

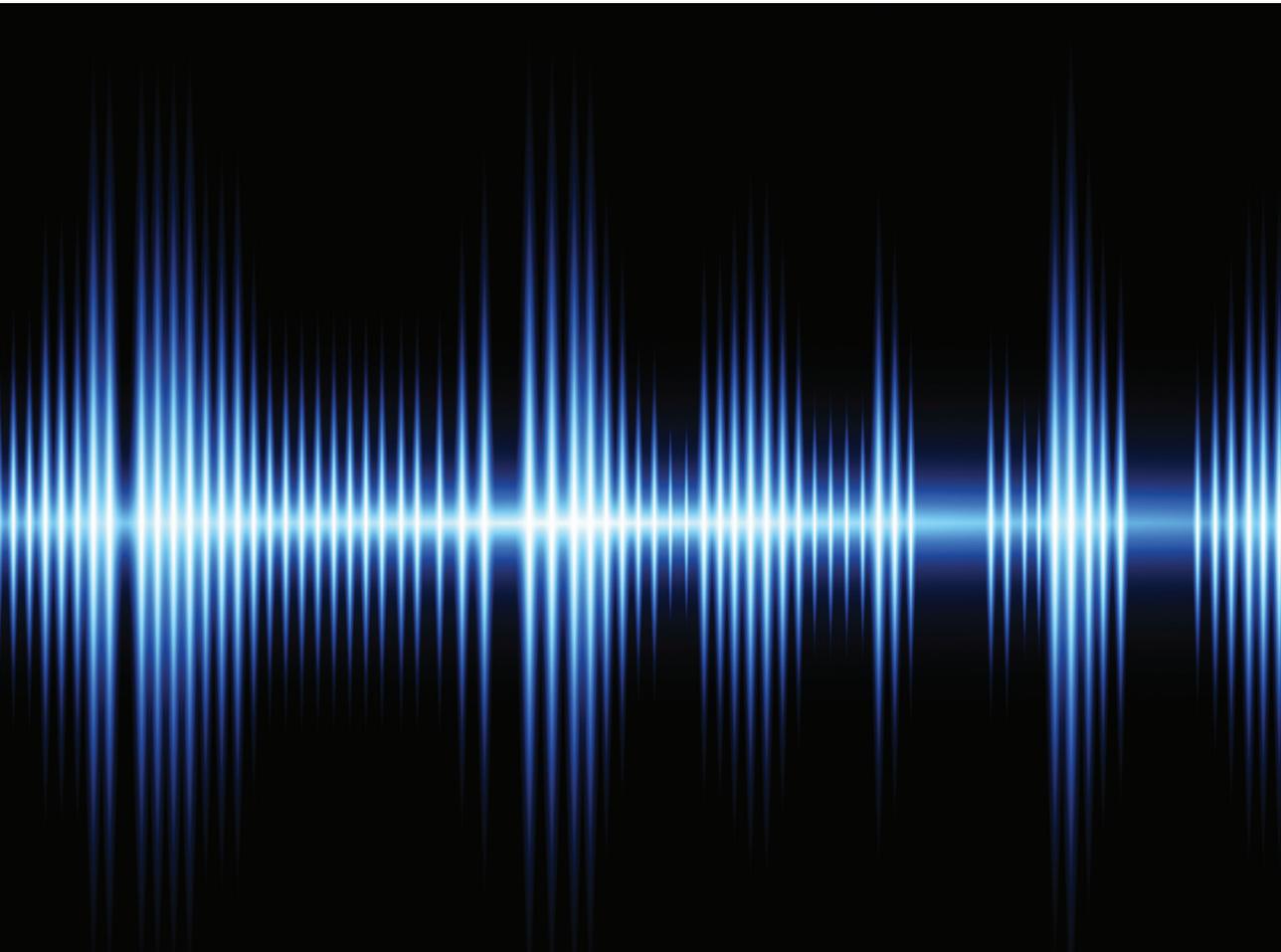
Raquel Meister Ko. Freitag

Luciana Lucente

organizadoras

# PROSÓDIA DA FALA:

pesquisa e ensino



Raquel Meister Ko. Freitag  
Luciana Lucente  
(organizadoras)

# Prosódia da fala: pesquisa e ensino

Livro financiado pelo Edital CAPES/PAEP/2016

**Blucher**

*Prosódia da fala: pesquisa e ensino*

© 2017 Raquel Meister Ko.Freitag, Luciana Lucente (organizadoras)

Editora Edgard Blücher Ltda.

---

# Blucher

---

Rua Pedroso Alvarenga, 1245, 4º andar

04531-934 – São Paulo – SP – Brasil

Tel 55 11 3078-5366

**contato@blucher.com.br**

**www.blucher.com.br**

Segundo Novo Acordo Ortográfico, conforme 5. ed.  
do *Vocabulário Ortográfico da Língua Portuguesa*,  
Academia Brasileira de Letras, março de 2009.

É proibida a reprodução total ou parcial por quaisquer  
meios, sem autorização escrita da editora.

---

Todos os direitos reservados pela editora  
Edgard Blücher Ltda.

## FICHA CATALOGRÁFICA

---

Prosódia da fala : pesquisa e ensino [livro eletrônico] /  
Raquel Meister Ko.Freitag, Luciana Lucente (orgs.). –  
São Paulo : Blucher, 2017.  
162 p. : PDF ; il. color.

### Bibliografia

ISBN 978-85-8039-259-3 (e-book)

ISBN 978-85-8039-258-6 (impresso)

1. Linguística 2. Análise prosódica (Lingüística) 3.  
Entonação (Fonética) I. Freitag, Raquel Meister Ko. II.  
Lucente, Luciana.

17-0832

CDD 414.6

---

Índices para catálogo sistemático:  
1. Linguística  
2. Análise prosódica (Lingüística)

# Apresentação

No Brasil, a área “dura” nos estudos linguísticos, especialmente fonética e fonologia, ainda é muito carente de recursos humanos e pesquisa, situação que não é diferente na Universidade Federal de Sergipe. Com a chancela da Associação Luso-Brasileira de Ciências da Fala (LBASS), a Universidade Federal de Sergipe realizou a IV Escola de Prosódia, no período de 24 a 26 de outubro de 2016, em Aracaju/SE.

Atualmente, a instituição tem 8 habilitações de cursos de Letras (Português, Inglês, Francês, Espanhol e Libras), na modalidade presencial e a distância, dois cursos de Fonoaudiologia (nos campi de São Cristóvão e Lagarto), três programas de mestrado em Letras (dois deles em ensino - Profletras), e recentemente iniciou o curso de Doutorado em Letras

Por ser de pequeno porte e ainda incipiente na área de pesquisa nos domínios linguísticos, a realização da IV Escola de Prosódia propiciou a indução de novas linhas de investigação na Universidade Federal de Sergipe, especialmente as de interface teórica, como prosódia e variação linguística (sociolinguística), prosódia e polidez (pragmática), assim como interfaces de pesquisa com alto poder de inserção social, como na área de saúde (Fonoaudiologia) e de ensino (Profletras).

A importância dada à pesquisa sobre a prosódia do português brasileiro tem sido traduzida em uma série de iniciativas realizadas nos últimos anos. Em agosto de 2007, organizado por João Antônio de Moraes, ocorreu na Universidade Federal do Rio de Janeiro o 1º Colóquio Brasileiro de Prosódia da Fala, o qual reuniu pesquisadores de vários centros universitários do país. Em fevereiro de 2007 surge a Luso Brazilian Association of Speech Sciences (LBASS) destinada num primeiro momento a apoiar oficialmente o Speech Prosody 2008, em maio de 2008. O Speech Prosody 2008, organizado por Plínio Almeida Barbosa, César Reis e Sandra Madureira, reuniu pesquisadores de várias partes do mundo. Em

2009, ocorreu o 2º Colóquio Brasileiro de Prosódia da Fala na Universidade Estadual de Campinas. Esse evento, promovido pela Luso Brazilian Association of Speech Sciences foi organizado pelo Grupo de Estudos de Prosódia da Fala, liderado por Plínio Almeida Barbosa. Nesse colóquio, surgiu a proposta de se organizar a I Escola de Prosódia, que aconteceu em 2010 na PUC-SP, organizada por Sandra Madureira e Zuleica Camargo. Cumprindo os objetivos propostos para a divulgação e ampliação dos estudos em prosódia no Brasil, bianualmente são realizados o Colóquio Brasileiro de Prosódia da Fala e a Escola de Prosódia. Em 2011, foi realizado o 3º Colóquio Brasileiro de Prosódia da Fala, na Universidade Federal de Minas Gerais, organizado por César Reis, e, em 2012, foi realizada a II Escola de Prosódia, na Universidade Federal do Espírito Santo, organizada por Alexssandro Meireles. Em 2013, foi realizado o 4º Colóquio Brasileiro de Prosódia da Fala, na Universidade Federal de Alagoas, organizado por Miguel Oliveira Jr. Em 2014, foi realizada a III Escola de Prosódia, na Unicamp, mais uma vez organizada por Plínio Barbosa e pelo Grupo de Estudos de Prosódia da Fala. Em 2015, foi realizado o 5º Colóquio Brasileiro de Prosódia da Fala, na Universidade de Brasília, por Aveliny Lima, e, em 2016, a IV Escola de Prosódia foi realizada na Universidade Federal de Sergipe”, por Raquel Meister Ko. Freitag e Luciana Lucente.

