UFRJ
PROGRAMA DE PÓS-GRADUAÇÃO
EM LETRAS VERNACULAS

DIADORIM

Revista Diadorim
e-ISSN: 2675-1216
v.25, n.1, e61715, 2023
DOI: 10.35520/diadorim.2023.v25n1a61715

Artigo Original

O uso da prosódia na percepção de fronteiras de eventos narrados em português brasileiro

The use of prosody in the perception of event boundaries narrated in Brazilian Portuguese

Ebson Wilkerson Rocha Silva[1]
Arthur Ronald Brasil Terto[2]
Miguel Oliveira Jr [3]

[1]Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada, Serra Talhada, PE, Brasil.
[2] Universidade Federal de Alagoas, Faculdade de Letras, Programa de Pós-Graduação em Linguística e Literatura. Maceió, AL, Brasil.
[3] Universidade Federal de Alagoas, Faculdade de Letras. Maceió, AL, Brasil.
E-mail: ebson.wilkerson@ufrpe.br
E-mail: arthur.ppgll@gmail.com
E-mail: miguel@fale.ufal.br

Editor-chefe
Marcia dos Santos Machado Vieira

Autor correspondente
Ebson Wilkerson Rocha Silva
ebson.wilkerson@ufrpe.br

Recebido: 27/10/2023
Aceito: 08/05/2024

Como citar:
SILVA, Ebson Wilkerson Rocha; TERTO, Arthur Ronald Brasil; OLIVEIRA JR, Miguel. O uso da Prosódia na percepção de fronteiras de eventos narrados em português brasileiro. *Revista Diadorim*, v.25, n.1, e61715, 2023. doi: https://doi.org/10.35520/diadorim.2023.v25n1a61715

Resumo

A Teoria da Cognição de Eventos define eventos como ações que têm um objetivo; que são situadas no tempo e no espaço; que envolvem pessoas e objetos; e que podem representadas em narrativas (Zacks; Tversky, 2001; Zacks *et al.*, 2007; Zacks, 2020). Paralelamente, estudos têm demonstrado que a prosódia marca as relações semântico-pragmáticas entre as subpartes que compõem o texto falado (Grosz; Hirschberg, 1992; Oliveira Jr., 2000; Oliveira Jr.; Cruz; Silva, 2012; Swerts, 1996; Swerts; Geluykens, 1994; Terto; Oliveira Jr., 2021). Visto que eventos podem ser representados na fala; e que a prosódia é o elemento que organiza o texto falado, objetivamos analisar o papel da prosódia na segmentação da estrutura de eventos nessa modalidade, examinando as características acústicas de

This content is licensed under a Creative Commons Attribution 4.0 International License.

fronteiras prosódicas percebidas como fronteiras de eventos em um experimento de segmentação de narrativas espontâneas. Hipotetizamos que haveria características específicas de diferença de tom, de intensidade, de duração e de pausas silenciosas nas fronteiras prosódicas percebidas como fronteiras de eventos. Selecionamos trinta narrativas orais espontâneas de um *corpus* previamente construído. Participaram do estudo 26 participantes falantes nativos do português brasileiro. Nas análises acústicas das narrativas, comparamos medidas de F0, de duração e de intensidade entre as sílabas tônicas adjacentes às fronteiras prosódicas, bem como a ocorrência de pausas silenciosas. Os resultados sugerem que as fronteiras prosódicas percebidas como fronteiras de eventos apresentam maior diferença de tom, valores de intensidade mais baixos na sílaba tônica imediatamente anterior à fronteira prosódica e presença significativa de pausas silenciosas.

Palavras-chave

Prosódia; Teoria de eventos; Percepção; Segmentação; Narrativas orais.

Abstract

The Theory of Event Cognition defines events as actions that have a goal; that are situated in time and space; that involve people and objects; and can be represented in narratives (Zacks; Tversky, 2001; Zacks *et al.*, 2007; Zacks, 2020). Meanwhile, studies have shown that prosody marks the semantic-pragmatic relations between the subparts that make up spoken text (Grosz; Hirschberg, 1992; Oliveira Jr., 2000; Oliveira Jr.; Cruz; Silva, 2012; Swerts, 1996; Swerts; Geluykens, 1994; Terto; Oliveira Jr., 2021). Given that events can be represented in speech; and that prosody is the element that organizes spoken text, we aim to analyze the role of prosody in the segmentation of event structure in this modality, examining the acoustic characteristics of prosodic boundaries perceived as event boundaries in a spontaneous narrative segmentation experiment. We hypothesized that there would be specific differences in pitch, intensity, duration, and silent pauses at the prosodic boundaries perceived as event boundaries. We selected thirty spontaneous oral narratives from a previously constructed *corpus*. The study involved 26 native speakers of Brazilian Portuguese. In the acoustic analyses of the narratives, we compared measures of F0, duration, and intensity between the stressed syllables adjacent to the prosodic boundaries, as well as the occurrences of silent pauses. The results suggest that the prosodic boundaries perceived as event boundaries exhibit a greater difference in pitch, lower intensity values in the stressed syllable immediately preceding the prosodic boundary, and a significant presence of silent pauses.

Keywords

Prosody; Event theory; Perception; Segmentation; Oral narratives.

Introdução

Ao longo do dia, realizamos diversas ações: acordamos, tomamos café da manhã, vamos ao trabalho, almoçamos, saímos do trabalho, chegamos em casa, jantamos e, por fim, dormimos. Essa seria a narração prototípica de um dia comum no cotidiano da maioria das pessoas. Percebemos as ações da vida cotidiana como uma série de "pedaços" que podem ser identificados no tempo, os quais chamamos de eventos (Zacks *et al.*, 2007). Eventos são ações (i) que têm um propósito; (ii) que estão situadas no tempo, ou seja, têm um começo e um fim; (iii) que ocorrem em um espaço específico; e (iv) que envolvem pessoas e objetos (Zacks; Tversky, 2001; Radvansky; Zacks, 2014).

Um número considerável de estudos tem investigado a relação entre linguagem verbal e processamento de eventos. Nesse contexto, existe uma linha de pesquisa dedicada a compreender como funciona a representação de eventos na manifestação linguística de modalidade oral. Sabendo que eventos podem ser representados na linguagem (Zacks, 2020), surge a seguinte questão: quais pistas prosódicas seriam utilizadas pelo ouvinte para criar modelos de eventos a partir de uma narrativa oral?

