.

De acordo com Wolfson (1979), as pessoas sabem que regras são apropriadas em entrevistas como eventos de fala. Isso explicaria o fato de que algumas narrativas, quando eliciadas, são mais propensas a assumir a forma de resumos, pois são geralmente mais curtas, pontuais e exibem pouquíssimos detalhes, assim como respostas às perguntas em uma entrevista devem ser.

Para o nosso estudo, utilizamos um *corpus* de entrevistas espontâneas (WOLFSON, 1979). Na ocasião, os participantes foram solicitados a falar livremente sobre qualquer tema a partir de uma lista composta por 28 tópicos (Oliveira, 2000). Um total de 17 narrativas apareceram naturalmente nas entrevistas, na maioria das vezes, como uma ilustração de um dado argumento ou tópico.

Desse *corpus*, dez narrativas foram segmentadas, mas apenas seis foram selecionadas aleatoriamente para o estudo: duas foram utilizadas à guisa de ilustração e quatro narrativas foram separadas para o experimento de percepção³. As narrativas selecionadas derivam de entrevistas de quatro diferentes sujeitos: um homem e três mulheres, variando de 25 a 37 anos de idade, naturais de diferentes regiões do Brasil. No momento da coleta de dados, todos eram estudantes de pós-graduação.

As narrativas selecionadas tinham uma duração média de 69 segundos. Essas narrativas foram enumeradas (de 1 a 4), transcritas ortograficamente, de uma

³ As narrativas podem ser ouvidas no seguinte repositório: http://www.archive.org/details/Pesquisa_273.

maneira linear, sem marca de pontuação e sem indicação de paragrafação, para que a organização da estrutura do texto escrito não influenciasse os participantes na segmentação manual das narrativas. A Figura 1 mostra um exemplo de como as narrativas eram apresentadas aos participantes da pesquisa.

Figura 1 – Exemplo de narrativa espontânea transcrita ortograficamente. Exemplo 1

com a naty já aconteceu de tudo já nesse lance uma vez ela tava descendo na grouse fazendo um ziguezague daí tinha um uma criancinha também tava descendo né pequeninha acho que tinha uns quatro ou cinco anos fazendo ziguezague também daí no bem no meio da montanha os dois se encontraram assim se chocaram de frente assim daí caiu um pra cada lado assim daí uma pessoa levantou chacoalhou tirou a neve e foi embora né a outra ficou chorando então a naty ficou chorando lá a criancinha eu pensei que ela f tinha machucado né que nada levantou tirou a neve e foi embora e a naty ficou deitada lá chorando que pensou que tinha machucado a criança

3. MONTAGEM DO EXPERIMENTO DE PERCEPÇÃO

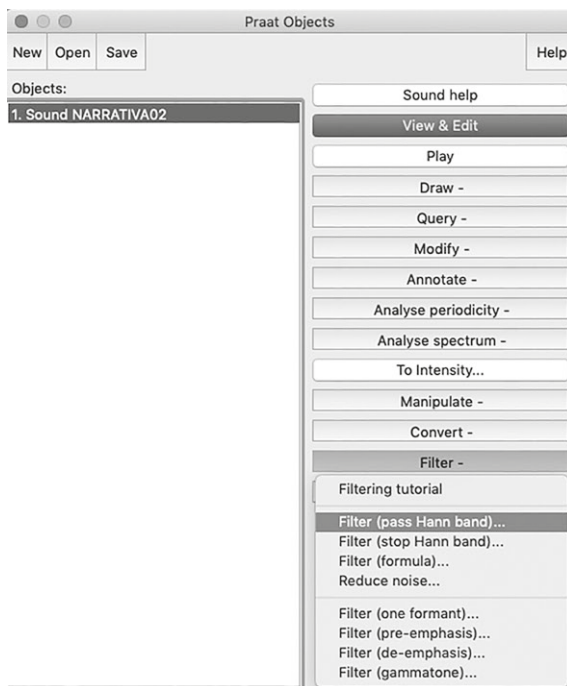
Seguindo com o exemplo do estudo que serviu como modelo para esse artigo, explicaremos os passos para montar o experimento de percepção utilizado na pesquisa.