Como resultado, esta coletânea reúne trabalhos de pesquisa em prosódia e suas aplicações, especialmente para o ensino de línguas. Ao congregar pesquisadores juniores e seniores, contribuímos para o desenvolvimento e consolidação do campo da Prosódia no Brasil.

# Conteúdo

Introdução à análise entoacional .....	7
<i>Luciana Lucente (UFAL)</i>	
O uso do magnetômetro (EMA) na análise de dados articulatórios da prosódia da fala.....	27
<i>Alexsandro Rodrigues Meireles (UFES)</i>	
Aspectos de produção e percepção de estilos de elocução profissionais e não profissionais em quatro línguas .....	43
<i>Plínio A. Barbosa (Unicamp)</i>	
A interface entre prosódia e discurso no inglês como língua estrangeira .....	61
<i>Rodrigo Garcia Rosa (USP)</i>	
O ensino de prosódia nas aulas de espanhol como língua estrangeira (E/LE).....	77
<i>Cristiane Conceição Silva (UFBA)</i>	
Escrita, prosódia e leitura .....	103
<i>Vera Pacheco (UESB)</i>	
Análise variacionista de pistas prosódicas em fronteiras de constituintes .....	117
<i>Raquel Meister Ko. Freitag (UFS), Bruno Felipe Marques Pinheiro (UFS) &amp; Lucas Santos Silva (UFS)</i>	

Prosódia e Fonoaudiologia: do fonoestilo ao transtorno da linguagem .....	133
<i>Aveliny Mantovan Lima (UnB) &amp; Ana Carolina Constantini (Unicamp)</i>	
Contribuições sobre as características prosódicas de interrogativas totais neutras produzidas por sergipanos .....	145
<i>Vanessa Gonzaga Nunes (UFS)</i>	

## Introdução à análise entoacional

Luciana Lucente  
Universidade Federal de Alagoas

### 1 Introdução

Este capítulo, como o próprio título indica, visa funcionar como uma introdução para o iniciante no estudo da entoação da fala.

A entoação é um dos aspectos que compõem a prosódia da fala, e é definida pela “combinação de características tonais em unidades estruturais maiores associadas ao parâmetro acústico da frequência fundamental ( $F_0$ ) e suas variações distintivas no processo da fala” (BOTINIS et al., 2001, p. 264).

Por sua vez, a prosódia compreende aspectos suprasegmentais da fala, como características temporais (duração) e dinâmicas (intensidade e frequência fundamental).

O termo suprasegmental se refere às propriedades da fala que estão sobrepostas aos segmentos, ou seja, que ocorrem em paralelo à sequência dos segmentos que compõem os sons da fala. Sendo assim, no momento em que pronunciamos uma sentença, paralelamente aos segmentos sonoros que a compõem (fones/fonemas) estão os componentes prosódicos, que possibilitam, por exemplo, a acentuação das sílabas tônicas e a expressão comunicativa. A pesquisa prosódica parte da matéria da fala relacionada à atribuição de características pragmáticas e comunicativas, sem deixar de relacionar esta descrição aos aspectos segmentais.

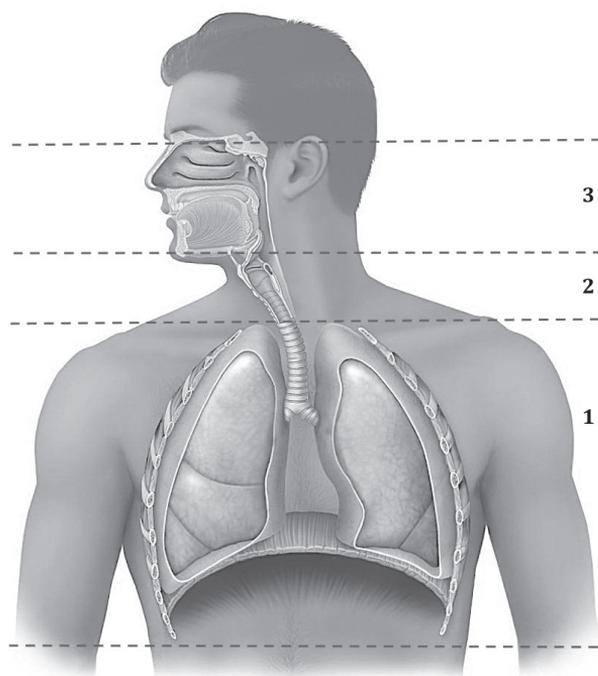
Sendo assim, a descrição prosódica lida com diferentes aspectos da fala em relação aos utilizados na fonética segmental, e esse fator nos leva ao desenvolvimento de diferentes métodos de análise entoacional, como, por exemplo, métodos de notação entoacional.

A seguir são apresentados mais detalhadamente conceitos importantes para o estudo da entoação, como: frequência fundamental, entoação e prosódia.

## 1.1 Frequência fundamental forças aerodinâmicas do sistema respiratório

A frequência fundamental (doravante  $F_0$ ), medida em Hz (Hertz), é definida pelo número de vezes por segundo em que as pregas vocais completam um ciclo de vibração. Esse ciclo é controlado pelos músculos da laringe que determinam a tensão nas pregas vocais, como também por forças aerodinâmicas do sistema respiratório (BOTINIS et al., 2001).

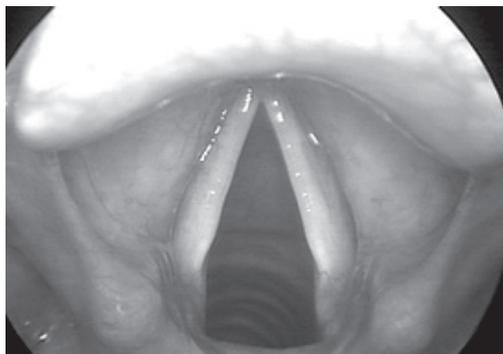
Para entendermos melhor como ocorre a produção da  $F_0$ , é preciso entender que, no sistema fonador, estão envolvidos outros dois subsistemas: o subsistema respiratório (sublaríngeo), composto por caixa torácica, pulmões, traqueia e musculatura de suporte; o subsistema articulatório (supralaríngeo), composto pela faringe, trato nasal e trato vocal; e a laringe (GICK et al., 2013). Como a nomenclatura indica, esses sistemas são interconectados pela laringe – órgão de múltiplas funções que funciona como uma válvula que impede a passagem de ar durante a deglutição e a entrada de partículas nas vias respiratórias, e responsável pela produção da voz. Esses sistemas podem ser vistos na Figura 1 a seguir.



**Figura 1** Sistema fonador e suas subdivisões: 1 – subsistema respiratório; 2 – laringe; 3 – subsistema articulatório.

A laringe é uma estrutura constituída por cartilagens, ligamentos, músculos e membranas, e se encontra ligada a estruturas ósseas por meio de outros músculos,

chamados músculos extrínsecos. Na parte interna da laringe se situam as pregas vocais, que são controladas por músculos intrínsecos, que controlam a abertura e fechamento da glote – que é o espaço entre as pregas vocais – e por ligamentos vocais. São esses ligamentos que efetivamente vibram com a passagem do ar vindo do sistema respiratório pela laringe (BARBOSA; MADUREIRA, 2015). A abertura e o fechamento glotal podem ser vistos nas figuras 2 e 3.