A literatura tem constatado que a prosódia é um fenômeno linguístico que serve para organizar as unidades discursivas dentro do texto falado[1] (De Pijper; Sanderman, 1994; Swerts; Geluykens, 1994; Swerts, 1996; Oliveira Jr., 2000; Barbosa, 2012; Oliveira Jr.; Cruz; Silva, 2012; Silva, 2017; Terto; Oliveira Jr., 2021, entre outros). Esse papel da prosódia não é constatado apenas do ponto de vista da produção da fala, mas também do ponto de vista da percepção. Os falantes/ouvintes usam as informações prosódicas para identificar e processar a estrutura do texto falado. Portanto, hipotetiza-se que a prosódia desempenha um papel relevante na percepção e no processamento da estrutura de eventos no texto oral, mais especificamente em narrativas orais. É necessário, então, investigar quais elementos prosódicos são mais salientes para o ouvinte na tarefa de percepção das fronteiras de eventos, simbolizadas, em termos estruturais, pelas fronteiras prosódicas.

Tendo em vista esse problema, o objetivo deste estudo foi investigar o papel dos elementos prosódicos no processamento da estrutura de eventos em narrativas orais espontâneas. Esse objetivo se desdobrou em outros mais específicos, a saber: (i) coletar e segmentar dados de narrativas orais espontâneas; (ii) desenvolver um experimento comportamental *on-line* em um *software* gratuito para a segmentação de narrativas

[1] Neste estudo, intercambiamos entre as expressões "discurso falado" e "texto falado" para nos referirmos à materialidade linguístico-estrutural que está acima do nível das sentenças. O "discurso", seria, nessa concepção, um dos níveis da gramática da língua: ele estaria acima do nível das sentenças. Assim, a unidade imediatamente inferior ao discurso seriam as sentenças, que, conectadas entre si por relações semântico-pragmáticas, formariam o todo maior, ou seja, o discurso. Para conhecer outras concepções de discurso, consulte Wichmann (2014).

gravadas em uma variedade do português brasileiro; e (iii) descrever as características acústicas dos elementos prosódicos que atuam em fronteiras prosódicas associadas pelos ouvintes a fronteiras de eventos dentro do texto narrativo.

A prosódia e a segmentação do discurso

A forma como o texto escrito é organizado costuma ser percebida porque convenções ortográficas, como os sinais de pontuação, a paragrafação e o espaçamento, facilitam o reconhecimento da estruturação desse texto. No entanto, essas pistas não são úteis para a apreensão da estruturação do discurso falado. Na cadeia de produção oral, por exemplo, a percepção do limite entre palavras pode ser mais difícil, como no caso das palavras "afim" e "a fim", que não apresentam diferenças oralmente, mas, na escrita, sim: o espaço em branco torna-se útil para esclarecer que, no segundo exemplo, temos duas palavras e não apenas uma, como no primeiro exemplo. Portanto, o discurso falado precisa utilizar outros mecanismos para fazer evidente a forma como está estruturado.

A organização do discurso falado em segmentos menores tem sido analisada em uma ampla gama de estudos. Grande parte deles aponta a prosódia como um dos principais elementos delimitadores de estruturas discursivas na oralidade (Grosz; Hirschberg, 1992; Swerts; Geluykens, 1994.; Passonneau; Litman, 1997). Nesse sentido, as unidades discursivas da fala podem ser reconhecidas a partir de certos elementos linguísticos, pois, segundo Grosz e Sidner (1986), há uma interação bidirecional entre a estrutura do discurso e os enunciados que o compõem. Isso significa que expressões linguísticas podem ser usadas para transmitir informações sobre a própria estrutura discursiva e, por outro lado, essa estrutura restringe a interpretação do próprio discurso, afetando, assim, o que o falante diz e como o ouvinte percebe o que é dito.

Oliveira Jr. (2000) mostrou que certas variáveis prosódicas exercem papel crucial na estruturação discursiva, dividindo o texto narrativo em seções semanticamente independentes. Por exemplo, encontram-se, com mais frequência, a ocorrência e a duração de pausas, a diferença de tom e os tons de fronteira marcando o final de uma seção semanticamente independente. O tom, aparentemente, não produz esse efeito se for observado isoladamente (Pereira, 2014), mas pode ser mais comumente encontrado em regiões de fronteira discursiva com características de tom mais baixo, ao invés de não baixos (Oliveira Jr., 2002).

Esses estudos mostram que, do ponto de vista da produção, quanto menos variáveis usadas nos limites do discurso, menores são as chances de tais fronteiras coincidirem com uma fronteira de seção narrativa. Quanto à função relacionada aos conteúdos informacionais, tem-se observado que a pausa, a velocidade da fala e a variação de tom são frequentemente empregadas em seções narrativas com função avaliativa (Oliveira Jr., 2002).

É possível afirmar, portanto, que, em estudos de produção, as fronteiras entre unidades maiores das narrativas (ou fronteiras narrativas) são prosodicamente diferentes. Por exemplo, pausas que ocorrem mais sistematicamente no final das seções narrativas são frequentemente mais longas que em outras posições. Outros elementos prosódicos já investigados foram alongamento final de palavra em posição de pré-fronteira frasal (Turk; Shattuck-Hufnagel, 2007), maior diferença de tom demarcando fim/início de unidades frasais (De Pijper; Sanderman, 1994; Swerts, 1996; Lin; Fon, 2011; Truckenbrodt; Féry, 2015), alongamento da sílaba tônica anterior à sílaba de fronteiras frasais (Byrd; Riggs, 2008; Katsika *et al.*, 2014), e em fronteiras de palavras (Kim, 2020).

Estudos envolvendo aspectos prosódicos da fala, no entanto, como os citados acima, devem considerar não apenas o componente de produção, mas, sobretudo, a percepção desse fenômeno. A relevância das variáveis prosódicas na demarcação das estruturas do discurso só pode ser plenamente verificada após considerar sua validade do ponto de vista perceptivo.

Oliveira Jr., Cruz e Silva (2012) estenderam a análise apresentada em Oliveira Jr. (2000), examinando até que ponto os ouvintes podem se beneficiar dos marcadores prosódicos para inferir a estrutura de narrativas orais espontâneas. Para tanto, foi desenvolvido um protocolo experimental no qual 48 informantes receberam narrativas em diferentes condições (somente texto, texto e áudio, somente áudio e áudio filtrado). Os resultados mostraram que as pessoas identificaram consistentemente a estrutura discursiva dos estímulos, mesmo sem acesso a informações lexicais, sintáticas e semânticas, sugerindo que a prosódia desempenha um papel importante na percepção da estrutura discursiva de narrativas orais espontâneas.