Para o experimento de percepção, as narrativas foram divididas em quatro diferentes condições de apresentação (C):

- Na Condição 1 (C1), foi apresentada apenas a transcrição ortográfica da narrativa, feita de uma maneira linear, sem marca de pontuação e sem indicação de paragrafação.
- Na condição de apresentação 2 (C2), foram apresentados simultaneamente a transcrição da narrativa e o áudio a que a transcrição se refere.
- Na condição de apresentação 3 (C3), apenas o áudio da narrativa foi apresentado.
- Na condição de apresentação 4 (C4), foi apresentada uma versão deslexicalizada da narrativa.

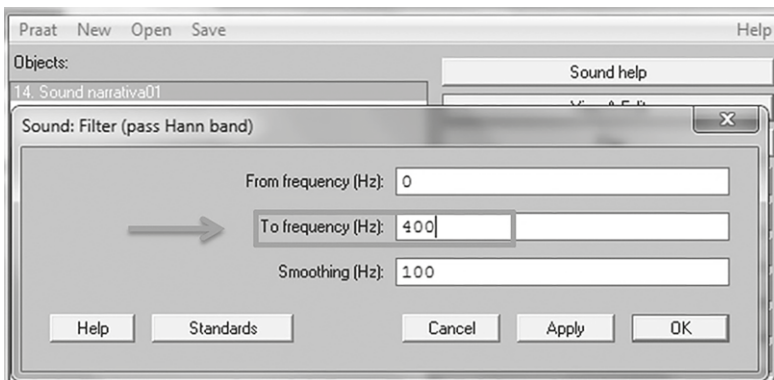
Para produzir a C4, ou seja, a versão deslexicalizada, o arquivo de áudio original foi filtrado com filtro passa-banda, selecionado através da função *filter* no aplicativo *Praat*, como ilustra a Figura 2:

Figura 2 – Janela do Praat: Utilizando filtro passa-banda.



A partir dessa filtragem, toda a informação de frequência superior a 400 Hz foi descartada (cf. Figura 3). Ou seja, a filtragem do áudio resulta em uma fala ininteligível, que preserva, no entanto, a informação prosódica.

Figura 3 – Seleção de valores de frequência, valores acima de 400 Hz são descartados.



As narrativas foram assim divididas para serem apresentadas aos participantes que fizeram a segmentação manual das narrativas. Sendo assim, teremos a segmentação feita da seguinte maneira: i) da leitura das narrativas (C1), que pode ser útil para verificar a atuação da prosódia implícita e ajudar a validar o modelo de intenções; ii) da leitura e audição da narrativa (C2); iii) exclusivamente da audição dos áudios das narrativas (C3); e iv) dos áudios *deslexicalizados*, que podem ser úteis para confirmar se os participantes basearam a sua segmentação por elementos prosódicos e se essa segmentação coincide com a segmentação das narrativas lidas e, além disso, se a segmentação manual coincide com a segmentação automática.

A seguir, explicaremos a sequência utilizada para segmentação manual, segmentação automática e seleção dos participantes.

4. SEGMENTAÇÃO MANUAL

Para a segmentação manual, nos baseamos no modelo teórico de Labov (1972). Nesse modelo, a narrativa é considerada um tipo de discurso que pode ser segmentado, tomando-se como princípio a função informacional de suas partes constituintes. De acordo com Labov (Idem), a estrutura de uma narrativa “bem formada” pode conter as seguintes seções: resumo, orientação, seção de complicação, avaliação, resolução e coda.

Entretanto, esse modelo laboviano precisa ser empiricamente validado, se quisermos considerá-lo como exemplo de modelo teórico estrutural de narrativas orais. Para isso, conforme Oliveira Jr. (2000), consideramos o nível suprasegmental do discurso narrativo. Um dos modelos estruturais mais influentes em estudos computacionais é aquele proposto por Passonneau e Litman (1993). Este modelo baseia-se na hipótese de que os falantes segmentam as narrativas que contam com o propósito – ou intenção – de deixar claro para o ouvinte a sua organização estrutural.

Esse modelo de intenções foi testado e validado para o português brasileiro no estudo piloto de Oliveira e Silva (2011). Além disso, foi verificado que, de modo geral, as pessoas reconhecem a localização de fronteiras discursivas e que a concordância tende a ser maior quando considerado o material sonoro das narrativas.