**Figura 2** Pregas vocais abduzidas – glote aberta.

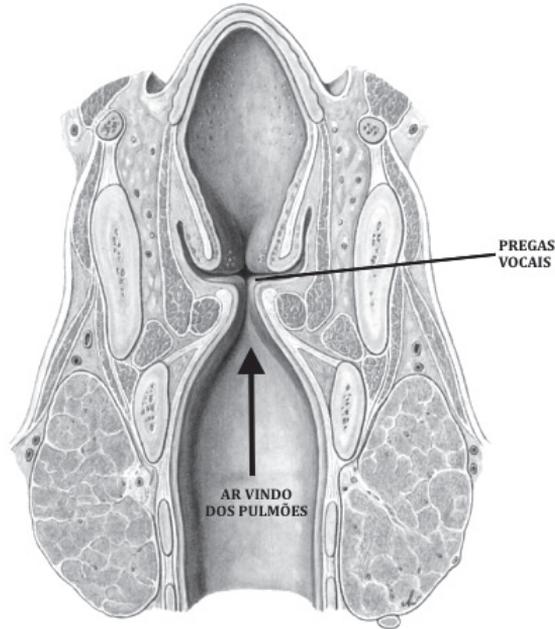


**Figura 3** Pregas vocais aduzidas – glote fechada.

A  $F_0$  da voz é produzida quando o ar vindo do subsistema respiratório passa pela glote fazendo com que as pregas vocais vibrem. Porém para que ocorra a vibração a musculatura da laringe deve agir provocando tensão nas pregas vocais, que funciona como uma barreira que a passagem do ar tem que vencer, como mostra a Figura 4. Ao passar pela glote, o ar vence essa pressão e produz vibração nas pregas vocais.

O sistema de produção da fala trabalha com diferenças de pressão entre os subsistemas que compõem o sistema fonador. Para se ter uma ideia desse funcionamento basta produzir uma vocalização e tentar continuar a vocalização fechando o nariz e a boca. Fazendo esse teste simples é possível perceber que em um determinado momento não é mais possível a vibração das pregas vocais, isso acontece por que ocorre um igualamento das pressões no sistema respiratório e no sistema articulatório.

Portanto a  $F_0$  é resultado do ciclo de vibração das pregas vocais provocado pela passagem do ar na glote. Uma voz feminina com  $F_0$  média de 400Hz indica que a passagem do ar faz as pregas vocais vibrarem 400 vezes em um segundo. Se num determinado momento essa mesma voz passa a vibrar numa frequência de 450Hz significa que a vibração das pregas aumentou, e para que isso ocorra é preciso que a pressão sublaríngea seja maior do que a pressão supralaríngea.



**Figura 4** Visão posterior da laringe; o fluxo de ar vindo dos pulmões tem que vencer a pressão supraglotal e atravessar as pregas vocais.

De acordo com Barbosa e Madureira (2015), é importante para a pesquisa e a análise entoacional que o pesquisador tenha o conhecimento de que o som da fala é resultado de um processo complexo que envolve o sistema respiratório, laríngeo e articulatório.

## 1.2 Prosódia e entoação

Como dito anteriormente, a prosódia engloba, entre outros fenômenos, a entoação. Porém, diferentes definições de prosódia e entoação adotam concepções diferentes destes fenômenos. A primeira definição de entoação aqui abordada se aproxima da definição de prosódia – termo mais amplo que envolve a descrição de características dinâmicas e temporais associadas a aspectos formais. Nesta abordagem, a definição explorada é proposta por Ladd (1996), para quem a entoação “se refere ao uso de características fonéticas suprasegmentais para expressar significados pragmáticos no nível da sentença de forma linguisticamente estruturada” (LADD, 1996, p. 6).

As características suprasegmentais a que se refere Ladd (1996) são:  $F_0$ , intensidade e duração, os correspondentes de ordem física (fonética), dos fenômenos psicofísicos (fonológicos), de *pitch*, volume e quantidade (LEHISTE, 1970).

Esta definição, muito comumente empregada em sistemas entoacionais que seguem a perspectiva da fonologia entoacional (PIERREHUMBERT, 1980; LADD, 1996), considera a entoação como resultante da atuação de características fonológicas subjacentes ligadas a características fonéticas na superfície. Esta relação entre aspectos fonéticos e fonológicos, ou físicos e psicofísicos, também aparece nas definições de outros autores.

Na definição de Botinis et al. (2001), apresentada na introdução deste capítulo, a entoação é definida pela combinação de características tonais em grandes unidades estruturais associadas ao parâmetro acústico da  $F_0$  e suas variações distintas no processo da fala. Nesta definição, ocorre um tipo de relação próxima à exposta por Ladd (1996), mesmo este assumindo uma definição mais ampla, que se aproxima da prosódia.

Em Hirst e Di Cristo (1998), o termo “entoação” é apresentado como portador de duas dicotomias entre sentidos amplos e restritos. A primeira dicotomia reside na definição da entoação em si. Em sentido amplo, a entoação compreende fatores como acento lexical, tom<sup>1</sup> e período, fatores que são excluídos quando a entoação é considerada em um sentido estreito, quando se refere à entoação propriamente dita, que se restringe a características supralexicais.

A segunda dicotomia apontada por Hirst e Di Cristo (1988) está nos níveis de análise e descrição, que podem ser feitos de acordo com níveis físicos ou formais, ou seja, fonético ou fonológico. Esta divisão de termos entre sentidos amplo e restrito faz uma distinção entre lexical versus não-lexical, e linguístico (formal ou fonológico) versus físico, porém, mesmo havendo esta distinção, é assumido de alguma forma que a entoação é definida por esta relação entre componentes funcionais na base e formais na superfície, mesmo que não haja para Hirst e Di Cristo (1998), assim como para Botinis et al. (2001), uma relação teórica entre suas definições e a fonologia entoacional.

Uma quarta definição para o fenômeno da entoação é apresentada por Xu (2005; 2006). Em sua explanação, o autor utiliza o conceito de melodia da fala, fenômeno para o qual convergem características tonais, como padrões de *pitch* relacionados ao acento lexical, e entoação propriamente dita, representada pela  $F_0$ . Este conceito de melodia da fala se aproxima do que propõem Ladd (1996), Hirst e DiCristo (1998) e Botinis et al. (2001) como definição de entoação em sentido amplo, o que, por sua vez, se aproxima da definição de prosódia: a relação entre componentes físicos e formais.

No entanto, o modelo de análise adotado por Xu (2005; 2006) considera a melodia da fala portadora de informações comunicativas que são produzidas

---

1 Pode-se entender por “tom” a diferença entre alturas de *pitch*. Por isso, nos referimos ao aumento de frequência como aumento de tom, tom alto ou agudo, e a frequências baixas como tom baixo ou grave.

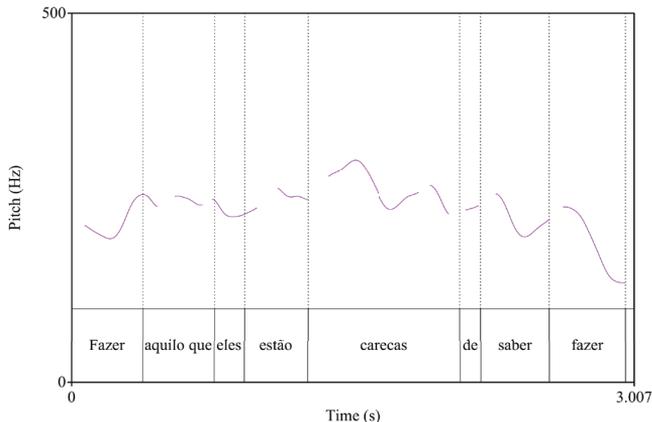
unicamente pelo sistema articulatório e codificadas e transmitidas por meio da  $F_0$ . Nesse sentido, podemos considerar que as funções que se encontravam separadas nas abordagens anteriores entre sistemas físico e psicofísico, se encontram unidas em um único componente articulatório nesta definição. Sendo assim, o que as outras abordagens consideram como entoação em amplo sentido, é definida para Xu dinamicamente apenas em termos da produção de  $F_0$ .