A teoria da segmentação de eventos

A maneira como percebemos e compreendemos o mundo ao nosso redor é influenciada por processos cognitivos que são moldados pelas experiências que vivenciamos. Essas experiências geralmente são vistas como eventos que ocorrem em lugares específicos, durante períodos de tempo definidos, envolvendo pessoas e objetos particulares, sendo impulsionadas por motivos que lhes conferem significado.

A Teoria de Segmentação de Eventos (Radvansky; Zacks, 2014; Zacks *et al.*, 2007; Zacks; Tversky, 2001) descreve e explica os processos envolvidos nas representações de eventos. Segundo essa teoria, existem cinco dimensões essenciais nessas representações: tempo, espaço, entidade (que pode ser uma pessoa ou um objeto), causalidade e intencionalidade (que envolve objetivos específicos) (Zwaan, 2008). Quando ocorrem mudanças em uma ou mais dessas dimensões, nós precisamos reconstruir ou atualizar o modelo situacional criado anteriormente, uma vez que

ele não consegue mais acomodar as informações previamente integradas (Speer; Zacks; Reynolds, 2007). Isso geralmente resulta na percepção do encerramento de um evento e o início de outro, isto é, uma fronteira (Zacks *et al.*, 2007).

Considerando que os eventos podem ser narrados oralmente por meio da língua e de suas estruturas, é razoável pressupor que os indivíduos recorrem à prosódia para segmentar o discurso narrativo em unidades menores, pois a prosódia é um mecanismo inerente à fala natural humana (Barbosa, 2012). Portanto, pressupomos que os elementos prosódicos de diferença de tom, intensidade, duração e ocorrência de pausas silenciosas são utilizados pelos ouvintes como indícios linguísticos que sinalizam a estrutura de eventos no texto narrativo falado.

Métodos

Estímulos experimentais

Neste estudo utilizamos como estímulos experimentais um conjunto de narrativas retiradas de um *corpus* de entrevistas semiespontâneas gravadas ao longo do segundo semestre de 2019, em que pares de informantes foram solicitados a conversar entre si, a partir da sugestão de vídeos, de imagens e de tópicos espontâneos ou pré-definidos. Para a elicitação das narrativas, foram utilizados como estímulos dois quadrinhos retirados do acervo *McNaught Syndicate*, *Inc.*, *N.Y.* e duas animações[2]. Participaram das entrevistas 24 graduandos[3] em Letras-Português, sendo seis homens. Os entrevistados tinham idade média de 22 anos e todos ainda estavam cursando o ensino superior.

Retiramos desse *corpus* de entrevistas semiespontâneas 10 (dez) narrativas. Os critérios utilizados para a seleção dos estímulos foram estes: (i) inexistência de sobreposição de fala; (ii) as narrativas espontâneas deveriam apresentar mudanças situacionais, isto é, mudanças em uma das seguintes dimensões: tempo, espaço, objeto, personagem, causalidade e/ou objetivo; (iii) qualidade acústica; e (iv) duração máxima de 3 (três) minutos e 30 (trinta) segundos.

Depois de coletadas, essas narrativas foram enviadas para 2 pesquisadores experientes em segmentação e revisão de dados de fala gravados. O objetivo deste procedimento

[2] As animações são disponibilizadas pelos estúdios *The CG Bros* no Youtube e podem ser acessadas pelos links: Lovebites – https://youtu.be/katIJ9Oabb8?si=HLThe2lKDx3Q0JPD
Playmate – https://youtu.be/2fR9HqilmKM?si=4WRPyR6NKWwZnstx

[3] Todos os participantes deste estudo assinaram um Termo de Consentimento Livre e Esclarecido. O estudo foi avaliado e aprovado para realização em 01/05/2018 com parecer de número 2.408.979 e CAAE 79431617.3.0000.5013.

foi evitar a circularidade[4] na análise dos dados prosódicos. Esses pesquisadores colaboradores tiveram como tarefa (i) segmentar as narrativas em unidades de sentido semanticamente independentes; e (ii) revisar todo o processo de segmentação. Cada narrativa coletada tinha em média cerca de 2 (dois) minutos.

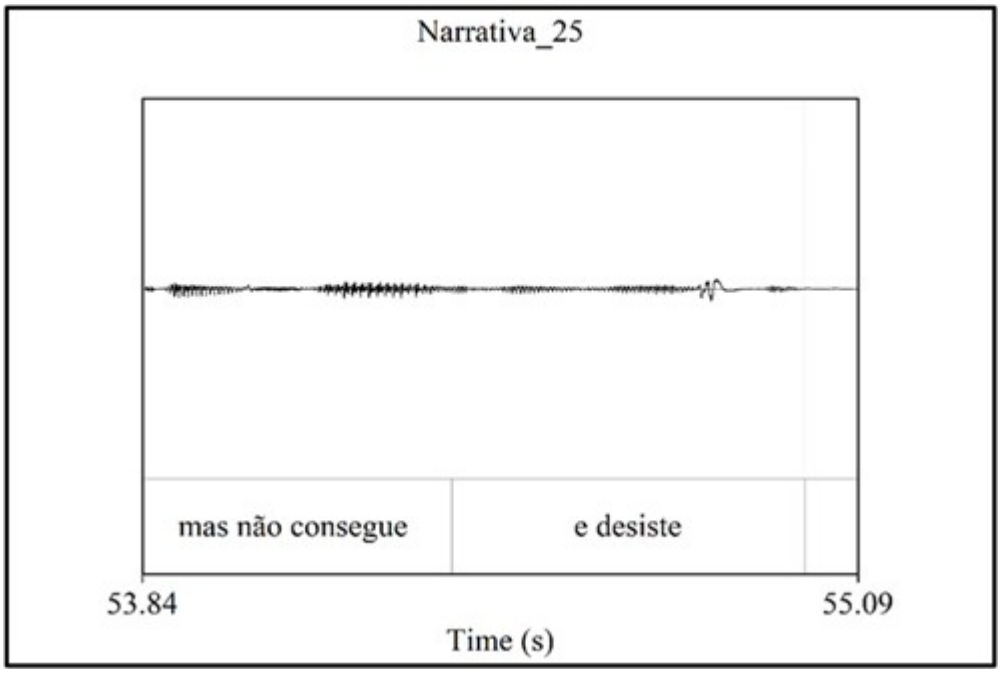

Figura 1. Exemplo de segmentação feita no Praat.
Fonte: Autores (2023).

