



Os efeitos prosódicos do uso da máscara de pano na leitura de uma narrativa

The prosodic effects of using fabric face mask in reading a narrative

Aline de Lima BENEVIDES*

RESUMO: O presente trabalho investiga se o uso de máscara de pano pode afetar parâmetros prosódicos por meio de tarefas de leitura de uma narrativa e, ainda, se tais mudanças se assemelham a disfarces de voz, como os utilizados por indivíduos em prática forense. Contrastaram-se, para tanto, três condições experimentais: *leitura normal* (condição controle), *leitura com máscara* e *leitura com lápis*, produzidas por quatro falantes nativos do PB, em um estudo piloto. Analisaram-se os seguintes parâmetros: taxa de elocução, taxa de articulação, pausas, duração das unidades V-V (BARBOSA, 2019), grupos acentuais e tons. Os resultados desta pesquisa mostram que as unidades V-V, o as pausas e o grupo acentual são os parâmetros prosódicos afetados pelo uso da máscara de pano. Além disso, observou-se que a *leitura com máscara* não apresenta comportamento similar à *leitura com lápis*, tampouco a *leitura normal*. Ao contrário, a *leitura com máscara* expõe comportamentos intermediários. Desse modo, estudos linguísticos que incidam sobre aspectos prosódicos devem considerar a máscara também como um disfarce de voz, uma vez que ela parece alterar alguns parâmetros prosódicos.

PALAVRAS-CHAVE: Prosódia. Fonética. Fonologia. Disfarce de voz. Máscara de pano.

ABSTRACT: This paper investigates whether fabric face mask can affect prosodic parameters through reading tasks of a narrative and, even, whether such changes resemble voice disguises, such as those used by individuals in forensic practice. For this purpose, three experimental conditions were contrasted: normal reading (control condition), reading with a mask and reading with a pencil, produced by four native speakers of BP, in a pilot study. The following parameters were analyzed: rate of speech, rate of articulation, pauses, duration of V-V units (BARBOSA, 2019), accent groups and tones. The results show that the V-V units, the pauses, and the accent group are the prosodic parameters affected by the use of the cloth mask. In addition, reading with a mask does not exhibit similar behavior to reading with a pencil, nor does normal reading. On the contrary, reading with a mask exposes intermediate behaviors. Thus, linguistic studies that focus on prosodic aspects should consider the mask also as a voice disguise, since it can change some prosodic parameters.

KEYWORDS: Prosody. Phonetic. Phonology. Voice disguise. Fabric face mask.

Artigo recebido em: 11.01.2023
Artigo aprovado em: 11.04.2023

* Doutora em Letras pela USP. Pesquisadora independente. benevides.aline12@gmail.com

1 Introdução

Diferentes instrumentos são correntemente utilizados por indivíduos que buscam disfarçar suas vozes. Künzel (2000) distingue-os em quatro tipos: i. disfarce em que o sujeito busca alterar a voz, por meio de rouquidão, sussurro etc.; ii. disfarce com a utilização de algum instrumento no trato vocálico que vise à modificação das características de ressonância; iii. disfarce de língua, em que o indivíduo busca ocultar ou imitar um sotaque, por exemplo; e iv. disfarce em que se modifica a forma de falar, com alteração no tempo de fala ou com a marcação de exageros.

O disfarce, segundo Gillier (2011), caracteriza-se como uma “ação deliberada de um falante que altera a sua voz, discurso ou língua com o propósito de esconder a sua identidade” (p. 20). Por ser um instrumento frequentemente empregado por criminosos na busca por camuflar a sua fala e, conseqüentemente, cercear a produção de provas, áreas como a Fonética Forense surgem com o propósito de investigar atributos específicos da fala de indivíduos, a partir do aparato teórico e metodológico da Fonética, sobretudo da Fonética Acústica. Sua investigação origina-se em uma análise comparativa de duas amostras, com o objetivo de indicar o quão similar ou dissimilar elas são, isto é, a natureza da variação intra e entre falantes a partir de vozes sonoramente similares (ROSE, 1996; GILLIER, 2011). Buscam-se, dessa maneira, características específicas na voz de um falante que permitam distingui-lo dos demais indivíduos (GILLIER, 2011).

Os estudos forenses relacionados à voz têm como objeto de análise escutas ou ligações telefônicas usadas na comunicação entre os indivíduos investigados e, também, em contato com as vítimas. Diante das frequentes inovações tecnológicas, diferentes aplicativos de compartilhamento de mensagens e de arquivos de áudios vêm substituindo a ligação em si. Não só porque se mostram formas mais práticas de transmitir mensagens, mas, nesses casos, por terem um alto índice de segurança do conteúdo enviado, devido ao sistema de criptografia de aplicativos como o WhatsApp e o Telegram.

Os parâmetros acústicos são os mais utilizados para a identificação dos falantes pela literatura da área forense, especificamente os formantes, a duração e a frequência fundamental (ROSE, 1996; GILLIER, 2011; PASSETTI; BARBORA, 2018). Contudo, informações prosódicas também podem contribuir de maneira significativa para a identificação dos falantes, tendo em vista que fenômenos de ritmo, de entoação e de acento também podem ser percebidos pelos ouvintes para discriminar um determinado falante. Uma simples mudança no sistema respiratório pode contribuir, por exemplo, para modificações na frequência fundamental (ADAMI, 2007). Uma alteração do fluxo respiratório pode ser motivada por fatores também muito simples, como o uso de uma máscara de pano. Esta, ainda, pode contribuir de maneira significativa para a restrição dos movimentos articulatorios.

Em momentos de pandemia, como o atualmente vivenciado, entre 2020 e 2022, que requerem o uso constante de uma máscara, o disfarce da voz nem sempre se torna uma ação deliberada do falante, tendo em vista a necessidade de proteção. Diante de tal cenário, o objetivo do presente trabalho é analisar se e quais parâmetros prosódicos podem ser influenciados pelo uso da máscara de tecido, a partir de um estudo piloto. Para tanto, busca-se tecer uma análise comparativa entre a *leitura com máscara* com outros dois tipos de leitura: a *leitura normal* e a *leitura com lápis* (entre os dentes). A *leitura normal* trata-se da leitura controle, a partir da qual será possível verificar os parâmetros habituais de cada um dos locutores. A *leitura com lápis* visa à averiguação de quais parâmetros prosódicos tendem a ser manipulados por cada locutor ao realizar um disfarce de voz. As variáveis prosódicas analisadas neste trabalho consistem em: taxa de elocução e de articulação, relacionadas ao ritmo; pausas, que são frequentemente empregadas para demarcar as unidades entoacionais; duração das unidades V-V (BARBOSA, 2019), que se mostram fundamentais para a composição dos grupos acentuais, e, por fim, os tons de fronteira.