É possível termos definições de entoação de caráter mais estreito e mais amplo. No entanto, a definição de entoação adotada neste capítulo é de caráter estreito e relaciona a entoação da fala diretamente ao parâmetro acústico da  $F_0$ , como propõe Xu (2005; 2006)

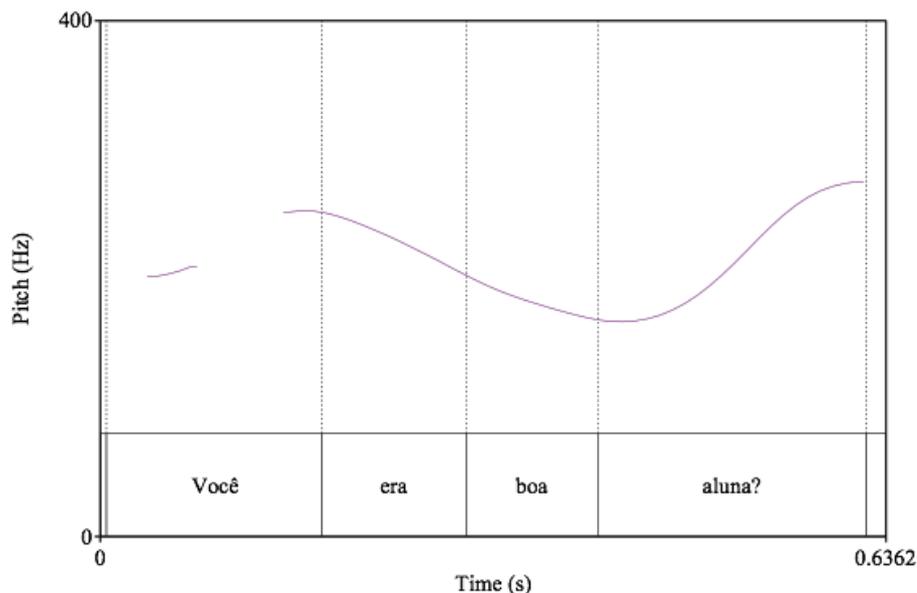
## 2 Análise entoacional

A  $F_0$ , apresenta variações de acordo com o sexo, a idade e a intenção dos falantes. Por exemplo, mulheres e crianças têm  $F_0$  maior em relação aos homens por uma questão anatômica, na qual as menores laringes produzem maiores frequências (KENT; READ, 2015). Quanto à idade dos falantes, percebe-se que pessoas mais velhas apresentam vozes mais graves, e isso se deve ao enfraquecimento dos músculos extrínsecos da laringe, aplicando, conseqüentemente, menor tensão nas pregas vocais (MIFUNE et al., 2007).

A intenção dos falantes se relaciona a mudanças na altura da  $F_0$  ao longo da fala. Por exemplo, para se atribuir ênfase a determinada palavra em um enunciado é preciso aumentar a  $F_0$  durante a articulação desta palavra, como ocorre na palavra “carecas” na Figura 5. Para produzir um enunciado interrogativo direto (em que partículas “qu” estão ausentes), é preciso um aumento da  $F_0$  alinhado ao final do enunciado, como ocorre na palavra “aluna” na Figura 6.



**Figura 5** Curva entoacional do enunciado “Fazer aquilo que eles estão carecas de saber fazer”; palavra “carecas” com  $F_0$  média de 264Hz.



**Figura 6** Curva entoacional do enunciado “Você era boa aluna? “; palavra “aluna” com  $F_0$  média de 217Hz.

Pelo fato da  $F_0$  apresentar movimentos distintos, ascendentes e descendentes, conforme podemos ver nas figuras 5 e 6, surgiu a necessidade de notações que representem esses fenômenos fonéticos. Por esta razão são propostos diversos sistemas de notação entoacional, cada qual com características próprias, como os sistemas ToBI (SILVERMAN et al., 1992; BECKMAN et al., 1994), Intsint (HIRST e DICRISTO, 1988), IPO (‘t HART, 1990), DaTo (LUCENTE, 2008; 2012). Esta representação dos contornos entoacionais por meio de sistemas de notação é conhecida como fonologia entoacional.

## 2.1 Fonologia entoacional

A fonologia entoacional surgiu da necessidade de representação dos fenômenos fonéticos entoacionais. Segundo Ladd (1996), a fonologia entoacional procura caracterizar os movimentos da  $F_0$  em termos de sequências de elementos categoricamente distintos; para tanto, são empregados sistemas ou formas de notação entoacional capazes de descrever tais elementos. Apresentamos a seguir como duas diferentes perspectivas teóricas interpretam os fenômenos entoacionais e a forma de representação que propõem.

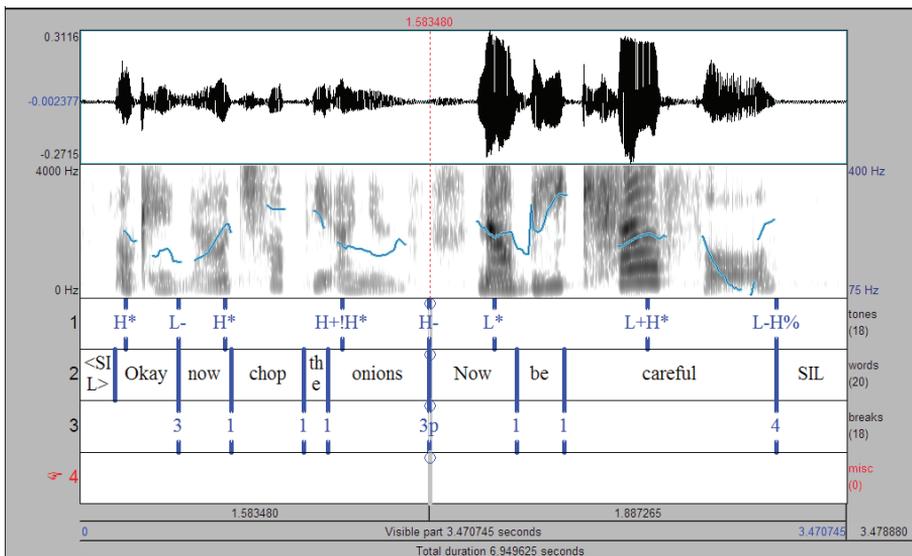
A primeira perspectiva segue os pressupostos da teoria métrica-autossegmental (teoria MA), representada no sistema ToBI de notação entoacional (BECKMAN et al, 1994; SILVERMAN et al, 1992); a segunda, uma perspectiva dinâmica de

análise da entoação, representada no sistema DaTo de notação entoacional (LUCENTE, 2008; 2012).

De acordo com Ladd (1996), a teoria MA ao tentar caracterizar adequadamente os elementos da estrutura tonal das línguas, leva em consideração fundamentos como: i) a linearidade da estrutura tonal; ii) a distinção entre proeminência (*pitch accent*) e acento lexical (*stress*); iii) a análise dos *pitch accents* em relação à altura dos tons; iv) e a relação entre características locais e tendências gerais.

Partindo da hipótese da linearidade da estrutura tonal, a teoria MA representa fonologicamente os contornos entoacionais como sequências de eventos entoacionais discretos (LADD, 1996). Sendo assim, a teoria assume claramente a distinção entre eventos e transições que ocorrem na estrutura tonal, reconhecendo partes da estrutura como linguisticamente importantes e outras como apenas transições preenchedoras da  $F_0$  entre eventos localizados. Os eventos considerados importantes do ponto de vista linguístico são os *pitch accents*, entendidos nesta teoria como eventos locais que assinalam proeminência no nível do enunciado, e associados à altura da curva entoacional nos pontos em que esses eventos ocorrem.

Apesar da teoria MA considerar os *pitch accents* como os fenômenos mais importantes da curva entoacional, estes são analisados à luz da teoria apenas como sequências ou combinações de tons altos e baixos, sendo representados por meio dos rótulos H (*high*) e L (*low*), e pela combinação destes, como se observa na Figura 7.



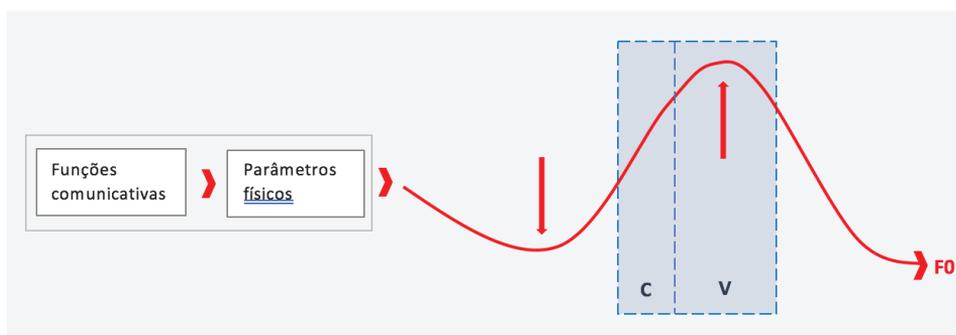
**Figura 7** Exemplo de notação do sistema ToBI utilizando o *software* Praat.

Seguindo esses pressupostos, o sistema ToBI (BECKMAN et al., 1994) apresenta uma notação que reflete a linearidade da estrutura tonal e a marcação dos *pitch accents*. Tal sistema foi desenvolvido por um grupo de pesquisadores com o intuito inicial de produzir uma ferramenta de transcrição que pudesse auxiliar no entendimento e no desenvolvimento de modelos computacionais de síntese e reconhecimento de fala que requerem grande quantidade de enunciados transcritos prosodicamente (SILVERMAN et al., 1992).