O processo de segmentação e revisão das narrativas se deu no ambiente do *software Praat*. A Figura 1 acima é um exemplo de segmentação feita no *Praat*, na primeira linha é apresentado o oscilograma do áudio das gravações; e, logo abaixo, um exemplo de transcrição ortográfica da narrativa com as segmentações feitas por especialista. Tomamos a segmentação independente como base para analisar, *a posteriori*, a percepção de fronteiras por parte dos sujeitos da pesquisa no experimento comportamental, descrito a seguir.

O *software* utilizado no experimento

Para aplicar o experimento comportamental, utilizamos a plataforma de experimentos *on-line Gorilla Experiment Builder*. Trata-se de um *software* livre desenvolvido para criar tarefas e experimentos científicos na *internet*. Os principais benefícios em utilizar tal *software* são: (i) sua interface altamente intuitiva; (ii) o fato de não exigir

[4] A circularidade ocorre quando o pesquisador de um estudo de segmentação usa seus próprios conhecimentos sobre prosódia para identificar as fronteiras prosódicas presentes no *corpus* que irá analisar (Swerts; Geluykens, 1994). Para evitar tal risco, é necessário que todas as segmentações sejam feitas de modo independente.

conhecimentos prévios em programação; (iii) a coleta de dados se dar com alta precisão do tempo de reação dos participantes; (iv) os experimentos serem construídos gratuitamente[5], dentre outros.

No tópico a seguir, falaremos mais sobre a utilização de plataformas *on-line* de experimentos em pesquisas.

Participantes, experimento e métodos de análise

O experimento foi encaminhado por meio de um link para os participantes. Após concluírem a tarefa experimental, o sistema do *software Gorilla Experiment Builder*, automaticamente, computava suas respostas e seus tempos de reação aos estímulos sonoros que haviam ouvido, as narrativas. Participaram do experimento 27 estudantes do curso de Letras da Universidade Federal de Alagoas (UFAL), com idades entre 18 e 30 anos, recrutados a partir de convites feitos em duas turmas do curso de Letras-Português. Os critérios para a seleção dos participantes foram (i) ser falante nativo do português brasileiro; (ii) não ter problemas auditivos; (iii) autodeclarar a inexistência de distúrbios neurológicos; (iv) não participar de nem conduzir pesquisas em fonética, mais especificamente em estudos sobre prosódia; e (v) ser estudante de graduação ou já ter concluído o ensino superior.

Para acessar o experimento e realizar a tarefa experimental, eles receberam um link por e-mail com uma mensagem de boas-vindas e com o Termo de Consentimento Livre e Esclarecido. Após a assinatura do Termo, acessaram uma tela de instruções contendo um exemplo de segmentação de narrativa acompanhada de áudio. Em seguida, iniciaram o experimento. Todos foram instruídos a utilizar fones de ouvido para realizar a tarefa. A tarefa experimental consistia em apertar a tecla espaço no teclado do computador a cada vez que julgassem ter percebido uma fronteira de eventos. Ao solicitarmos essa tarefa, porém, não explicamos o conceito de *evento*. O principal objetivo era verificar se os participantes concordavam entre si nas segmentações feitas. Algumas narrativas foram acompanhadas de perguntas com vistas a garantir a atenção ao conteúdo narrado.

Depois que os dados foram coletados, a taxa de concordância entre os participantes foi computada. Para tanto, foram estabelecidas janelas temporais ao longo das fronteiras prosódicas verificadas nas narrativas. Para cada fronteira, criamos uma janela cujos limites eram de 300 ms antes do início da fronteira e 300 ms após o fim da fronteira prosódica. A adoção desse ponto de corte deu-se por haver fortes evidências de que pausas inferiores a 250 ms não refletem necessariamente lacunas

[5] Registre-se, contudo, que, embora a construção das tarefas que compõem o experimento seja gratuita, o pesquisador deve pagar por cada participante que constituirá sua amostra de pesquisa.

na fala (Oliveira Jr., 2000). Toda segmentação feita dentro dessa janela temporal foi computada para fins do cálculo de concordância entre os sujeitos.

Em seguida, estabelecemos as fronteiras prosódicas proeminentes de cada narrativa a partir dos resultados do teste de segmentação. Sullivan (2014), ao avaliar diferentes métodos de análise de concordância, concluiu que uma taxa de 40% entre os participantes é um limiar adequado para julgar como significativo um dado nível de concordância. Assim, adotando esse limiar, julgamos como proeminentes as fronteiras em que houvesse ao menos 40% de concordância, significando que foram percebidas como fronteiras de eventos dentro do texto narrativo; já aquelas fronteiras prosódicas cujo limiar de concordância foi abaixo de 40% foram categorizadas como sendo não proeminentes, isto é, como fronteiras prosódicas comuns.

Após a categorização das fronteiras prosódicas a partir do limiar de concordância, iniciamos a análise acústica dos dados. Além da anotação de segmentação realizada independentemente, outras quatro anotações foram feitas para cada narrativa: (i) Janelas (anotação usada para o cálculo da taxa de concordância entre os sujeitos no que se refere à tarefa de segmentação das narrativas); (ii) VVsNP (anotação utilizada para indicar as sílabas fonéticas em fronteiras não proeminentes); (iii) VVsP (anotação utilizada para indicar as sílabas fonéticas das fronteiras proeminentes); e (iv) Silence (anotação usada para indicar a presença ou ausência de pausas silenciosas).

O script AnalyseTier foi usado durante a análise acústica. Trata-se de um script desenvolvido por Hirst e Di Cristo (1998) para coletar informações dos correlatos acústicos dos enunciados. Essa coleta tem como ponto de partida as segmentações e anotações feitas nas camadas do *Praat*. As medidas discretas de frequência fundamental (diferença de tom), de amplitude (intensidade) e de duração (duração das sílabas tônicas) foram coletadas de forma automática. O script coletou os valores mínimo, médio e máximo de F0 e de amplitude para cada unidade segmentacional feita no *Praat*. No entanto, levamos em consideração apenas o pico de F0 e de intensidade para cada sílaba.

A análise de diferença de tom e de diferença de intensidade foi realizada a partir da comparação entre a F0 máxima e a intensidade máxima, respectivamente, das sílabas tônicas adjacentes às fronteiras prosódicas. A análise de duração, por sua vez, foi baseada na comparação da duração das sílabas tônicas imediatamente anteriores aos dois tipos de fronteiras prosódicas (proeminentes e não proeminentes). Utilizamos o conceito de sílabas fonéticas para essa análise de duração (Barbosa, 2019). Por fim, o parâmetro escolhido para a marcação de pausas foi o período de silêncio igual ou superior a 150 ms (Oliveira Jr., 2000). Foram escolhidas 2 etiquetas de anotação em relação às pausas: Yes (quando havia pausa); No (quando não havia pausa).