Assim, com o modelo validado, convidamos os participantes para a pesquisa. Vale ressaltar que, no entanto, em pesquisas que envolvem percepção, é necessário estabelecer critérios de inclusão dos participantes. Para que os participantes fizessem o solicitado nas instruções para que não houvesse dificuldades para

realizar a tarefa, utilizamos como critérios para inclusão do participante o grau de escolaridade – terceiro grau completo ou incompleto – e saúde auditiva.

A tarefa solicitada aos participantes foi indicar, nas narrativas, os pontos em que julgassem haver intenção, por parte do falante, de finalizar uma unidade comunicativa. Nenhuma definição para a noção de unidade comunicativa foi apresentada. Os participantes foram instruídos a julgar quais eram as fronteiras dessas unidades em caráter puramente subjetivo, tendo como base, apenas, os exemplos de narrativas segmentadas por intenção, que eram apresentados antes de fazerem as suas próprias segmentações, conforme exemplo apresentado no Quadro 1. O objetivo foi comparar esse julgamento dos informantes com uma segmentação puramente prosódica.

Todos os participantes leram e assinaram um Termo de Consentimento Livre e Esclarecido (TCLE), de acordo com a exigência do Ministério da Saúde e do Comitê de Ética em Pesquisa.⁴

Após a assinatura do TCLE, os participantes receberam instruções gerais impressas acerca do experimento, seguidas de exemplos: duas narrativas adicionais, extraídas do *corpus* supramencionado.

Os exemplos traziam narrativas segmentadas, seguidas de uma análise de suas partes constitutivas, considerando-se as possíveis intenções do falante, conforme o Quadro 1. Desse modo, o participante entenderia o que foi solicitado, quando estivesse diante das condições em que teria acesso às informações da narrativa. As segmentações e análises foram feitas pelos pesquisadores apenas à guisa de ilustração, e não como modelos a serem seguidos. Também, a título de ilustração, uma das narrativas foi apresentada em sua versão filtrada, para que os informantes se familiarizassem com o que ouviriam na C4.

⁴ Estudo avaliado pelo Comitê de Ética em Pesquisa, aprovado para realização em 24/11/2010 (Processo n. 013526/2010-01).

Quadro 1 – Modelo de narrativa apresentada como exemplo aos participantes, segmentada em unidades comunicativas com base em intenções. Exemplo 2

pai também teve uma experiência assim de ele tava passando por uma rua a gente chama de portão do gelo esse lugar		<i>introduz a narrativa</i>
e agora ele tá bem mudado fizeram muita casa mas antes ele era muito cheio de árvore de mato não de árvore de mato era matagal		<i>descreve o lugar atual vs. passado em que aconteceu a narrativa</i>
e pai vinha por ali de noite e ele começou a ver a sentir ouvir passos atrás dele né e ele olhou pra trás olhou pro lado e não viu ninguém atrás dele aí ele começou a correr assustado e sentindo que a coisa vinha continuando perseguindo ele né vinha atrás dele correndo e ele correndo e a coisa correndo também		<i>apresenta o que aconteceu a seu pai</i>
uma experiência paranormal		<i>conclui dando um tema para a narrativa</i>

Os participantes foram então informados de que receberiam quatro diferentes narrativas em quatro diferentes condições (cf. Seção 3.1 O *corpus*). Foram solicitados a ouvir, ler, ou ler e ouvir cada uma das narrativas com atenção e, em seguida, segmentá-las de acordo com o julgamento que fizessem acerca da intenção do falante de finalizar uma unidade comunicativa. Eles poderiam levar o tempo que julgassem necessário para fazer a segmentação de cada uma das narrativas, podendo também ouvir as versões em áudio quantas vezes julgassem necessário para completar a tarefa.