### 3 Sistema DaTo

Enquanto a teoria MA, representada pelo sistema ToBI, se baseia em aspectos lineares da estrutura tonal, na identificação dos *pitch accents* e no alinhamento abstrato dessa estrutura com o material linguístico, o sistema DaTo de notação entoacional (LUCENTE, 2008; 2012) se concentra na convergência de aspectos fonéticos – velocidade, intensidade, altura, duração – da curva entoacional a fim de atingir um alvo ou desempenhar uma tarefa linguística por meio dos contornos entoacionais, da gama de variação tonal e do alinhamento específico com o material linguístico.

O sistema DaTo foi desenvolvido com base na entoação do português brasileiro (PB) e trabalha com o conceito de contorno dinâmico, que é definido em Lucente (2012, p. 99) como “uma unidade tonal que contém elementos comunicativos expressos em uma trajetória ideal da curva entoacional, especificada por um alvo a ser atingido e associada a uma unidade segmental linguística”. O esquema ilustrativo deste sistema é apresentado na Figura 8.



**Figura 8** Em um contorno entoacional, as (i) funções comunicativas são codificadas em parâmetros físicos expressos em uma (ii) trajetória ideal de  $F_0$  com um alvo a ser atingido (setas); e (iii) alinhado a uma unidade segmental (sílabas tônicas).

Cada contorno dinâmico, marcador de proeminência, tem movimento e altura relacionados à altura e ao movimento de  $F_0$  que o antecede e o segue, e sua posição na estrutura tonal é determinada pelo alinhamento com o material

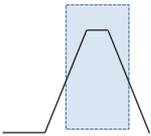
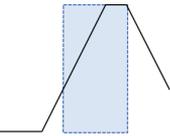
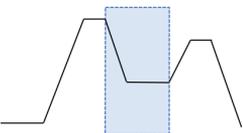
linguístico, mais precisamente com as sílabas tônicas, ou seja, com o acento. O contorno dinâmico na notação DaTo não se refere à adição de movimentos com diferentes alturas de  $F_0$ , mas a um movimento da curva entoacional ao qual se atribui marcação de proeminência.

O sistema DaTo, além de apresentar uma notação dos fenômenos entoacionais do PB a partir de uma perspectiva dinamicista, procura também apresentar informações pragmático-comunicativas e aspectos fonéticos do ritmo da língua em suas camadas de notação.

Os contornos dinâmicos, como definido anteriormente, são movimentos contínuos de  $F_0$ , que percorrem a transição de um ponto a outro na curva entoacional até atingir o seu alvo. Este movimento de transição carrega a informação tonal que antecede e segue seu alvo. Nessa concepção, um contorno dinâmico torna relevantes as transições até então consideradas pela teoria MA como apenas transições preenchedoras entre eventos estáticos.

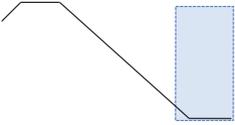
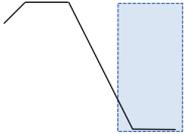
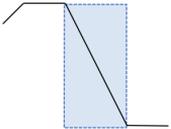
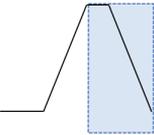
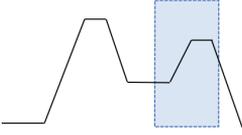
No sistema DaTo, os contornos dinâmicos se dividem entre ascendentes, descendentes e níveis de fronteira, e são representados por rótulos que ilustram o movimento de  $F_0$  pelas letras H (high) e L (low).

O conjunto de contornos entoacionais do sistema DaTo são anotados levando em consideração seu padrão de movimento e alinhamento com as sílabas tônicas partindo primeiramente da percepção de ênfase e posteriormente de uma inspeção visual da curva entoacional que determinará o rótulo a ser empregado. A Figura 9 indica o rótulo empregado para notação de acordo com o movimento da curva entoacional e seu alinhamento com a vogal tônica.

Contornos dinâmicos		
Rising	LH	
Late rising	>LH	
Compressed rising	vLH	

(continua)

**Tabela 1** rótulos do sistema DaTo de acordo com padrão de movimento e alinhamento com a vogal tônica. (continuação)

Contornos dinâmicos		
Rising falling	LHL	
Falling	HL	
Late falling	>HL	
Compressed falling	vHL	
Falling rising	HLH	
Níveis de fronteira		
Low	L	
High	H	

**Figura 9** Rótulos do sistema DaTo de acordo com padrão de movimento e alinhamento com a vogal tônica, ilustrada pelo retângulo azul.

## 4 Análise entoacional

Para exemplificar a análise entoacional com o uso da notação entoacional do sistema DaTo, foram usadas gravações de leituras em voz alta de dois sujeitos, uma mulher na faixa de 30 a 35 anos e um homem na faixa de 25 a 30 anos,

ambos pós-graduandos em linguística na época da gravação, e com conhecimento em fonética.

As orações gravadas pelos sujeitos são parte de uma história popular que circula na internet, na qual se diz que um homem muito rico morreu e deixou um bilhete escrito sem nenhuma pontuação, como em (1). Como o conteúdo do bilhete estava sem pontuação, cada pessoa que lia interpretava da forma mais conveniente. Sendo assim, as pontuações das orações de (2) a (5) teriam sido feitas, respectivamente, por sua irmã, seu sobrinho, o padeiro e pelos pobres da cidade. O resultado do emprego das diferentes pontuações resulta na mudança de sentido e na reestruturação das orações em sintagmas distintos, nisso reside o humor da história, no entanto na leitura das orações em voz alta é a mudança na entoação da fala que produz os diferentes sentidos.

- (1) “Deixo meus bens à minha irmã não a meu sobrinho jamais será paga a conta do padeiro nada dou aos pobres”
- (2) “Deixo meus bens à minha irmã, não a meu sobrinho. Jamais será paga a conta do padeiro. Nada dou aos pobres.”
- (3) “Deixo meus bens à minha irmã? Não! A meu sobrinho. Jamais será paga a conta do padeiro. Nada dou aos pobres.”
- (4) “Deixo meus bens à minha irmã? Não! A meu sobrinho? Jamais! Será paga a conta do padeiro. Nada dou aos pobres.”
- (5) “Deixo meus bens à minha irmã? Não! A meu sobrinho? Jamais! Será paga a conta do padeiro? Nada! Dou aos pobres”.

Os exemplos de leitura gravados são bastante úteis para uma análise da relação entre forma e função do contorno entoacional, uma vez que exemplos idênticos do ponto de vista segmental são produzidos ora com entoação declarativa, ora com entoação interrogativa. Por exemplo, em (2) temos “Deixo meus bens a minha irmã, não a meu sobrinho”, e em (3) “Deixo meus bens à minha irmã? Não! A meu sobrinho”.

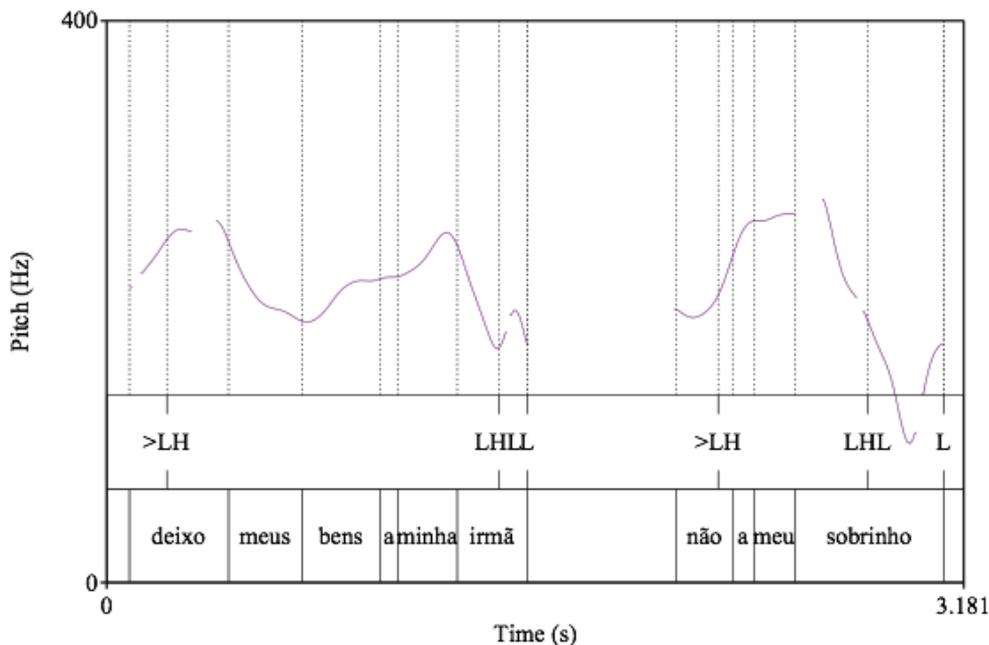
Entoacionalmente, a estrutura dos sintagmas se apresenta da seguinte forma: i) no final da primeira oração, quando esta é declarativa, a notação corresponde a um movimento final descendente da  $F_0$ , notado com o contorno LHL e com o nível de fronteira L, como em (6); ii) quando esta passa a ser interrogativa, a notação do movimento final fica >LH, com nível de fronteira H, como em (7); iii) a oração subordinada “não a meu sobrinho” em (6’) apresenta o movimento ascendente inicial atrasado, notado por >LH (padrão no início de enunciados após pausa), e o movimento descendente notado por LHL e L na fronteira final; iii) quando esta oração se divide entre os sintagmas “Não” e “A meu sobrinho”, como em (7’) e (7’), a partícula negativa ganha status de oração independente e é notada vLH e L, enquanto “A meu sobrinho” recebe a notação >LH, LHL e L, muito comumente encontrada em declarativas (LUCENTE, 2012); iv) de forma