Para além dos estudos forenses, análises que envolvem os parâmetros prosódicos - e, também, acústicos - mostram-se relevantes diante do cenário incerto de

pandemia global. As pesquisas científicas que envolvem a linguagem, por exemplo, têm buscado maneiras diversas de adaptação de suas metodologias, em decorrência da impossibilidade de contato físico e/ou do uso de técnicas laboratoriais. Algumas pesquisas podem se valer de gravações via aplicativos como o Google Meet, o Skype e o Zoom; no entanto, outras são impossibilitadas em decorrência da necessidade de instrumentais específicos, o que poderia ser viabilizada pelo uso de máscara por todos os participantes da situação experimental - desde, claro, que nenhum dos atores envolvidos estejam no grupo considerado de risco -, se ela não se mostrar influente na produção da fala.

Tendo em vista tais fatos, o presente estudo investiga a influência da máscara de pano em relação aos parâmetros prosódicos descritos. Para tanto, apresentamos, conforme a literatura da área, na seção 2, uma breve descrição de cada um dos parâmetros analisados (taxa de elocução, taxa de articulação, pausas, duração das unidades V-V, grupos acentuais e tons); na seção 3, expomos as decisões metodológicas adotadas e o método empregado para a análise dos dados; em sequência, na seção 4, explicitamos os resultados encontrados e os analisamos; e, por fim, na seção 5, apresentamos as considerações finais.

2 Parâmetros prosódicos

Ao longo desta seção, apresentam-se breves definições a respeito dos parâmetros prosódicos investigados nesta pesquisa.

2.1 Taxa de elocução e de articulação

As taxas de elocução e de articulação consistem em medidas de análise prosódica relacionadas ao ritmo da fala. A taxa de elocução refere-se à quantidade de unidades linguísticas produzidas por um falante em um determinado tempo, a qual tende a indicar não especificamente a velocidade da fala, mas, em comparação à

quantidade de unidades fonológicas, pode indicar a frequência das elisões, de possíveis epênteses e das hipoarticulações. Para calculá-la, deve-se dividir a quantidade de unidades fonéticas pelo tempo dedicado a produção da fala. A taxa de articulação, por sua vez, refere-se, especificamente, à quantidade de unidades linguísticas produzidas por um falante em um dado tempo, descontando-se, neste caso, a duração das pausas silenciosas ou preenchidas (BARBOSA, 2019).

2.2 Pausas

As pausas caracterizam-se por trechos de silêncios ou de alongamento inseridos pelo falante ao longo da cadeia da fala. Há dois tipos de pausas: as pausas demarcativas e as pausas hesitativas. As pausas demarcativas, como o nome sugere, ocorrem entre sintagmas, com o intuito de delimitá-los e, também, de estabelecer a hierarquia conceitual/ informacional do texto. As pausas hesitativas, por sua vez, tendem a ser inseridas no interior de unidades entoacionais (MERLO, 2012). Nenhuma delas engloba o silêncio decorrente da obstrução da passagem do ar das consoantes oclusivas.

2.3 Duração de unidades V-V

A duração das unidades V-V, segundo Barbosa (2019), consiste em uma medida mais objetiva do que a duração silábica, que visa à relação da duração silábica com o padrão rítmico da língua. A unidade V-V caracteriza-se por unidades fonéticas que se delimitam do início de uma vogal ao início de outra vogal, isto é, “a fronteira de unidade silábica para fins de estudo prosódico” (BARBOSA, 2019, p. 55). O pico dessas unidades é utilizado para marcar os grupos acentuais.

2.4 Grupo acentual

Segundo Barbosa (2019), o grupo acentual trata-se do agrupamento de uma ou mais palavras fonológicas. Quando há mais de uma, somente uma das palavras que compõe o grupo acentual que tem proeminência se estabelece no enunciado. Por ser o português, uma língua de ritmo acentual com cabeça à direita da palavra, a proeminência dos grupos acentuais ocorre de modo semelhante: com proeminência na margem direita da palavra fonológica.

2.5 Tons de fronteira

Os tons de fronteira consistem em unidades entoacionais que auxiliam na segmentação da cadeia da fala em enunciados e sintagmas (BARBOSA, 2019). Quando em conjunção com o acento de *pitch* e com o ritmo, permitem a marcação de atitudes, expressões etc. Adota-se, aqui, para fins de codificação dos dados em análise, o sistema de notação do DaTo (LUCENTE; BARBOSA, 2009), o qual possui seis contornos dinâmicos (LH, > LH, HLH, HL, > Hl e LHL) e dois tons de proeminência e de fronteira (L e H), este fundamental para a presente pesquisa.

3 Metodologia

3.1 Participantes

Participaram deste estudo 4 falantes nativos do português brasileiro, residentes na região metropolitana de São Paulo¹, sendo 2 homens e 2 mulheres, com idades entre

¹ Deve-se ressaltar que o número restrito de falantes decorre de, no momento de coletas de dados, o contexto de pandemia que não ter permitido o contato com um número abrangente de falantes, dado que o estudo tem que ser realizado presencialmente e com a utilização dos mesmos instrumentos de coleta de dados.

27 e 34 anos. Nenhum dos falantes relatou qualquer problema cognitivo, auditivo ou fonoarticulatório².

3.2 Método

A tarefa experimental consistiu na leitura de um texto narrativo, especificamente uma adaptação da fábula de Esopo, *O lobo e a cegonha*, tal como está transcrita abaixo, de três modos distintos. A primeira, denominada de *leitura normal*, consistiu na leitura habitual do informante, sem uso de qualquer instrumento. A segunda, denominada de *leitura com máscara*, consistiu na leitura com a utilização de uma máscara de tecido modelo ninja (conforme demonstra a Figura 1), sendo a mesma para todos os informantes³. A terceira, denominada de *leitura com lápis*, consistiu na leitura com o uso de um lápis, especificamente do lápis Bic Evolution Hb n.2, entre os dentes incisivos. As leituras foram gravadas na mesma ordem descrita acima.