Nas condições C1 e C2, em que receberam uma transcrição da narrativa, os participantes foram solicitados a fazer a segmentação na transcrição, por meio de barras transversais no papel. Nas condições C3 e C4, em que foi exposta apenas a versão em áudio da narrativa, os participantes foram solicitados a fazer a segmentação simultaneamente à execução do áudio, pressionando a tecla *Enter* de um teclado de computador. As respostas aos estímulos apresentados nessas duas condições (3 e 4) foram registradas no aplicativo computacional *ELAN* (BRUGMAN e RUSSEL, 2004).⁵

Durante todo o experimento, seguiu-se de forma semialeatória a ordem de apresentação das narrativas em diferentes condições aos informantes, ou seja, cada uma das quatro narrativas que compõem o *corpus* do estudo, em diferentes condições experimentais, apareceu três vezes em cada ordem de apresentação

⁵ O *ELAN* é um aplicativo computacional desenvolvido pelo Max Planck Institute for Psycholinguistics: <http://tla.mpi.nl/tools/tla-tools/elan/>.

(4 narrativas x 4 condições x 3 ordens = 48 grupos de apresentação únicos). A apresentação semialeatória dos estímulos assegura a generalidade dos resultados, conforme descrito no Quadro 2.

Quadro 2 – Lista randomizada de apresentação das narrativas e suas respectivas condições

PARTICIPANTE	Ordem de apresentação das narrativas			
	NAR_CONDIÇÃO	NAR_CONDIÇÃO	NAR_CONDIÇÃO	NAR_CONDIÇÃO
1	N1_C1	N3_C4	N2_C3	N4_C2
2	N1_C2	N4_C1	N2_C4	N3_C3
3	N1_C3	N4_C2	N3_C1	N2_C4
4	N1_C4	N4_C3	N3_C2	N2_C1
5	N2_C1	N4_C4	N3_C3	N1_C2
6	N2_C2	N1_C1	N3_C4	N4_C3
7	N2_C3	N1_C2	N4_C1	N3_C4
8	N2_C4	N1_C3	N4_C2	N3_C1

O Quadro 2 apresenta a ordem e a condição em que foram apresentadas as narrativas aos participantes. Por exemplo, o informante 1 recebeu a Narrativa 1, na C1 – em que teria acesso apenas à transcrição; depois ouviria a Narrativa 3, na condição 4, em que teria acesso ao áudio deslexicalizado e assim por diante.

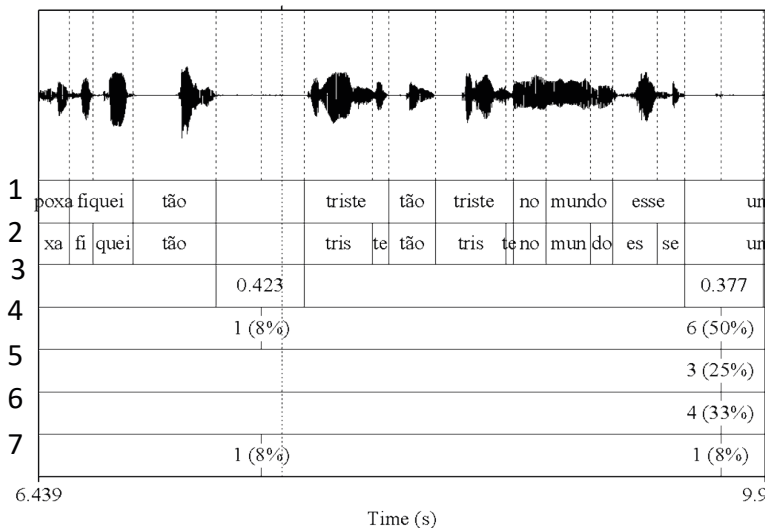
Em seguida, as respostas dos participantes foram salvas, separadas por narrativas e por condições. Cada uma das narrativas foi transcrita ortograficamente em uma tabela do Excel, segmentada por palavras, conforme é possível verificar no Tabela 1, a seguir. Os dados foram, então, computados, isto é, foram marcadas, ao final de cada palavra, a porcentagem de informantes por condição que julgou haver ali uma fronteira de unidade comunicativa. Também foram informadas na tabela a ocorrência e a duração de pausas em cada narrativa.