similar ao que ocorre em (7) com “Deixo meus bens à minha irmã?”, em (8”) a oração “A meu sobrinho” assume a modalidade interrogativa e recebe notação >LH, >LH e H.

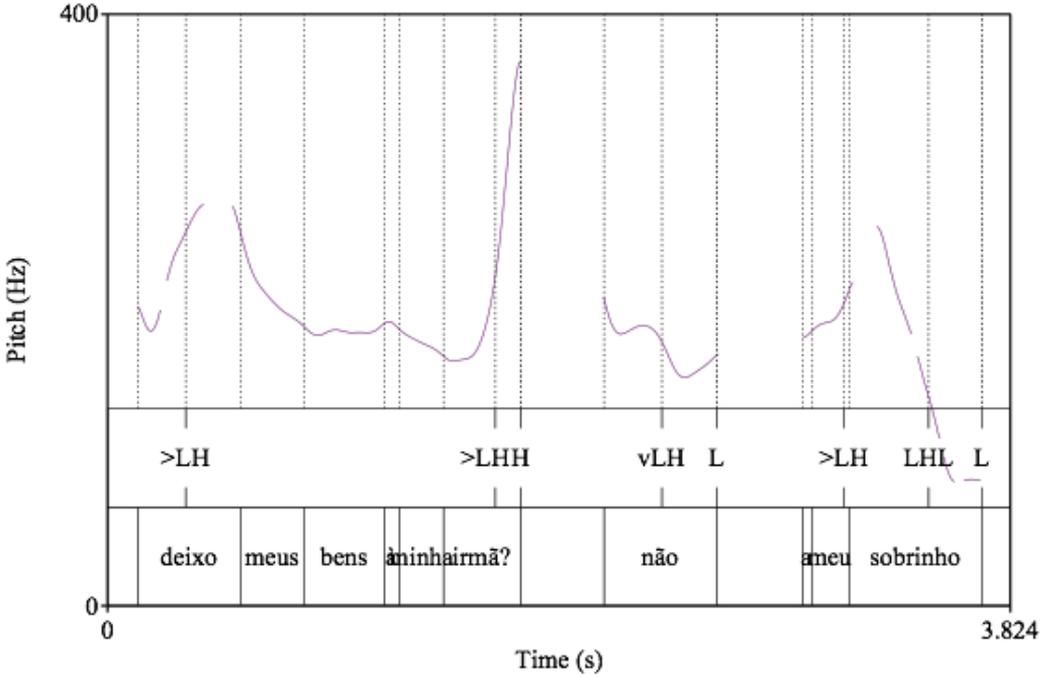
- (6) Deixo meus bens à minha irmã, (6’) não a meu sobrinho.  
 >LH LHL L >LH LHL L  
 (7) Deixo meus bens à minha irmã? (7’) Não. (7”) A meu sobrinho.  
 >LH >LH H vLH L >LH LHL L  
 (8) Deixo meus bens à minha irmã? (8’) Não (8”) A meu sobrinho?  
 >LH >LH H vLH L >LH >LH H

Observando melhor a partícula negativa, sua estrutura apresenta a seguinte forma: i) quando esta está inserida na oração, como em (6’), recebe a notação padrão de início de enunciados, que é >LH, em que toda a extensão da subida da F<sub>0</sub> está alinhada à vogal tônica, no caso o [ãw]; ii) quando “Não” é uma oração independente, como em (7’) e (8’), assume outro contorno de F<sub>0</sub>, no qual a vogal tônica está em movimento descendente, notado por vLH, seguido de uma fronteira intermediária, que resulta na notação do nível L. Esse pequeno exemplo mostra como a entoação desempenha papel fundamental na atribuição de sentido à fala.

As figuras 10 e 11 ilustram o movimento da curva de F<sub>0</sub> e o emprego dos rótulos do sistema DaTo para a notação destes exemplos.