O lobo e a cegonha (trecho 1)

Um lobo devorou sua caça tão depressa, com tanto apetite, que acabou ficando com um osso entalado na garganta (trecho 2).

Cheio de dor, o lobo começou a correr de um lado para outro soltando uivos, e ofereceu uma bela recompensa para quem tirasse o osso de sua garganta (trecho 3). Com pena do lobo e com vontade de ganhar o dinheiro, uma cegonha resolveu enfrentar o perigo (trecho 4).

Depois de tirar o osso, quis saber onde estava a recompensa que o lobo tinha prometido (trecho 5).

– Recompensa? – berrou o lobo (trecho 6). – Mas que cegonha pechinchona! (trecho 7) Que recompensa, que nada! (trecho 8) Você enfiou a cabeça na minha boca e em vez de arrancar sua cabeça com uma dentada deixei que você a tirasse lá de dentro sem um arranhãozinho (trecho 9). Você não acha que tem muita sorte? (trecho 10) Seu bicho insolente! (trecho 11) Dê o fora e se cuide para nunca mais chegar perto de minhas garras! (trecho 12)

² Este estudo piloto está compreendido em um projeto de pesquisa registrado e aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH/USP), cujo número do processo é 3.332.409.

³ Salienta-se que a máscara foi higienizada entre as gravações, tendo em vista o contexto de pandemia.

Moral: Não espere gratidão ao mostrar caridade para um inimigo (trecho 13).

Figura 1 — Máscara de tecido modelo ninja.



Fonte: arquivo da autora.

Todas as gravações foram realizadas em sala comum, por meio do aplicativo do WhatsApp no celular LG K10. O propósito do uso de tal aplicativo foi buscar simular um contexto forense em que os indivíduos se valessem desse instrumento como meio de comunicação, devido ao seu sistema de criptografia e de segurança. As gravações teriam, dessa maneira, interferências acústicas semelhantes a uma situação real.

3.3 Tratamento e análise dos dados

Os arquivos de áudio foram convertidos em .wav através do aplicativo Audacity, versão 2.4.1 (Audacity Team, 2019) e segmentados e transcritos no aplicativo Praat, versão 6.0.37 (BOERSMA; WEENINK, 2019). A transcrição considerou a marcação das unidades V-V e das pausas e a delimitação da narrativa em enunciados. As medições foram realizadas automaticamente por meio do uso do *script Prosody Descriptor Extractor*, de Barbosa (2020).

Após o processo de codificação, os dados foram analisados no *software* R (R Core Team, 2019), com a utilização do modelo de regressão logística, através da função *glm*.

4 Resultados e análise

4.1 Taxa de elocução e de articulação

A seguir, expõem-se os dados referentes à taxa de elocução e de articulação, obtidos nas três condições experimentais por falante, calculadas a partir de unidades V-V fonéticas por segundo:

Tabela 1 — Taxa de elocução e de articulação por falante, condição de leitura e unidades V-V.

Falante	Condição	Taxa de elocução	Taxa de articulação	Unidades V-V
H1	normal	3,7	4,6	253
	máscara	3,8	5,0	263
	lápiz	3,5	4,3	241
H2	normal	4,1	5,1	247
	máscara	4,3	5,2	250
	lápiz	3,6	4,7	243
M1	normal	3,5	4,2	260
	máscara	3,9	4,7	268
	lápiz	3,4	4,2	259
M2	normal	4,1	4,9	255
	máscara	4,2	5,3	261
	lápiz	3,9	4,7	258

Fonte: a autora.

Observe que, descritivamente, os valores de taxa de elocução e de articulação apresentam pouca variação entre as condições experimentais e, até mesmo, entre os falantes: de 3,4 a 4,3 de taxa de elocução e de 4,2 a 5,3 de taxa de articulação. Análises estatísticas, por meio de modelos de regressão logística, não apontaram diferenças significativas entre as taxas de elocução e de articulação ($p > 0.05$), entre os falantes ($p > 0.05$) e, ainda, em relação à condição de leitura ($p > 0.05$). Nota-se, assim, que o ritmo de fala, no que se refere às taxas de elocução e de articulação, não sofreu alteração em

decorrência da utilização ou não do disfarce de voz. Os valores expressos na Tabela 1 seguem as taxas médias descritas na literatura da área, que apontam de 4 a 6 sílabas por segundo, sujeitas à variação conforme as características dos falantes e ao estilo de elocução (BARBOSA, 2019).

Embora as taxas obtidas nas condições experimentais não apresentem diferenças significativas, é interessante notar que os menores índices, seja de taxa de elocução, seja de taxa de articulação, em todos os falantes, ocorrem na *leitura com lápis*; em contrapartida, a *leitura com máscara* apresenta maiores índices. Hipotetiza-se que a dificuldade que a *leitura com lápis* impõe aos falantes, ao requerer simultaneamente a leitura da narrativa e a permanência do lápis entre os dentes, induz a uma leitura mais lenta e, também, menos cuidadosa. A menor articulação na *leitura com máscara* é constatada ao se comparar a quantidade de unidades V-V produzidas nesta condição em relação às demais - exceto na falante M2 -, em decorrência da supressão de algumas unidades fonéticas, especialmente na leitura dos homens.

4.2 Pausas

Conforme expõe a Tabela 2, a dificuldade no processo de leitura da *leitura com lápis* é verificada, ainda, através da quantidade de pausas inseridas, estratégia essa empregada por todos os falantes, embora nem o tempo das pausas ($p > 0.05$) nem a quantidade delas ($p = 0.002$) tenham se mostrado significativos. Apesar disso, observou-se que, de modo geral, as pausas na *leitura normal* e na *leitura com máscara* incidiram nas fronteiras dos sintagmas fonológicos - também denominadas de pausas demarcativas (MERLO, 2012) -, como uma maneira dos indivíduos intensificarem o fluxo respiratório, enquanto que, na *leitura com lápis*, a pausa se dava continuamente - mesmo dentro das unidades entoacionais.

Tabela 2 – Quantidade e tempo das pausas por falante e por condição de leitura.