Tabela 1 – Computação dos dados de fronteiras percebidas pelos informantes

TRANSCRIÇÃO NARRATIVA	OCORRÊNCIA DE PAUSA	DURAÇÃO	FRONTEIRAS N4_01	FRONTEIRAS N4_02	FRONTEIRAS N4_03	FRONTEIRAS N4_04
Acidentes	Não	**	25%	8%	**	**
Aconteceu	Não	**	25%	16%	**	**
Um	Não	**	**	**	**	**
Acidente	Não	**	41%	8%	**	**
Mas	Não	**	**	**	**	**
Eu	Não	**	**	**	**	**
Não	Não	**	**	**	**	**
Lembro	Não	**	50%	25%	8%	**
Eu	Não	**	**	**	**	**
Tinha	Não	**	**	**	**	**
Dois	Não	**	**	**	**	**
Anos	Sim	0.837	41%	75%	100%	91%

A Tabela 1 apresenta um exemplo de segmentação para a narrativa 4, nas 4 condições. A tabela serve apenas como ilustração, a tabela completa foi publicada na seção de anexos em Silva, 2017. Após esta etapa, os dados foram transcritos no *Praat*.

Figura 4 – Tabela de transcrição do *Praat*.



Na Figura 4, temos uma tela do *Praat* que apresenta: na linha 1, a transcrição segmentada em palavras; na linha 2, a segmentação em sílabas; na linha 3 há a indicação de duração de pausas. Nas linhas 4, 5, 6 e 7, estão os valores de concordância (em porcentagem) nas quatro condições experimentais do estudo.

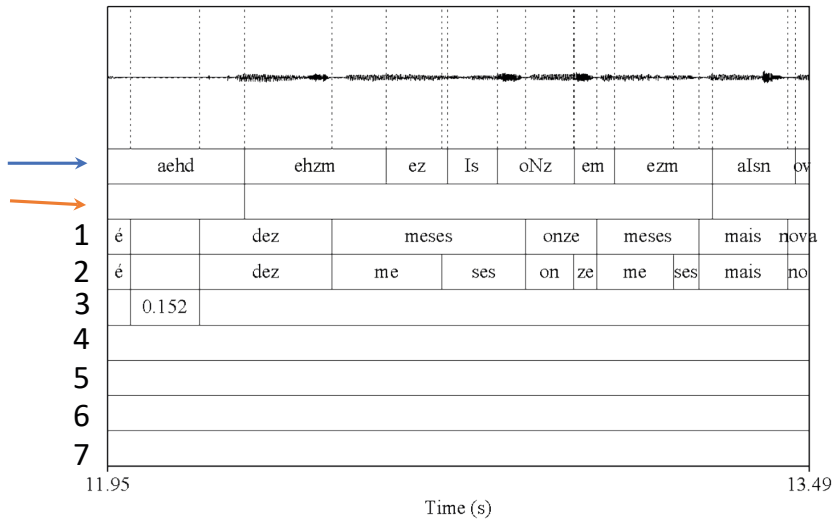
5. SEGMENTAÇÃO AUTOMÁTICA

Além da segmentação dos participantes, a qual foi realizada considerando as intenções do falante de terminar uma unidade comunicativa, também foi realizada uma segmentação por *scripts*. O uso dos *scripts* justifica-se por fornecer uma segmentação semiautomática baseada apenas em elementos prosódicos.

Para a segmentação dos grupos acentuais, utilizamos o *SG Detector* (BARBOSA, 2006). O *SG Detector* segmenta a narrativa considerando aspectos prosódicos, logo esta segmentação nos foi útil para comparar com a segmentação dos participantes. Entretanto, para que o *script* funcionasse corretamente, foi necessário fazer uma segmentação em unidades VV, por isso foi utilizado o *script BeatExtractor*, aplicativo implementado por Cummins e Port (1998) e adaptado para o *Praat* por Barbosa (2006). Ao segmentar o áudio em unidades VV, o *BeatExtractor* pode apresentar erros que precisam ser corrigidos manualmente. Assim, a segmentação foi semiautomática. Os conceitos de grupos acentuais e unidade VV e o funcionamento dos *scripts* serão discutidos na próxima seção.

Como é possível observar na Figura 5, a seguir, além das linhas já conhecidas na Figura 4, há também dados de transcrição segmentados em unidades VV (como indicado na linha em que temos a seta azul) e informações de grupos acentuais (*stress groups*) (é possível identificar a linha com a marcação dos grupos pela seta vermelha). A segmentação em unidades VV foi obtida através do *script Beat Extractor* e a segmentação em grupos acentuais através do *SG Detector*.