Após as anotações dos trechos e o armazenamento dos dados coletados pelos scripts em planilhas eletrônicas, as análises estatísticas foram feitas no *software* R. O tratamento dos dados no R foi conduzido por meio de alguns pacotes operacionais, dentre os principais estes: *readr, readxl, dplyr, ggplot2, lme4* e *lmerTest*.

Os dois testes utilizados em nossas análises foram o modelo linear misto e a regressão logística binomial.

O teste estatístico de modelo linear misto (LLM) é usado na análise de dados para investigar a relação entre variáveis independentes e dependentes, levando em consideração a estrutura hierárquica dos dados. Dizer que os dados têm uma estrutura hierárquica significa dizer que foram obtidos a partir de observações repetidas dentro um grupo de sujeitos ou unidades amostrais (por exemplo, quando coletamos diferentes amostras de segmentação de narrativas feitas pelo mesmo participante). Já a regressão logística binomial é uma técnica estatística usada para modelar a relação entre uma variável dependente binária e uma ou mais variáveis independentes. Trata-se de um teste utilizado a fim de compreender como variáveis independentes influenciam a probabilidade de um resultado binário (por exemplo, quando testamos se a presença de uma fronteira de eventos influenciava a ocorrência de pausas silenciosas = sim ou não).

Resultados

Diferença de tom entre unidades prosódicas

Postulamos que aquelas fronteiras prosódicas percebidas pelos ouvintes como sendo correspondentes a fronteiras de eventos dentro das narrativas teriam uma maior diferença tonal entre o valor de F0 verificado na última sílaba tônica da unidade prosódica finalizante e o valor de F0 verificado na primeira sílaba tônica da unidade prosódica seguinte. Esse postulado decorre do fato de estudos prévios (De Pijper; Sanderman, 1994; Oliveira Jr., 2000, entre outros) terem constatado que é comum haver esse padrão tonal quando fronteiras prosódicas coincidem com a percepção de mudança de tópico discursivo no texto narrado, resultando, assim, em uma maior diferença de tom entre unidades prosódicas.

Para verificar o postulado, ajustamos um modelo linear misto com *semitons (ref. 100 Hz)* como variável dependente e *posição de sílaba* (níveis: última sílaba, primeira sílaba), *proeminência* (níveis: não proeminente, proeminente) e a interação entre essas duas últimas variáveis como efeitos fixos. O modelo também continha interceptos aleatórios por janelas[6]. O Gráfico 1 abaixo mostra a diferença de tom verificada nos dois tipos de fronteiras prosódicas (NP: não proeminentes [não percebidas como fronteira de eventos]; e P: proeminentes [percebidas como fronteira de eventos]).

[6] Veja a seção de métodos para entender a noção de janelas temporais.

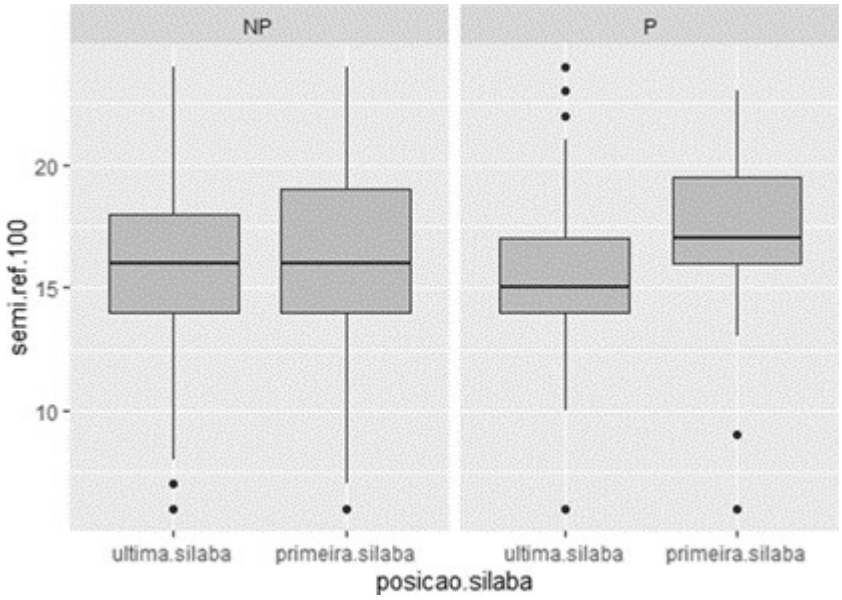

Gráfico 1. Diferença de F0 entre sílabas tônicas adjacentes às fronteiras prosódicas em ambos os contextos de proeminência.
Fonte: Autores (2023).

Ao aplicarmos um teste ANOVA[7] para comparação por modelos aninhados, constatamos que o modelo com interação entre as variáveis *posição de sílaba* e *proeminência* era significativo para explicar os dados (p = 0.02). Um teste post-hoc mostrou que não houve diferença de F0 significativa entre as sílabas tônicas quando a fronteira era não proeminente (não percebida como fronteira de eventos) (p = 0.39). Porém, quando a fronteira era proeminente (percebida como correspondente a uma fronteira de eventos), a diferença de F0 entre as sílabas era significativa (p = 0.03). Portanto, esses resultados indicam a confirmação do postulado mencionado acima: a diferença de F0 entre as unidades prosódicas adjacentes, em um contexto de fronteira prosódica percebida como fronteira de eventos, é um elemento prosódico relevante para a percepção da estrutura de eventos narrados oralmente.

Diferença de intensidade entre unidades prosódicas

Postulamos que aquelas fronteiras prosódicas percebidas pelos ouvintes como sendo correspondentes a fronteiras de eventos teriam uma diferença significativa de intensidade entre a última sílaba tônica da unidade prosódica finalizante e a primeira sílaba tônica da unidade prosódica seguinte. Isso porque a queda no valor de intensidade, no fim de uma unidade prosódica, e seu reinício, na unidade prosódica seguinte, também é um elemento prosódico relevante para a indexicalização

[7] O teste ANOVA (Análise de Variância) é uma técnica estatística usada para comparar as médias de dois ou mais grupos independentes para determinar se há diferenças estatisticamente significativas entre eles. Ele testa se as médias dos grupos são iguais ou se pelo menos um grupo é diferente dos outros.

de segmentações no discurso falado (Oliveira Jr., 2000; Almeida, 2017; Silva, 2017; Barbosa, 2019; Terto; Oliveira Jr., 2021).