**Figura 10** Curva de F<sub>0</sub> e emprego da notação DaTo para as orações (6) e (6’).

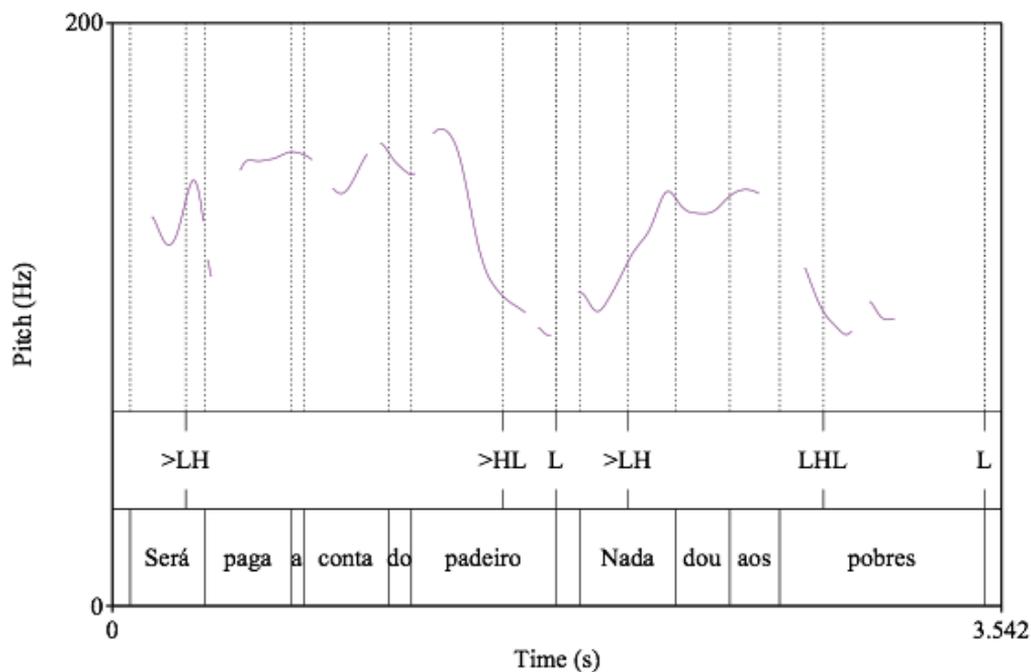


**Figura 11** Curva de  $F_0$  e emprego da notação DaTo para as orações (7), (7') e (6").

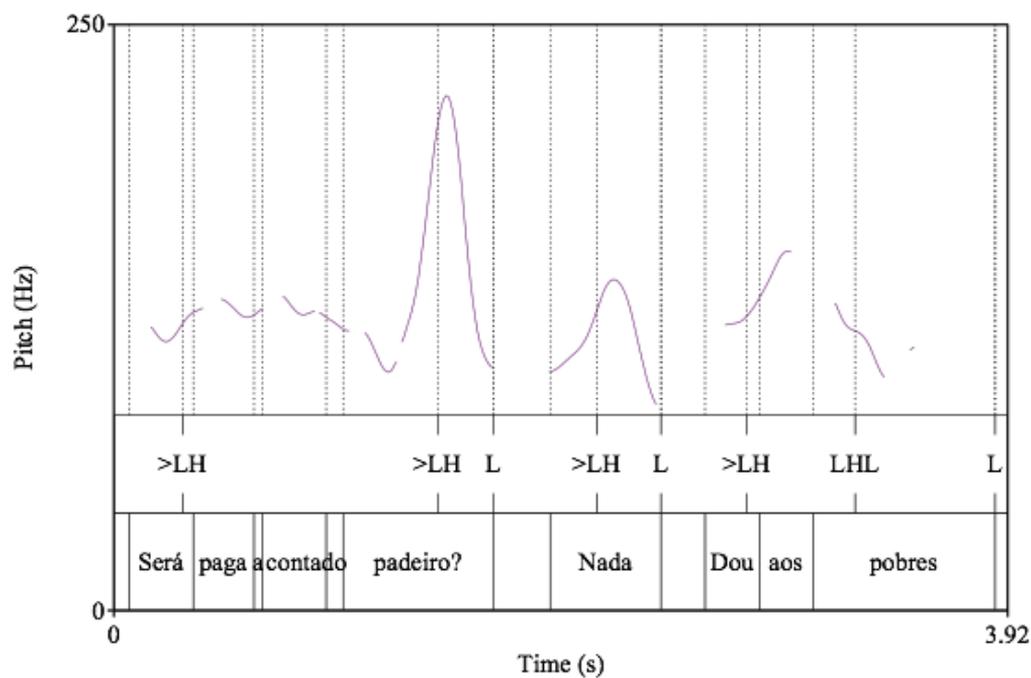
O comportamento entoacional das outras orações da gravação se apresenta de forma bastante similar aos exemplos já observados: i) em (10) a oração “Será paga a conta do padeiro?”, que em (9) apresenta os contornos >LH, >HL e L, assume a modalidade interrogativa, passando a apresentar os contornos >LH, >LH e L; ii) a separação da oração independente “Nada dou aos pobres” apresentada em (9') nos dois sintagmas independentes “Nada” e “Dou aos pobres”, como em (10') e (10"), seguem as mesmas características dos exemplos (6'), (7') e (7"), em que o padrão declarativo >LH, LHL e L se desmembra em >LH e L em (10') e >LH, LHL e L em (10") reestruturando um padrão entoacional declarativo em uma oração menor, como também ocorre em (7"). As figuras 12 e 123ilustram o movimento da curva de  $F_0$  e o emprego dos rótulos do sistema DaTo para a notação destes exemplos.

(9) Será paga a conta do padeiro. (9') Nada dou aos pobres.  
 >LH >HL L >LH LHL L

(10) Será paga a conta do padeiro? (10') Nada! (10") Dou aos pobres.  
 >LH >LH L >LH L >LH LHL L



**Figura 12** Curva de  $F_0$  e emprego da notação DaTo para as orações (9) e (9').



**Figura 13** Curva de  $F_0$  e emprego da notação DaTo para as orações (10), (10') e (10'').

Os exemplos apresentados mostram algumas diferenças quanto à notação entoacional; em alguns casos, os contornos finais das sentenças declarativas podem ser notados por LHL, em outros casos, por >HL. Como descrito em Lucente (2012), ambos rótulos servem para a notação de contorno descendentes, porém LHL é reservado para descidas espacialmente mais longas e graduais da  $F_0$ , enquanto os rótulos HL e >HL representam descidas mais acentuadas e mais curtas. Em termos funcionais, essa diferença pode denotar uma ação concluída em oposição a uma ação a ser concluída: nos exemplos aqui apresentados o uso de >HL se restringe à oração declarativa “Será paga a conta do padeiro”, que assume a função de ordem, enquanto em “Deixo meus bens à minha irmã” e “Nada dou aos pobres” essa função é ausente.

Nas modalidades interrogativas, em (7) e (10), a marcação da função interrogativa está associada aos rótulos >LH e H e >LH e L, respectivamente. Essas orações, que são perguntas retóricas, são seguidas por negativas em (7') e (10'), que apresentam os padrões vLH e L e >LH e L, como se pode ver nas figuras 11 e 13. Essa diferença nos padrões dos níveis de fronteira e no início dos sintagmas negativos pode ser atribuída a diferença de acentuação das partículas negativas, pois, enquanto “não” é um monossílabo tônico, “nada” é uma paroxítone. Em “não”, não há tempo para a realização de um contorno ascendente, pois a negação tem que realizar uma descida obrigatória no final do enunciado declarativo, como se pode verificar na Figura 11.

Esse fator pode explicar a ocorrência de um nível de fronteira alto (H) na interrogativa que o antecede. Dessa forma o movimento de descida da  $F_0$  que se observa no nível de fronteira L em (10') está ocorrendo durante o sintagma “Não” em (7'). Em contrapartida, como pode ser observado na Figura 13, em (10) temos um nível de fronteira baixo (L) que possibilita a realização do contorno >LH alinhado à sílaba tônica de “nada”, como é previsto em início de enunciados declarativos no PB.

## 5 Forma e função dos contornos entoacionais

A análise apresentada, além de servir como método introdutório à análise entoacional, mostra a relação existente entre forma e função dos contornos entoacionais. O sistema DaTo (LUCENTE, 2008; 2012) de notação trabalha com a forma dos contornos entoacionais, estabelecendo rótulos de notação de acordo com os padrões de movimento da  $F_0$  e o alinhamento destes com as sílabas tônicas, como já se apresentou na Figura 9. Esses contornos se apresentam bastante consistentes no que podemos chamar de tonotaxe do PB, ou seja, eles ocorrem com frequência, seguem uma ordem de combinações e estão relacionados a funções específicas na fala.

A afirmação de que os contornos entoacionais do PB seguem uma ordem de combinações se baseia em fatos como: i) a frequente ocorrência do contorno >LH em início de enunciados, sendo que LH também ocorre nesta posição em frequência bem menor (LUCENTE, 2015); ii) ao contorno LH se reserva a marcação de proeminências no nível do enunciado, assim como o contorno HLH, que ocorre alinhado a palavras paroxítonas com mais de três sílabas, sempre tendo o primeiro pico alinhado à primeira sílaba da palavra e o segundo à sílaba tônica (LUCENTE, 2012); iii) os rótulos LHL, HL e >HL se reservam à marcação de final de enunciado, precedendo fronteiras de nível L, sendo que, como dito anteriormente, LHL descreve uma descida mais suave e prolongada da  $F_0$ ; iv) o contorno vLH tem como propriedade descrever movimentos da  $F_0$  em que não há espaço para a realização de um pico como em LH, como ocorre no exemplo (7'), por isso é chamado de contorno ascendente comprimido; v) de forma similar, o contorno vHL também descreve um contorno comprimido, porém descendente, pois nos casos em que ele se aplica restrições espaço-temporais impedem que haja uma descida da  $F_0$  até o nível que HL alcança, por exemplo (LUCENTE, 2012).

Essa descrição de elementos formais da entoação caminha em conjunto com aspectos funcionais da entoação, porém, tais aspectos são atribuídos às essas formas considerando diversas combinações entre estas. O exemplo (9) ilustra uma possível relação entre o contorno HL e o fato desta oração apresentar uma ordem.

São notórios os estudos que atribuem aspectos pragmáticos e comunicativos à entoação da fala. Em Ward e Hirschberg (1995), são tratados aspectos de implicaturas de incerteza relacionados à entoação descendente-ascendente no inglês americano. Neste trabalho, os autores utilizaram como base as implicaturas conversacionais de Grice (1975) para analisar as relações entre forma e função dos contornos entoacionais, metodologia similar pode ser vista em Hirschberg (2002). Em Moraes e Colamarco (2007), são analisados padrões entoacionais de pedidos e perguntas no PB, considerando-os como atos ilocutórios diretivos, de acordo com Searle (1995). Estes são apenas alguns exemplos de trabalhos que tratam a entoação de um ponto de vista formal-funcional.

Os aspectos entoacionais abordados neste capítulo contribuem para uma introdução à metodologia de análise entoacional, servindo de modelo para futuras análises, ou mesmo para o estudo teórico da entoação.

## Referências

BARBOSA, P. A., MADUREIRA, S., *Manuel de fonética acústica experimental: aplicação a dados do português*. São Paulo: Cortêz, 2015.