Figura 5 – Tabela de transcrição do Praat.



5.1 Grupos acentuais

Os grupos acentuais ou *stress group* (SG) são definidos em Lucente (2012, p. 96) como “unidades delimitadas por dois acentos frasais consecutivos”. O acento frasal é definido em Barbosa (2006) como uma proeminência do domínio da produção da fala, caracterizada por um pico local de duração ao longo do enunciado.

Barbosa (2006) apresenta três tipos de técnica que podem ser empregadas para delimitar os grupos acentuais a partir da localização dos acentos frasais, dependendo do interesse da pesquisa. A primeira é de base perceptiva e pode ser feita pela escuta do *corpus* de pesquisa por uma bateria de ouvintes, aos quais se solicita marcar as palavras que lhes parecem proeminentes. No entanto, essa técnica possui algumas dificuldades, a principal delas é definir a partir de qual porcentagem é possível considerar que há uma proeminência correspondente a um acento frasal que corresponde a uma fronteira de grupo acentual.

Outra técnica parte dos picos de duração de unidades VV ao longo do enunciado para selecionar, entre as palavras em que ocorreram os picos, as que são reconhecidas por pelo menos 20% dos ouvintes. Essa escolha também é arbitrária, porém é guiada por uma seleção inicial fundamentada em critérios de produção.

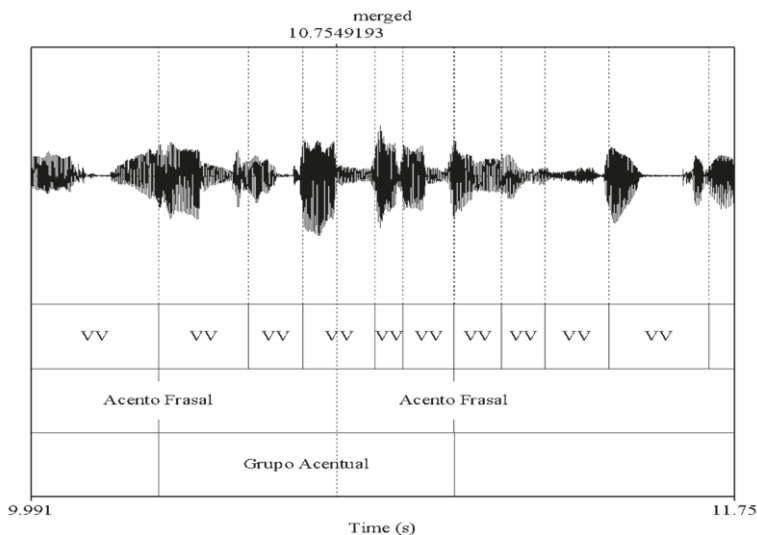
A terceira técnica é fundamentada apenas em dados de duração produzida e, segundo o autor, é perfeitamente adequada para o estudo da produção da organização temporal do enunciado. Essa técnica possui três etapas automáticas,

mas para isso é necessário um procedimento de segmentação e etiquetagem de unidades VV, o que justifica o uso do *BeatExtractor* nesta pesquisa.

As etapas automáticas correspondem a: a) normalização da duração das unidades VV, que consiste em obter um valor de *z-score*;⁶ b) suavização, que, de acordo com Barbosa (2006), consiste em atenuar as variações locais de duração advindas da queda de duração em unidades VV pós-tônicas e/ou de durações de fones muito distintos da relação de durações dos fones do português brasileiro; e c) cálculo dos picos de máximo do *z-score* estendido suavizado através da detecção das posições de transição de derivada positiva para negativa (ponto de máximo).

São essas três etapas que foram implementadas no *script SG Detector*, desenvolvido por Barbosa (2006), para segmentar automaticamente o enunciado em grupos acentuais. O *SG Detector* também gera uma camada com a segmentação dos grupos acentuais no *Praat*, conforme a Figura 6:

Figura 6 – Unidades VV, acentos frasais e grupos acentuais.