Falante	Condição	Quantidade de Pausas	Tempo das pausas (s)
H1	normal	20	14,26
	máscara	24	15,83
	lápiz	27	12,53
H2	normal	23	12,37
	máscara	22	10,82
	lápiz	24	15,31
M1	normal	19	12,41
	máscara	24	11,51
	lápiz	30	14,95
M2	normal	23	10,18
	máscara	25	11,78
	lápiz	31	10,23

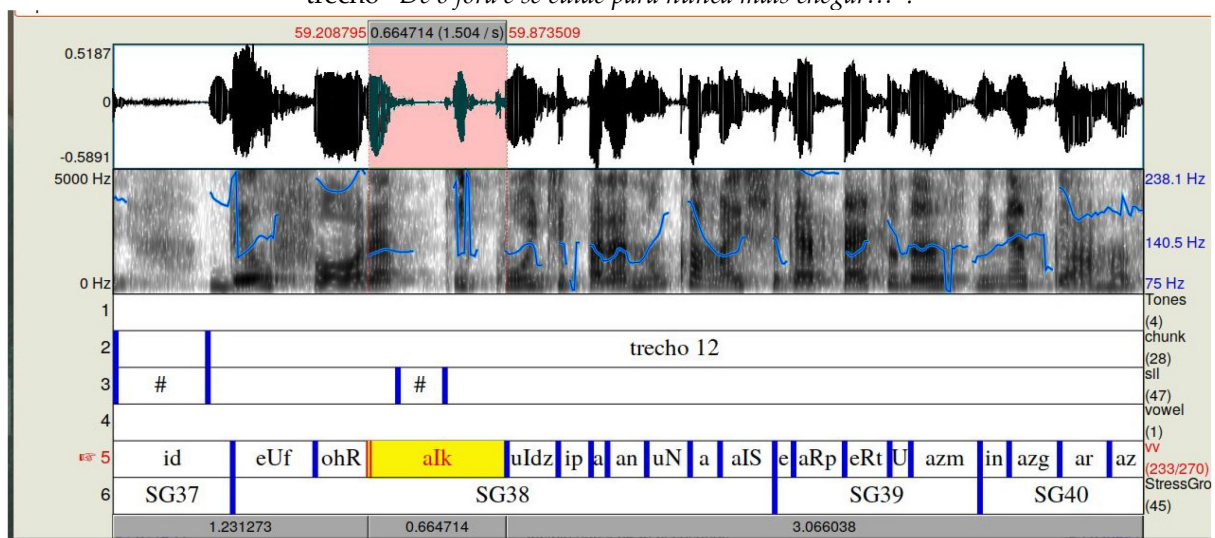
Fonte: a autora.

Em uma análise comparativa entre a *leitura normal* e a *leitura com máscara*, observa-se que o uso da máscara faz com que os falantes, com exceção do H2, insiram mais pausas, os quais apresentam diferenças significativas entre eles ($p < 0.05$) e entre as condições de leitura ($p < 0.05$). Essa diferença mostra-se significativa mesmo na comparação entre *leitura com máscara* e *leitura com lápis*, seja por falante ($p < 0.05$), seja por condição ($p < 0.05$), indicando, dessa maneira, que a pausa é um recurso utilizado de modo distinto pelos falantes e nessas duas condições de leitura. Grosso modo, a pausa, na *leitura com máscara*, não expõe resultados que a aproximem, quer da *leitura normal*, quer da *leitura com lápis*.

Em uma análise mais minuciosa, percebe-se que a falante M1, por exemplo, não só inclui uma quantidade expressivamente maior de pausas na *leitura com lápis* do que na *leitura normal*, como também se vale do alongamento das unidades segmentais como uma maneira de conter o lápis, a saliva e a respiração. Observe, na Figura 2, que

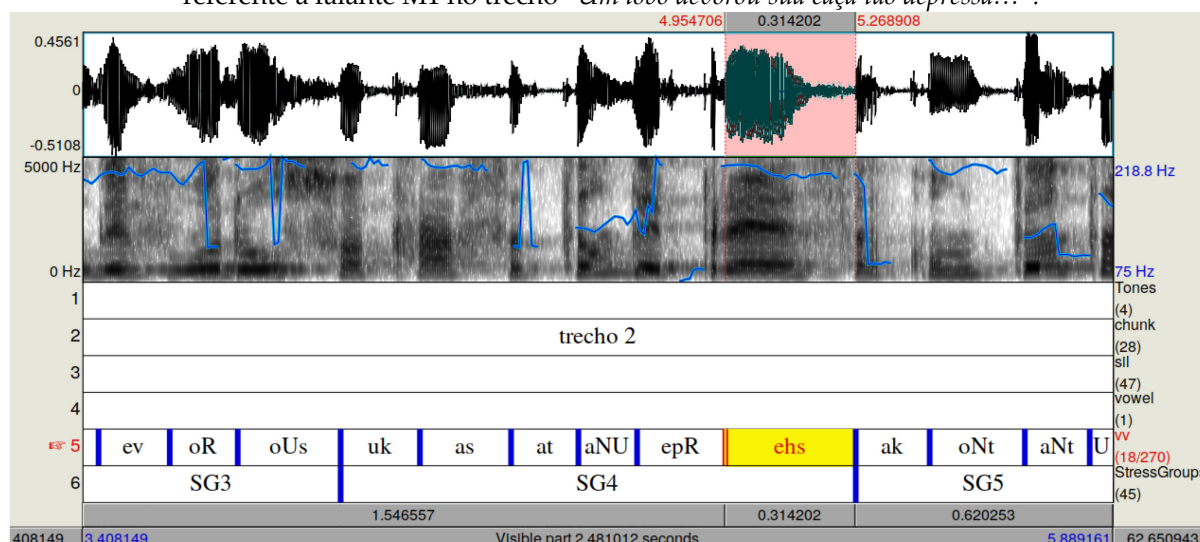
a participante insere breves espaços temporais entre as unidades de fala e, como no trecho em destaque, alonga a vogal em [ik], que deveria ser uma vogal reduzida – alongamento esse que pode ser considerado uma pausa preenchida. Embora, na Figura 2, os alongamentos ocorram dentro de um grupo acentual, eles, em geral, são também utilizados como elementos delimitadores dos grupos acentuais (MERLO, 2012 apud BARBOSA, 2006), como pode ser visualizado na Figura 3. Na *leitura com lápis*, portanto, as pausas delimitam não só as fronteiras dos sintagmas fonológicos, mas também são usadas internamente a eles. Há, na *leitura com lápis*, a presença frequente de pausas tanto demarcativas como delimitativas.