Tendo esse postulado em vista, ajustamos um modelo linear misto com *intensidade (dB)* como variável dependente e *posição de sílaba* (níveis: última sílaba, primeira sílaba), *proeminência* (níveis: não proeminente, proeminente) e a interação entre essas duas últimas variáveis como efeitos fixos. O modelo também continha interceptos aleatórios por janelas. Os contrastes entre os níveis das variáveis foram ajustados para o tipo *dummy coding*. O Gráfico 2 abaixo mostra a diferença de intensidade verificada nos dois tipos de fronteiras prosódicas (NP: não proeminentes [não percebidas como fronteira de eventos]; e P: proeminentes [percebidas como fronteira de eventos]).

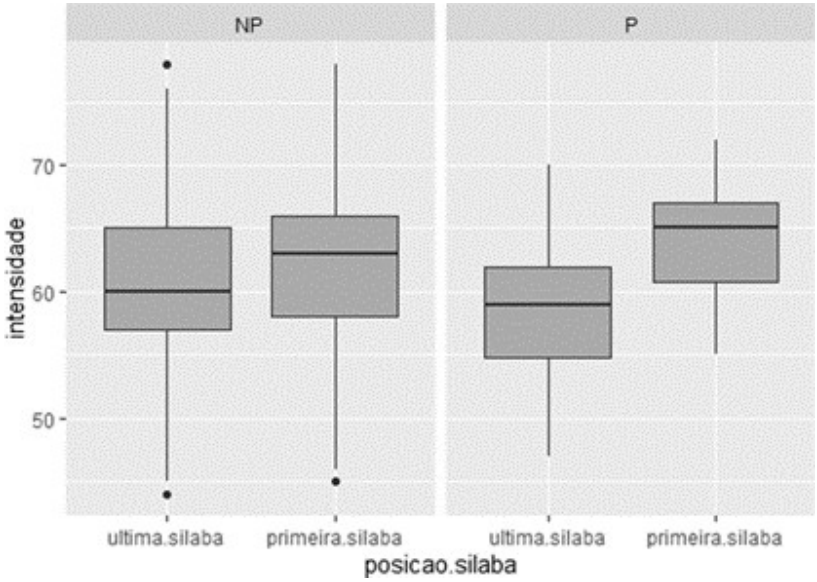

Gráfico 2. Diferença de intensidade entre sílabas tónicas adjacentes às fronteiras prosódicas em ambos os contextos de proeminência.
Fonte: Autores (2023).

Ao aplicarmos um teste ANOVA para fazer uma comparação por modelos aninhados, verificamos que o modelo com interação entre *posição de sílaba* e *proeminência* era significativo para explicar os dados de intensidade ($p < 0.001$). Um teste post-hoc mostrou que houve diferença significativa de intensidade entre as sílabas tônicas adjacentes às fronteiras prosódicas quando estas eram percebidas como proeminentes (percebidas como sendo também uma fronteira de eventos) ($p < 0.001$). No entanto, observamos os mesmos resultados para a condição de fronteiras não proeminentes (isto é, percebidas como não sendo uma fronteira de eventos) ($p <0.001$). Esses resultados parecem sugerir que, independentemente de uma fronteira prosódica estar associada a uma fronteira de eventos, é esperado que a unidade prosódica finalizante tenha registros de intensidade menores que aqueles observados no início de uma nova unidade. É razoável notar, entretanto, a diferença visivelmente maior na condição de fronteira proeminente.

Duração da última sílaba tônica anterior aos dois tipos de fronteiras prosódicas

Postulamos que as sílabas tônicas finais anteriores a fronteiras prosódicas percebidas como fronteiras de eventos teriam uma duração maior que as sílabas tônicas finais anteriores a fronteiras prosódicas não proeminentes (ou seja, percebidas como não sendo associadas a uma fronteira de eventos). Esse postulado partiu da premissa de que as fronteiras prosódicas corresp*onde*ntes a uma fronteira de eventos estariam finalizando unidades discursivas, uma vez que estudos prévios demonstraram que, quanto maior for o nível linguístico da segmentação (por exemplo, uma segmentação entre unidades discursivas e não entre sentenças), maior a duração da sílaba tônica precedente à fronteira prosódica (Swerts; Geluykens, 1994; Swerts, 1996).

Para testar esse postulado, ajustamos um modelo linear com *intensidade_US* (intensidade verificada na última sílaba) como variável dependente e *proeminência* (níveis: não proeminente, proeminente) como efeito fixo. O Gráfico 3 abaixo mostra os valores de intensidade observados nas sílabas tônicas imediatamente anteriores às fronteiras prosódicas nos dois tipos de fronteira prosódica (NP: não proeminentes [não percebidas como fronteira de eventos]; e P: proeminentes [percebidas como fronteira de eventos]).

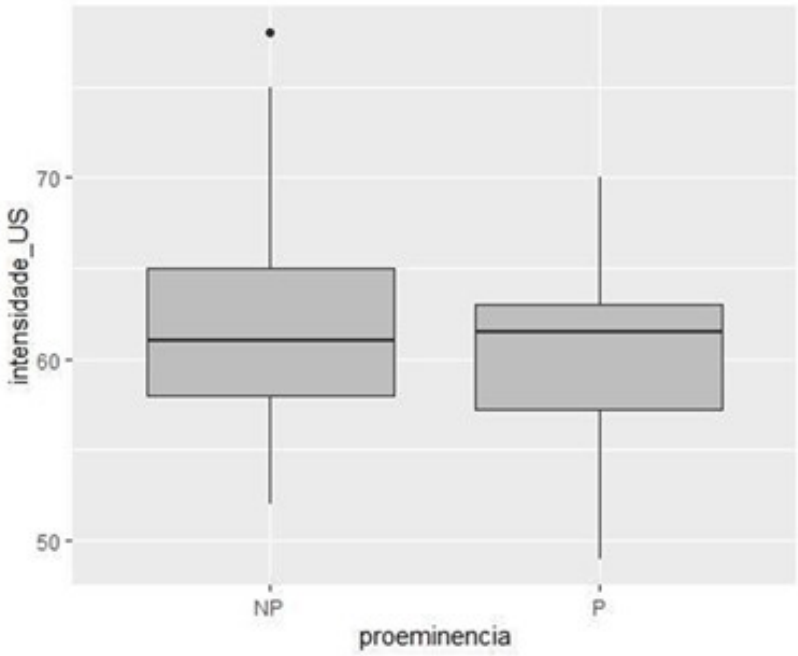

Gráfico 3. Valores de intensidade observados nas últimas sílabas tônicas em ambos os tipos de fronteira prosódica.
Fonte: Autores (2023).