Descritos os passos tomados para a obtenção dos dados da pesquisa, têm-se, por fim, duas segmentações, uma obtida através do julgamento de informantes

⁶ Essa normalização, entretanto, corresponde a uma minimização dos efeitos da duração intrínseca (um segmento pode durar mais porque é simplesmente intrinsecamente mais longo) e da variabilidade da duração (quando um segmento dura mais simplesmente porque é mais variável).

sobre a intenção do falante em terminar uma unidade comunicativa, isto é, uma segmentação baseada em intenções, e outra obtida através do uso dos *scripts* desenvolvidos por Barbosa (2006), o *BeatExtractor* e o *SG Detector*. A segmentação com o uso dos *scripts* é de caráter estritamente fonético, ou seja, apenas elementos fonéticos serão considerados para a segmentação semiautomática. Essa segmentação é de grande relevância para estudos que pretendem analisar, a partir de uma segmentação automática da narrativa, elementos prosódicos associados a essa segmentação.

5.2 Unidades VV

Como explicado em seções anteriores, para proceder à segmentação das narrativas em grupos acentuais utilizando o *SG Detector*, fez-se necessário segmentá-las em unidades VV. Portanto, cabe aqui uma breve explicação sobre as unidades VV e o *script* utilizado para fazer a segmentação.

Para Barbosa (2021), a unidade VV (*vowels onset*) é a unidade mínima da programação rítmica da fala, a partir dela se organiza o grupo acentual. Essa unidade é do tamanho da sílaba e contempla a sequência de intervalos a partir do *onset* da vogal até o *onset* da vogal imediatamente seguinte, incluindo todas as consoantes e glides entre as duas. Estudos como os de Pompino-Marschall (1991) e Barbosa (1996) mostram a importância da transição CV para a produção e percepção da fala. Devido à necessidade de manter estável a sequência de transição CV, há uma tendência de preservação da duração das unidades entre as transições (unidades VV). Com o intuito de manter a estrutura rítmica e entoacional do enunciado, essas unidades resistem à variação mais do que a sílaba, mantendo os demais parâmetros constantes. No que diz respeito à percepção, observando o modo como a fala é processada, utilizar a unidade VV seria metodologicamente mais adequado do que a sílaba, de acordo com estes estudos.

Para segmentar as narrativas em unidades VV, é preciso utilizar, como dito anteriormente, o *script BeatExtractor*. Esse *script* foi adaptado por Barbosa (2004) para uso no *Praat* e está disponível na forma de programa aberto, via solicitação ao autor. O *script* detecta automaticamente *onsets* de vogal, através de um algoritmo sugerido por Scott (1993), a partir de quatro etapas que, de acordo com Barbosa (2004, p. 25), são as seguintes:

- filtragem do sinal de fala pelo uso de um filtro passa-faixas com valores default das frequências de corte para um locutor masculino fixadas entre 1000 e 1800 Hz. Os valores default para um locutor feminino são 1150 e 2100 Hz. Essa filtragem preserva a energia na faixa entre o primeiro e segundo formantes. O filtro

default é o filtro Butterworth de segunda ordem, mas a escolha por um filtro Hanning com faixas de transição mais abruptas do que a do Butterworth pode ser escolhida opcionalmente;

- retificação do sinal previamente filtrado;
- filtragem passa-baixas num valor default da frequência de corte de 20 Hz (preserva transições mais bruscas de amplitudes no PB – por exemplo, no caso de “tap” [ɾ] seguido de vogal);
- localização das fronteiras pela recuperação dos pontos que satisfazem duas restrições: (1) ponto de derivada máxima local positiva superior a 15% do valor máximo de derivada do sinal obtido após a aplicação da etapa anterior, e (2) ponto de amplitude superior a 12% da amplitude máxima obtida após a aplicação da etapa anterior.

Após a localização das fronteiras, elas foram gravadas em um arquivo do *Praat* para registro de marcação de fronteiras e rótulos de segmento, na extensão *TextGrid*. Durante a localização das unidades VV, no entanto, devido a ruídos do áudio, comentários do interlocutor etc., o *script* pode apresentar erros, por exemplo, a não marcação do começo ou término de uma unidade, ou pode ocorrer também a marcação equivocada. Estes erros, no caso desta pesquisa, foram corrigidos manualmente.