Figura 2 — Exemplo de inserção de alongamento vocálico no espectrograma referente à falante M1 no trecho “Dê o fora e se cuide para nunca mais chegar...”.



Fonte: a autora.

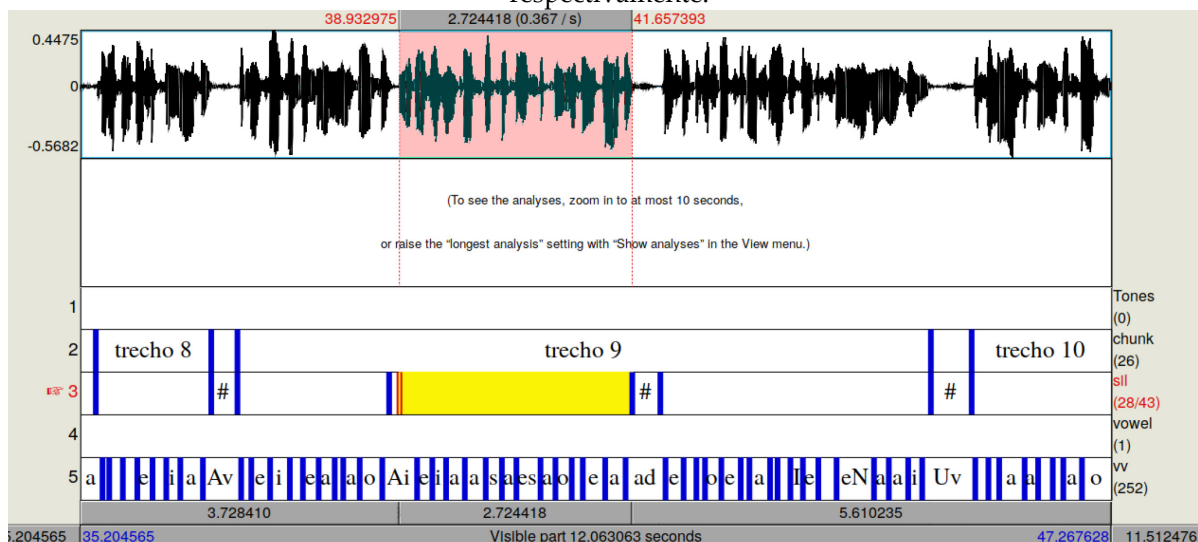
Figura 3 — Exemplo de inserção de alongamento em fronteira de grupo acentual no espectrograma referente à falante M1 no trecho “Um lobo devorou sua caça tão depressa...”.



Fonte: a autora.

A *leitura normal* e a *leitura com máscara*, em contrapartida, apresentam apenas pausas demarcativas, que delimitam os sintagmas fonológicos, como exposto na Figura 4. Note que o falante insere dentro de um mesmo trecho duas pausas, todas elas para demarcar a ruptura de sintagma, com pausas silenciosas. A marcação de pausas mostra-se, desse modo, uma característica que aproxima a *leitura com máscara* da *leitura normal*. A diferença entre elas incide, sobretudo na necessidade de maiores rupturas sintagmáticas na *leitura com máscara* em decorrência da obstrução respiratória que a máscara provoca. Tal constatação deve-se a uma observação realizada pela pesquisadora durante a coleta de dados, momento em que se observou uma leitura mais ofegante e cansativa por parte dos falantes na *leitura com máscara*, motivada, neste caso, pela obstrução da circulação do ar. Observa-se, diante disso, que, em textos maiores ou num ato de fala mais duradouro, por exemplo, em uma aula ou em situações experimentais mais duradouras, o uso da máscara pode trazer mais restrições e a incorporação de mais pausas do que a fala normal.

Figura 4 — Exemplo de inserção de pausas demarcativas entre unidades entoacionais referente ao falante H2 (leitura com máscara) na produção dos trechos “Você enfiou a cabeça na minha boca”, “e em vez de arrancar sua cabeça com uma dentada” e “deixei que você a tirasse lá de dentro sem um arranhãozinho”, respectivamente.

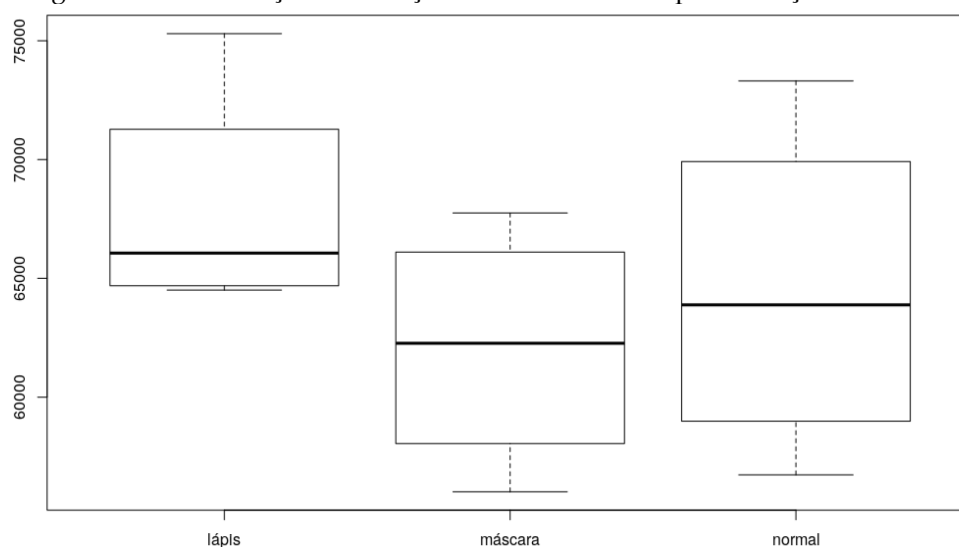


Fonte: elaborado pela autora.