O modelo ajustado mostrou que não houve diferença significativa entre os valores de duração das sílabas tônicas anteriores a fronteiras proeminentes e os valores de duração das tônicas anteriores a fronteiras não proeminentes (p = 0.61).

Ocorrência de pausas silenciosas nas fronteiras prosódicas

Por fim, postulamos que haveria uma maior probabilidade de ocorrência de pausas silenciosas nas fronteiras prosódicas percebidas como sendo também fronteiras de eventos dentro das narrativas. Esse postulado tem como fundamento achados de estudos anteriores que constataram o papel crucial das pausas silenciosas na organização da estrutura do texto falado, sendo este um dos elementos prosódicos mais atuantes na indexicalização de fronteiras discursivas (Swerts; Geluykens, 1994; Swerts, 1996; Oliveira Jr., 2000; Oliveira Jr.; Cruz; Silva, 2012; Silva, 2017; Terto; Oliveira, 2021).

Para testar esse postulado, ajustamos uma regressão logística binomial com *pausa* (níveis: não, significando não ocorrência; e sim, significando ocorrência) como variável dependente e *proeminência* (níveis: não proeminente, proeminente) como efeito fixo. O contraste dos níveis da variável dependente (não, sim) foi ajustado para contrastes do tipo *sum coding*, com "Nao" (que indica a não ocorrência de pausas) como nível de referência da variável dependente. O Gráfico 4 abaixo mostra a relação entre a ocorrência e a não ocorrência de pausas silenciosas nos dois tipos de fronteiras prosódicas (NP: não proeminentes [não percebidas como fronteira de eventos]; e P: proeminentes [percebidas como fronteira de eventos]).

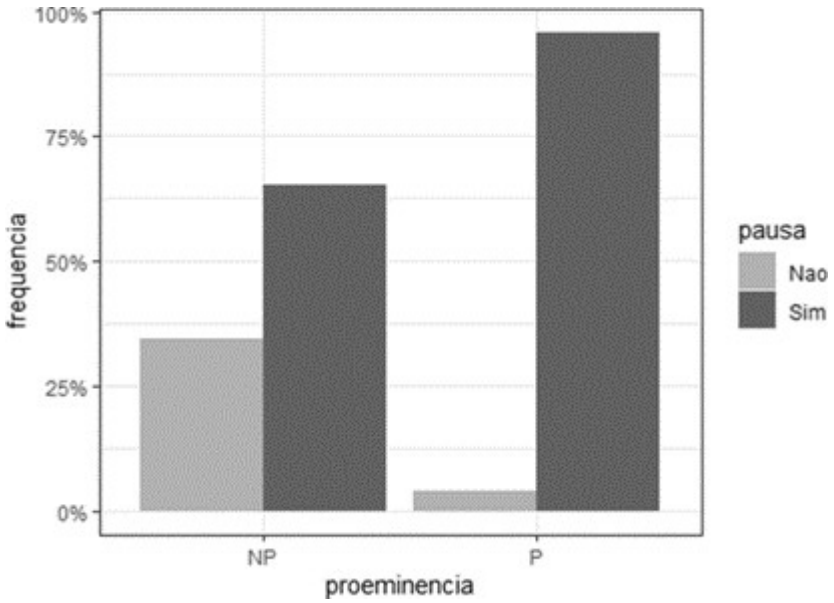

Gráfico 4. Relação entre ocorrência e não ocorrência de pausas silenciosas nos dois tipos de fronteiras prosódicas.
Fonte: Autores (2023).

A regressão mostrou que as fronteiras percebidas como proeminentes (associadas a fronteiras de eventos) têm chances muito significativas de conter uma pausa silenciosa ($p < 0.001$). No entanto, as chances de ocorrência de pausas silenciosas nas fronteiras não proeminentes (percebidas como não sendo fronteiras de eventos) também se mostraram significativas ($p < 0.001$). Esses resultados ratificam, assim, o papel crucial das pausas silenciosas na marcação de quaisquer fronteiras prosódicas, sejam elas percebidas como fronteiras de eventos/discursivas ou não. Entretanto, é razoável notar uma diferença visivelmente maior entre a ocorrência e a não ocorrência de pausas na condição de fronteira proeminente, com pouquíssimos casos em que a pausa não se faz presente.

Discussão geral

A prosódia desempenha um papel crucial na organização do discurso oral, sendo um dos principais meios pelos quais os ouvintes percebem as fronteiras de eventos e estruturas discursivas. Este artigo buscou investigar os elementos prosódicos (diferença de tom, intensidade, duração e pausas silenciosas) na percepção das fronteiras de eventos na fala. Ao analisar a interação complexa entre esses elementos, buscamos entender como eles influenciam a delimitação das unidades discursivas e a marcação de eventos na língua falada.

Observamos que a diferença de tom emergiu como um fator relevante para os ouvintes discernir as fronteiras de eventos nas narrativas estudadas. Essa diferença tonal foi particularmente notável nas sílabas tônicas adjacentes às fronteiras prosódicas, indicando uma diferença de tom mais significativa associada à percepção de eventos. Em relação à intensidade, mesmo que as diferenças não tenham atingido significância estatística, foi observado que as fronteiras prosódicas relacionadas a eventos tendem a apresentar valores de intensidade ligeiramente mais baixos, sinalizando possíveis "reinícios" de intensidade ao finalizar unidades discursivas.

Por outro lado, a duração não se mostrou um elemento crucial na percepção das fronteiras de eventos, destacando que, ao contrário do que foi constatado em estudos anteriores, a variação temporal não é determinante nesse contexto. Em relação às pausas silenciosas, estas se mostraram distintivas na estrutura discursiva em geral, mas são particularmente marcadas em fronteiras prosódicas associadas a eventos, indicando uma transição nítida entre eventos narrados.