6. CONSIDERAÇÕES FINAIS

Recentemente tem havido um crescente interesse em investigar elementos prosódicos (pausa, duração, tons de fronteira, por exemplo) que são utilizados para delimitar macroestruturas discursivas em narrativas orais de diferentes gêneros. Esses elementos prosódicos contribuem para a identificação semântica das unidades discursivas, assim como explicitam quais intenções do falante para com sua audiência, desse modo, o processo de comunicação é facilitado.

Estudos que envolvem aspectos prosódicos da fala devem considerar não apenas a parte da produção, mas sobretudo a percepção de tal fenômeno. Neste capítulo, apresentamos um guia para pesquisas de percepção da estrutura do discurso narrativo, assim como também para fazer a segmentação de narrativas em unidades VV. Utilizamos a metodologia aqui descrita para coletar dados para nossa pesquisa de mestrado sobre a relação entre produção e percepção de pistas prosódicas na segmentação de narrativas espontâneas apresentada ao Programa de Pós-Graduação em Linguística e Literatura da Universidade Federal de Alagoas, em 2017.

REFERÊNCIAS

- BARBOSA, P. A. *Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação*. Revista Estudos Linguísticos, Belo Horizonte, v.20, n.1, p.11- 27, jan./jun. 2012.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas, SP: Pontes ed.: FAPESP. 2006.
- BRUGMAN, H.; RUSSEL, A. *Annotating Multi-media / Multi-modal resources with ELAN*. In: *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC 2004)*. Paris: European Language Resources Association. 2004. pp. 2065-2068.
- LABOV, W. *The transformation of experience in narrative syntax*. In: LABOV, W. (ed.) *Language in the inner City*. Philadelphia: University of Pennsylvania Press, 1972. p. 219-247.
- LUCENTE, Luciana. *Aspectos dinâmicos de fala e da entonação no português brasileiro*. 2012. 204f. Tese (Doutorado em Letras). Universidade Estadual de Campinas, Campinas, 2012.
- OLIVEIRA JR., M. *Prosodic Features in Spontaneous Narratives*. 2000. 291 fl. Tese (Doutorado em Linguística) – Simon Fraser University, Vancouver, 2000.
- OLIVEIRA JR., M. *O uso do tom de limite como marca de segmentação da narrativa espontânea*. In *Revista Gelne*, vol.4, N 1, 2002. Disponível em: http://www.gelne.ufc.br/revista_ano4_no1_13.pdf. Acesso em: 25 jul. 2013.
- OLIVEIRA JR., M. *Percepção dos elementos prosódicos na narrativa espontânea. Projeto de pesquisa*. Faculdade de Letras, Universidade Federal de Alagoas, 2010.
- OLIVEIRA JR., M.; CRUZ, R.; SILVA, E. W. *A relação entre a prosódia e a estrutura de narrativas espontâneas: um estudo perceptual*. Revista Diadorim, Rio de Janeiro, vol. 12, p. 39-50, dez. 2012.
- PASSONNEAU, R. J.; LITMAN, D. J. *Intention-based segmentation: Human reliability and correlation with linguistic cues*. In: *Proceedings of the 31st annual meeting of the association for computational linguistics (acl-93)*, 1993, Columbus, Ohio. Columbus, Ohio, 1993. p. 148-155.

POMPINO-MARSCHALL, B. *The syllable as a prosodic unit and the so-called P-centre effect*. Munique, Alemanha, 1991.

SILVA, E. W.; OLIVEIRA JR., M. *A percepção dos elementos prosódicos como marca de estruturação de narrativas espontâneas*. In: *RESUMOS DO III COLÓQUIO BRASILEIRO DE PROSÓDIA DA FALA*, 2011, Belo Horizonte. Belo Horizonte, 2011. p. 15-17.

SILVA, E. W. *A relação entre produção e percepção de pistas prosódicas na segmentação de narrativas espontâneas*. 2017. 79 f. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Programa de Pós-Graduação em Letras e Linguística, Universidade Federal de Alagoas, Maceió, 2017.

WOLFSON, N. *Speech events and natural speech*. *Language in Society*, Cambridge, n.5, p.189-209, 1979.