4.3 Duração das unidades V-V

A seguir, na Figura 5, expõe-se a distribuição da duração média das unidades V-V por condição de leitura:

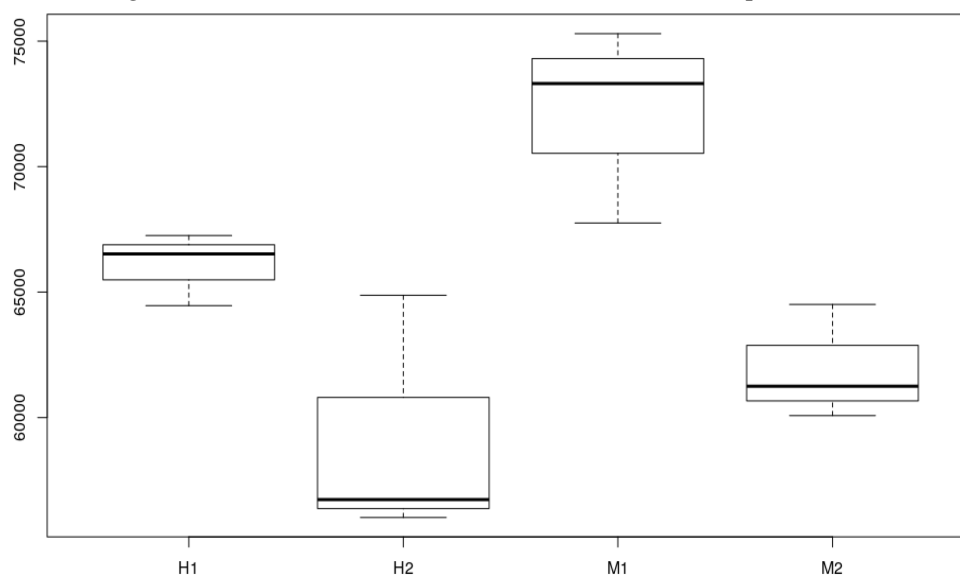
Figura 5 — Distribuição da duração das unidades V-V por condição de leitura.



Fonte: elaborado pela autora.

Conforme demonstram os dados da Figura 5, observe que a *leitura normal* é a condição que exibe uma maior variabilidade na duração das unidades V-V, a ser observada na abrangência do *boxplot* – note, ainda, que as medianas (linha mais escura no *boxplot*) são diferentes em todas as condições experimentais. Essa variabilidade está, sobretudo, associada aos diferentes comportamentos apresentados pelos falantes, como demonstra a Figura 6. Note que o falante H2 e a falante M1 apresentam atitudes distintas: o falante H2 tende a produzir tanto unidades fonéticas mais breves, como também a suprimir e a coarticular as unidades V-V; a falante M1, por sua vez, não só produz as unidades de forma mais articulada, como também com um maior alongamento. As *leituras com máscara e com lápis* evidenciam novamente resultados que vão em sentido opostos: as unidades em análise tendem a ser mais curtas na *leitura com máscara* e mais longas na *leitura com lápis*, tendo como referência à *leitura normal* (condição controle). A maior duração das unidades V-V na *leitura com lápis* já era esperada, na medida em que, como salientado, os falantes tenderam a produzir fones mais longos quando estavam com o lápis na boca. O prolongamento dos segmentos mostrou-se a estratégia encontrada pelos falantes para permanecer com o lápis na boca e, ainda, para controlar o fluxo respiratório. As análises inferenciais, a partir do modelo de regressão logística aplicado, indicam que tanto a condição como os falantes mostram-se fatores significativos para explicar a variabilidade de duração das unidades V-V, com valor de $p < 0,005$, em ambos os casos.

Figura 6 — Distribuição da duração em unidades V-V por falante.



Fonte: elaborado pela autora.

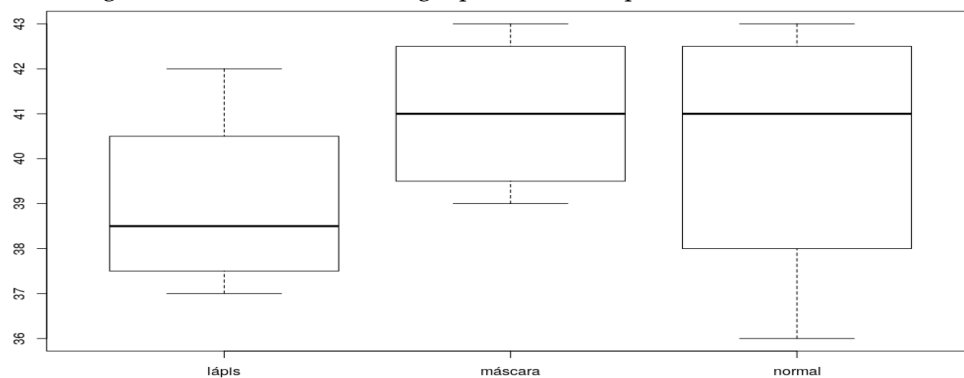
É interessante salientar, em suma, que a duração das unidades V-V destaca-se, aqui, como uma variável relevante para compor a discriminação dos falantes, na medida em que os comportamentos distintos entre todos os indivíduos investigados. Além disso, são essas unidades fundamentais para a determinação automática dos grupos acentuais, a serem descritos a seguir.

4.4 Grupos acentuais

A Figura 7 abaixo apresenta a distribuição de grupos acentuais por condição de leitura. Note que a *leitura com máscara* e a *leitura normal* apresentam uma quantidade maior de grupos acentuais e com semelhante mediana (veja a linha mais escura no *boxplot* na Figura 7), se comparado com a *leitura com lápis*. Deve-se destacar que, ainda que possa parecer que a maior duração das unidades V-V conduz a um maior número de grupos acentuais, estes são formados a partir do contraste forte e fraco, momento em que se destaca os picos duracionais, o que não implica que haja picos muito variáveis entre as durações na *leitura com lápis*; pode haver, no entanto, valores mais altos de forma mais homogênea. Os modelos estatísticos demonstram a significância

do grupo acentual para explicar a variabilidade nas condições de leitura e entre os falantes, com valor de $p < 0.005$.

Figura 7 — Distribuição dos grupos acentuais por condição de leitura.