Finalmente, ressaltamos que a prosódia, especialmente a diferença de tom e as pausas silenciosas, não apenas contribuíram para a percepção de eventos individuais, mas também desempenharam um papel fundamental na organização global do discurso oral. As complexas interações entre esses elementos prosódicos são essenciais para a delimitação das fronteiras de eventos, sinalizando aos ouvintes a estrutura de eventos dentro do discurso oral. Este entendimento aprofundado das nuances da

prosódia expande nosso conhecimento sobre como os elementos suprassegmentais fundamentam a representação dos eventos na língua falada, fornecendo uma visão mais abrangente da percepção linguística.

Considerações finais

Neste estudo, investigamos o papel dos elementos prosódicos no processamento da estrutura de eventos em narrativas orais espontâneas. O propósito era identificar que elementos prosódicos eram os mais salientes para o ouvinte na tarefa de percepção das fronteiras de eventos, simbolizadas, em termos estruturais, pelas fronteiras prosódicas. Constatamos que as fronteiras de eventos dentro do texto narrativo tinham características específicas: maior diferença de F0, valores de intensidade mais baixos na sílaba tônica imediatamente anterior à fronteira prosódica e presença significativa de pausas silenciosas.

Embora os elementos prosódicos analisados neste estudo sejam relevantes, ainda há outros que podem ser observados em uma análise que pretenda ser mais abrangente, como tons de fronteira, duração de pausas silenciosas, declínio de F0, entre outros. Acreditamos também que o uso de técnicas de experimentação *on-line*, como a pupilometria e a eletroencefalografia, é uma abordagem metodológica promissora e inovadora para investigar a relação entre a percepção de aspectos prosódicos e o processamento da estrutura de eventos narrados oralmente.

Referências

ALMEIDA, A. N. *Análise prosódica de agrupamentos numéricos no português do Brasil.* 2017. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Letras e Linguística, Universidade Federal de Alagoas, Maceió, 2017.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos da Linguagem*, v. 20, n. 1, p. 11-27, 2012.

BARBOSA, P. A. *Prosódia*. São Paulo, SP: Parábola Editorial, 2019.

BYRD, D.; RIGGS, D. Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association*, v. 38, p. 187-202, 2008.

DE PIJPER, J. R.; SANDERMAN, A. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America*, v. 96, n. 4, p. 2037-2047, 1994.

GROSZ, B.; SIDNER, C. L. Attentions, intentions, and the structure of discourse. *Computational Linguistics*, v. 85, p. 363-394, 1986.

GROSZ, B.; HIRSCHBERG, J. Some intonational characteristics of discourse structure. *In:* INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, 2., 1992, Banff. *Anais [...]* Banff: ISCA Archive, 1992. p. 429-432

HIRST, D.; DI CRISTO, A. *Intonation systems*: a survey of twenty languages. Cambridge: Cambridge University Press, 1998.

KATSIKA, A.; KRIVOKAPIC, J.; MOOSHAMMER, C.; TIEDE, M.; GOLDSTEIN, L. The coordination of boundary tones and its interaction with prominence**.** *Journal of Phonetics*. v. 44, p. 62-82, 2014.

KIM, J. *Individual Differences in the Production and Perception of Prosodic Boundaries in American English*. 2020. Thesis (Doctor in Philosophy) – The University of Michigan, Michigan, 2020.

LIN, Hsin-Yi; FON, J. The Role of Pitch Reset in Perception at Discourse Boundaries.*In:* INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 17, 2011, Hong Kong. *Anais [...]* Hong Kong: International Phonetic Association, 2011. p. 1242-1245.

OLIVEIRA JR., M. *Prosodic features in spontaneous narratives*. 2000. 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. *Revista do GELNE*, v. 4, n. 1, p. 1-6, 2002.

OLIVEIRA JR.; CRUZ, R.; SILVA, E. W. R. A relação entre a prosódia e a estrutura de narrativas espontâneas: um estudo perceptual. *Revista Diadorim*, v. 12, p. 38-53, 2012.

PASSONNEAU, R. J; LITMAN, D. J. Discourse segmentation by human and automated means. *Computational Linguistics*, v. 23, n. 1, p. 103-139, 1997.

PEREIRA, J. *Estudo perceptual da prosódia como elemento delimitador da estrutura de narrativas orais espontâneas*: a diferença de tom. 2014. Dissertação (Mestrado em Letras) – Universidade Federal do Pará, Belém, 2014.

RADVANSKY, G. A.; ZACKS, J. M. *Event cognition*. Oxford: Oxford University Press, 2014.

SILVA, E. W. R. *A relação entre produção e percepção de pistas prosódicas na segmentação de narrativas espontâneas*. 2017. Dissertação (Mestrado em Linguística) – Universidade Federal de Alagoas, Maceió, 2017.

SPEER, N.K.; ZACKS, J.M.; REYNOLDS, J.R. Human brain activity time-locked to narrative event boundaries. *Psychological Science*, v. 18, p. 449-455, 2007.

SULLIVAN, A. D. *Determining an inter-rater agreement metric for researchers evaluating student pathways in problem solving*. 2014. Dissertação (Mestrado em Ciências) – Iowa State University, Iowa, 2014.

SWERTS, M. Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America*, v. 101, n. 1, p. 514-521, 1996.

SWERTS, M.; GELUYKENS, R. Prosody as a marker of information flow in spoken discourse. *Language and Speech*, v. 37, n. 1, p. 21-43, 1994.

TERTO, A.; OLIVEIRA JR., M. A prosódia do metadiscurso: uma análise a partir dos dados do NURC Digital Recife. *Cadernos de Linguística*, v. 2, n. 4, p. 1-28, 2021.

TRUCKENBRODT, H. ; FERY, C. Hierarchical organization and tonal scaling. *Phonology*. v. 32. p. 19-47, 2015.

TURK, A. E.; SHATTUCK-HUFNAGEL, S. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, v. 35, p. 445-472, 2007.

WICHMANN, A. *Intonation in text and discourse*: beginnings, middles and ends. Edinburgh Gate: Pearson Education, 2000.

ZACKS, J. M. *The Importance of Events in Conception and Language*. Leiden, The Netherlands: Brill, 2020.

ZACKS, J.M.; SPEER, N. K.; SWALLOW, K. M.; BRAVER, T. S.; REYNOLDS, J. R. Event perception: A mind-brain perspective. *Psychological Bulletin*, v. 133, p. 273-293, 2007.

ZACKS, J. M.; TVERSKY, B. Event structure in perception and conception. *Psychological Bulletin*, v. 127, p. 3-21, 2001.

ZWANN, R.A. Time in language, situation models, and mental simulations. *Language Learning*, v. 58, n. 1, p. 13-26, 2008.