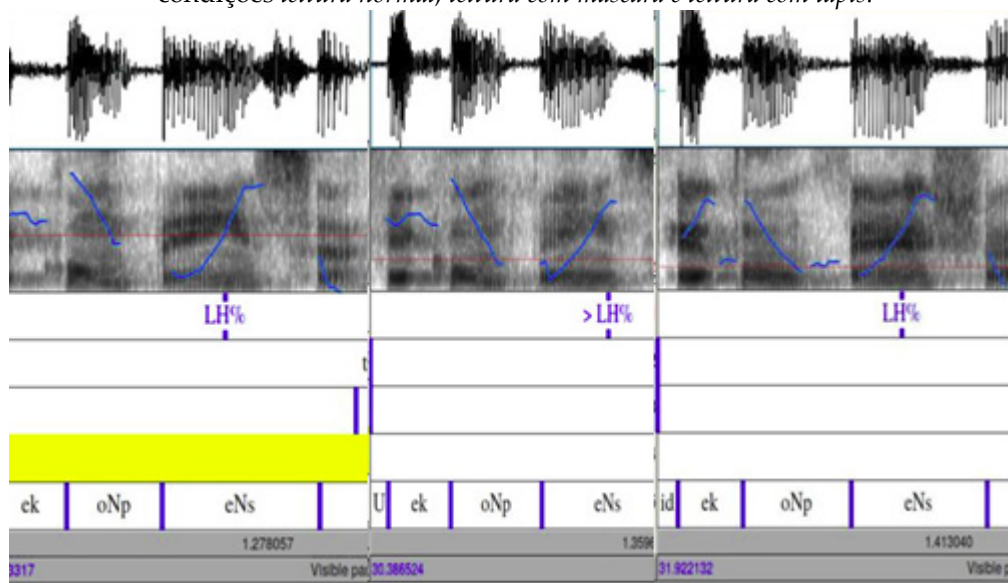


Fonte: elaborado pela autora.

4.5 Tons

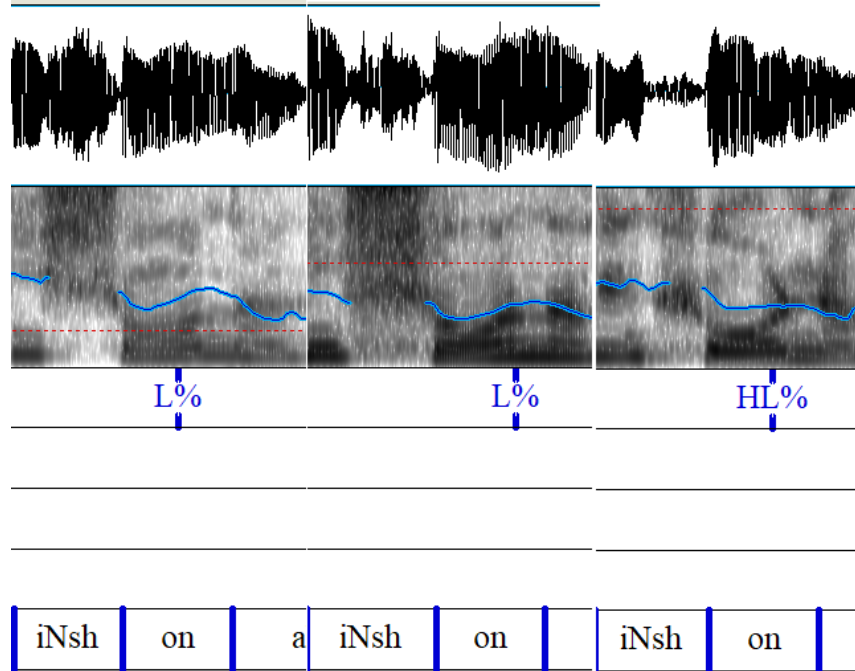
A seguir, apresenta-se, nas Figuras 8, 9 e 10, um exemplo amostral, referente à falante M2, da marcação de tons nas três condições de leitura, bem como na produção de perguntas, exclamações e afirmações, respectivamente.

Figura 8 — Marcação de tom referente à produção de M2 de “- Recompensa?”, respectivamente, nas condições *leitura normal*, *leitura com máscara* e *leitura com lápis*.



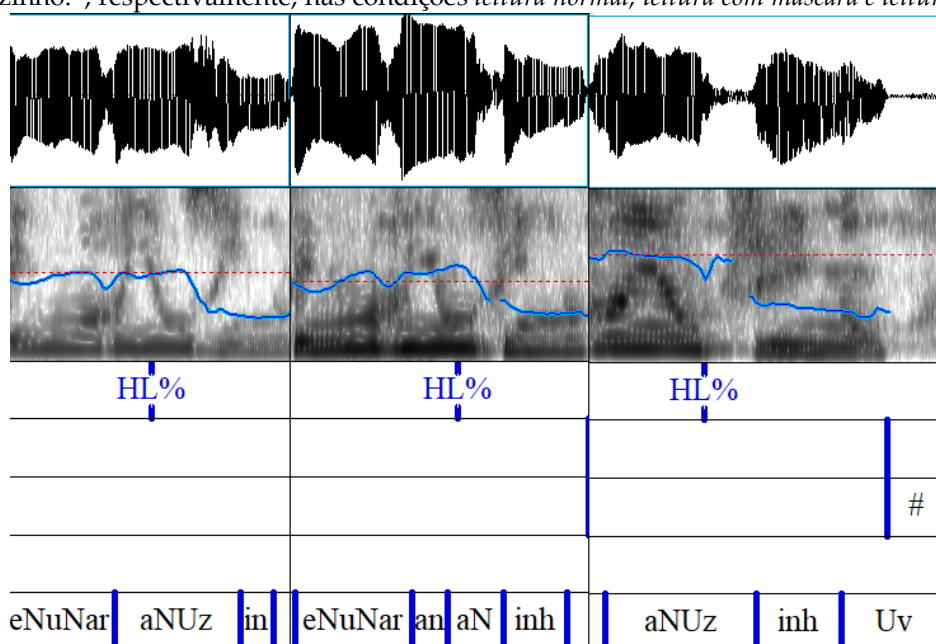
Fonte: elaborado pela autora.

Figura 9 — Marcação de tom referente à produção de M2 do trecho “- Mas que cegonha pechinchona!”, respectivamente, nas condições *leitura normal*, *leitura com máscara* e *leitura com lápis*.



Fonte: elaborado pela autora.

Figura 10 — Marcação de tom referente à produção de M2 de “(...) você a tirasse lá de dentro sem um arranhãozinho.”, respectivamente, nas condições *leitura normal*, *leitura com máscara* e *leitura com lápis*.



Fonte: elaborado pela autora.

Como se pode observar nas figuras expostas acima, marcaram-se as fronteiras entoacionais de três trechos (6, 7 e 9) selecionados na narrativa em análise, lidos nas

três condições de leitura (*leitura normal, leitura com máscara e leitura com lápis*, respectivamente). As figuras trazem a produção de uma mesma falante, a M2. Observe que os tons de fronteira atribuídos a cada um dos tipos de sentenças (pergunta, exclamação ou afirmação) são os mesmos independente da condição de leitura nos trechos apresentados. De modo geral, como exposto na Tabela 3, os falantes apresentam uma taxa média de concordância entre os tons de fronteira de 28,4%, considerando-se, neste caso, a comparação entre as condições experimentais e o tom por trecho para cada falante. Vale destacar que o trecho foi considerado divergente caso o falante produzisse em uma das diferentes condições outro tom de fronteira.

Tabela 3 – Percentual de convergência de tons de fronteira por falante entre os trechos e as condições experimentais.

Falante	% de convergência (número de trechos)
M1	53% (7)
M2	30% (4)
H1	7,7% (1)
H2	23,1% (3)
Média dos falantes	28,4% (15)

Os homens apresentam menores percentuais de divergência entre os tons. Eles, em geral, produziram recorrentemente tons baixos em quase todos os trechos da narrativa, independentemente de a sentença ser afirmativa, exclamativa ou interrogativa. Os trechos que apresentaram maior divergência foram os trechos 3, 4 e 10, sendo os dois primeiros afirmativos e o último, interrogativo. No comparativo entre todos os tons empregados em todos os trechos e as condições experimentais, há um percentual de apenas 9,6% de diferença, o que evidencia que o contorno tonal não parece ser um parâmetro relevante, seja para a discriminação do falante, seja da condição de leitura – não há sistematicidade em relação a esta. Tal fato era esperado, visto que um leitor proficiente tenderia a tentar preservar os contornos entoacionais característicos da língua, apesar de se valer de um instrumento para disfarçar a voz.

5 Considerações finais

O presente trabalho analisou quais os efeitos prosódicos do uso da máscara de pano em tarefas de leitura de uma narrativa. Para tanto, contrastaram-se três diferentes condições: *leitura normal* (condição controle), *leitura com máscara* e *leitura com lápis*, a partir da análise dos seguintes fenômenos prosódicos: taxa de elocução e de articulação, pausas, duração das unidades V-V, grupo acentual e tons de fronteira. O propósito era, também, observar a qual condição, *leitura normal ou com lápis*, a *leitura com máscara* se aproximava em termos prosódicos.

No decorrer deste trabalho piloto, verificou-se que a taxa de elocução, de articulação e os tons de fronteira não se mostraram significativos para a discriminação dos sujeitos e das condições de leitura. Entretanto, as demais variáveis, grupo acentual, duração das unidades V-V e pausas, segundo o modelo de regressão logística, são significativas para explicar a variabilidade encontrada nos dados. É interessante destacar que, apesar do uso de disfarce de voz em ambas as condições, os fenômenos prosódicos não exibiram tendências semelhantes entre a *leitura com máscara* e a *leitura com lápis*; ao contrário disso, elas parecem desencadear estratégias opostas. Por exemplo, na *leitura com máscara*, constatou-se que a duração das unidades V-V apresentava taxas significativamente menores do que as taxas verificadas na *leitura com lápis*. Além disso, constatou-se que o uso do *lápis* implicou não só na inserção de um maior número de pausas, como também em uma distribuição mais homogênea ao longo da leitura, o que resultou no uso tanto de pausas demarcativas como hesitativas, característica essa peculiar a essa condição de leitura.

Diante do exposto, verificou-se que a máscara de pano parece modificar parâmetros prosódicos da fala, tais como a duração das unidades V-V, o número de pausas e de grupos acentuais. Desse modo, a partir do que este estudo piloto indica, estudos linguísticos, que passem a incorporar esse instrumento como forma de proteção dos participantes da pesquisa, se o estudo incidir na prosódia, devem, antes, buscar um estudo mais aprofundado e detalhado a respeito dos efeitos das máscaras,

visto que ela também parece atuar como um disfarce de voz, já que alterou, neste estudo incipiente, alguns parâmetros prosódicos.

Referências

ADAMI, A. G. Modeling prosodic differences for speaker recognition. **Speech Communication**, 49, p. 277-291, 2007. DOI <https://doi.org/10.1016/j.specom.2007.02.005>

Audacity Team. **Audacity(R)**: Free Audio Editor and Recorder [Computer application]. Version 2.4.1. Disponível em: <https://audacityteam.org>. Acesso em: 2019.

BARBOSA, P. A. **Script Prosody Descriptor Extractor.psc**, 2020.

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

BOERSMA, P.; WEENINK, D. **Praat**: doing phonetics by computer (Versão 6.0.37). Disponível em: <http://www.praat.org>. Acesso em: 2019.

GILLIER, R. **O disfarce da voz em Fonética Forense**. 2011. 103f. Tese (Mestrado em Linguística). Departamento de Linguística Geral e Românica, Faculdade de Letras, Universidade de Lisboa, Lisboa, 2011.

KÜNZEL, H. Effects of voice disguise on speaking fundamental frequency. **Journal of Speech, Language and the Law**, n. 7, v. 2, p. 150-179, 2000. DOI <https://doi.org/10.1558/sll.2000.7.2.149>

LUCENTE, L.; BARBOSA, P. A. Sistema DaTo de notação entoacional do português brasileiro. Teoria e Funcionamento. **Cadernos de Pesquisas em Linguística**, v. 4, p. 41-66, 2009.

MERLO, S. **Dinâmica temporal de pausas e hesitações na fala semi-espontânea**. Tese (Doutorado em Linguística). 2012. 179f. Instituto de Estudos da Linguagem, Universidade de Campinas, Campinas, 2012. DOI <https://doi.org/10.20396/cel.v54i1.8636968>

PASSETTI, R. R.; BARBOSA, P. A. Análise fonético-acústica de disfarces vocais com aplicações para a fonética forense. **Cadernos de Estudos Lingüísticos**, Campinas, v.60, n.3, p. 1-17, 2018. DOI <https://doi.org/10.20396/cel.v60i3.8651760>

R Core Team. **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: <http://www.R-project.org/>. Acesso em: 2019.

ROSE, P. **Speaker verification under realistic forensic conditions**. Disponível em: http://philjohnrose.net/pubs/FVC_pubs/Rose_papers_on_FVC/Rose_1996_SST_Realistic_FSL.pdf. Acesso em: jun. 2020.