- BECKMAN, M. E., HIRSCHBERG, J., PITRELLI, J. F. *Evaluation of Prosodic Transcription Labeling Reliability in the ToBI Framework*. Disponível em [http://www.ling.ohio-state.edu/~tobi/ame\\_tobi](http://www.ling.ohio-state.edu/~tobi/ame_tobi), 1994.
- BOTINIS, A., GRANSTROM, B., MOBIUS, B. Developments and paradigms in intonation research. *Speech Communication* 33, 2001. p. 263-296.
- GICK, B., WILSON, I., DERRICK, D. *Articulatory Phonetics*. Wiley-Blackwell, 2013.
- GRICE, H., P. *Logic and conversation*. Syntax and Semantics, vol 3. New York: The Academic Press, 1975.
- HIRSCHBERG, J. The pragmatics of intonational meaning. *Speech Prosody 2002*. Aix-en-Provence, 2002, p.11-13
- HIRST, D. (2005) Form and Function in the Representation of Speech Prosody. *Speech Communication*, 46, p. 334-347, 2005.
- HIRST, D., DI CRISTO, A. *Intonational Systems, a survey of twenty languages*. Cambridge: Cambridge University Press, 1998.
- KENT, D. R., READ, C. *Análise acústica da fala*. São Paulo: Cortez, 2015.
- LADD, D. R. *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- LEHISTE, I. *Suprasegmentals*. Cambridge: MIT Press, 1970.
- LUCENTE, L. Funções comunicativas e restrições articulatórias na determinação do alinhamento de contornos dinâmicos. *Diadorim*, n. 17, vol, p. 52-66, 2015.
- LUCENTE, L. *Aspectos dinâmicos da fala e da entoação do português brasileiro*. Tese de Doutorado. Unicamp, 2012.
- LUCENTE, L. *DaTo: Um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea*. Dissertação de Mestrado. Unicamp, 2008.

MIFUNE E., JUSTINO VSS., CAMARGO Z., GREGIO F. Análise de  $f_0$  da voz do idoso. *Rev CEFAC*, São Paulo, v.9, n.2, 238-47, 2007.

MORAES, J., A., COLAMARCO, M. Você está pedindo ou perguntando? Uma análise entonacional de pedidos e perguntas no português do Brasil. *Revista de Estudos Linguísticos*, v. 15, n. 2, p. 113-126, 2007.

PIERREHUMBERT, J. *The Phonology and Phonetics of English Intonation*. Ph.D Thesis, MIT, 1980.

SILVERMAN, K., BECKMAN, M., PITRELLI, J., OSTENDORF, M., PIERREHUMBERT, J., HIRSCHBERG, J., PRICE, P. TOBI: A Standard Scheme for Labeling Prosody. *Proceedings of the International Conference on Spoken Language 92*, Banff, 1992.

SEARLE, J. *The Philosophy of Language*. Oxford: Oxford University Press, 1995.

HART, J., COLLIER, R., COHEN, A. *A perceptual study of intonation: an experimental-phonetic approach*. Cambridge: Cambridge University Press, 1990.

WARD, G., HIRSCHBERG, J., Implicating uncertainty: the pragmatics of fall-rise intonation. *Language*, vol. 61, n. 4, p. 747-776, 1985.

XU, Y. Speech melody as articulatory implemented communicative functions. *Speech Communication* 46, 2005. p. 220-251.

XU, Y. Speech prosody as articulated communicative functions. *Proceedings of Speech Prosody 2006*, Dresden, Germany, 2006.



## O uso do magnetômetro (EMA) na análise de dados articulatórios da prosódia da fala

Alexsandro Rodrigues Meireles  
Universidade Federal do Espírito Santo

### 1 O magnetômetro (EMA1)

O objetivo deste capítulo é explicar didaticamente como o magnetômetro (ou articulógrafo) funciona para aquisição de dados articulatórios em prosódia da fala. Para uma informação técnica sucinta de como o magnetômetro funciona, vejamos o resumo abaixo de Marino (2002):

Esta técnica, conhecida como articulografia eletromagnética (eletromagnetic articulography EMA), utiliza dispositivos de campo magnético alternativo que permitem monitorar movimentos de pontos onde os sensores se encontram fixados nos articuladores em relação ao tempo. Dentre os sistemas comercialmente disponíveis no mercado têm-se o Articulógrafo AG 100 (Cartens, Alemanha). Este sistema é constituído por três bobinas geradoras que são fixadas em um capacete especialmente construído para esse fim. Quando dispostas em posição fixa ao longo do plano médiosagital, cada uma das bobinas gera seu próprio campo magnético alternativo de alta frequência. Um conjunto de pequenas bobinas receptoras são aderidas temporariamente nos articuladores de interesse (lábios, língua, mandíbula, e/ou véu palatino) por meio de uma cola especial. Como em um transformador, os campos magnéticos alternativos das bobinas geradoras induzem tensão elétrica nas bobinas receptoras. A tensão elétrica induzida em cada uma das

---

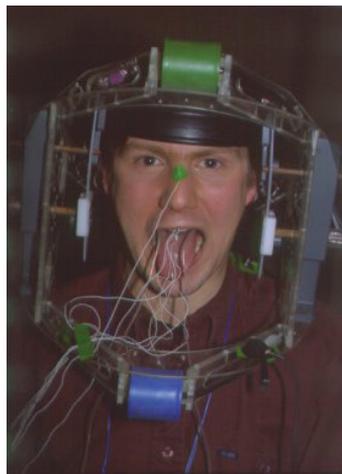
1 Também encontramos na literatura a sigla EMMA (*eletromagnetic midsagittal articulo-meter*, Perkell et al., 1992). No entanto, a sigla mais comum encontrada atualmente nos artigos é EMA (*eletromagnetic articulometer*).

bobinas receptoras é inversamente proporcional ao cubo de sua distância em relação às bobinas geradoras. Tais bobinas operam em frequências diferentes e geram campos magnéticos que se sobrepõem, portanto, o sistema é capaz de captar, simultaneamente, a distância entre cada bobina receptora e as bobinas geradoras (BAKEN; ORLIKOFF, 2000).

O articulógrafo foi desenvolvido inicialmente pela empresa alemã Cartens que produziu o Articulógrafo AG 100 citado por Marino (2002) e, posteriormente, os Articulógrafos 2-D AG 200 e 3-D AG 500. Recentemente, desenvolveram ainda os Articulógrafos 3-D AG501 e 3-D AG501 *twin*, os quais podem extrair informações articulatórias em três dimensões. Informações detalhadas do produto podem ser obtidas no seguinte website: <http://www.articulograph.de/>.

Devido ao seu custo e aplicação específica no mercado, somente encontramos esse equipamento em algumas universidades ao redor do mundo. Mesmo nos Estados Unidos, encontramos esse equipamento em poucas universidades, tais como: University of Southern Califórnia (USC), University of California at Los Angeles (UCLA), University of Connecticut, Haskins Laboratories, University of Texas at Dallas. Para outras universidades do mundo que utilizam o sistema da Cartens, visitar [http://www.articulograph.de/?page\\_id=26](http://www.articulograph.de/?page_id=26). No Brasil, não há nenhuma universidade que utilize o equipamento dessa empresa.

Ao realizar meu doutorado-sanduíche na USC em 2004, tive a oportunidade de ser um dos primeiros brasileiros a ter contato com essa técnica articulatória de análise. A Figura 1 é uma foto de um articulógrafo do mesmo modelo usado por mim nessa universidade, com vários sensores colocados na língua, nariz, e mandíbula. Essa foto foi obtida no endereço [www.phon.ox.ac.uk/~paula/emahome.html](http://www.phon.ox.ac.uk/~paula/emahome.html).



**Figura 1** Exemplo de um Articulógrafo 2-D AG-200 utilizado em meus experimentos na USC.

A partir de 2009, a empresa Northern Digital Inc. (NDI, <http://www.ndigital.com>), do Canadá, em parceria com os foneticistas da USC, desenvolveu um sistema de articulografia eletromagnética (EMA) tão eficiente quanto o da Carstens, porém bem mais barato do que o da empresa alemã. Além disso, o sistema era bem mais simples de se utilizar, não era necessário calibragem do equipamento e dispensava o uso do capacete enorme mostrado na Figura 1.

Também em 2009, pelo edital Pró-Equipamentos CAPES 2009, recebemos verba para adquirir um articulógrafo e, sob aconselhamento da Dra. Dani Byrd, da USC, compramos o equipamento da NDI WAVE. Esse equipamento encontra em pleno funcionamento no Laboratório de Fonética do Programa de Pós-Graduação em Linguística da Universidade Federal do Espírito Santo (PPGEL-UFES) e está disponível para utilização a todos os membros da comunidade acadêmica do Brasil, segundo normas de utilização multiusuário previstas no projeto da CAPES. Até a presente data, o equipamento já foi utilizado por pesquisadores da UNICAMP e da UFPA. Descrevemos, a seguir, as especificações técnicas do equipamento.

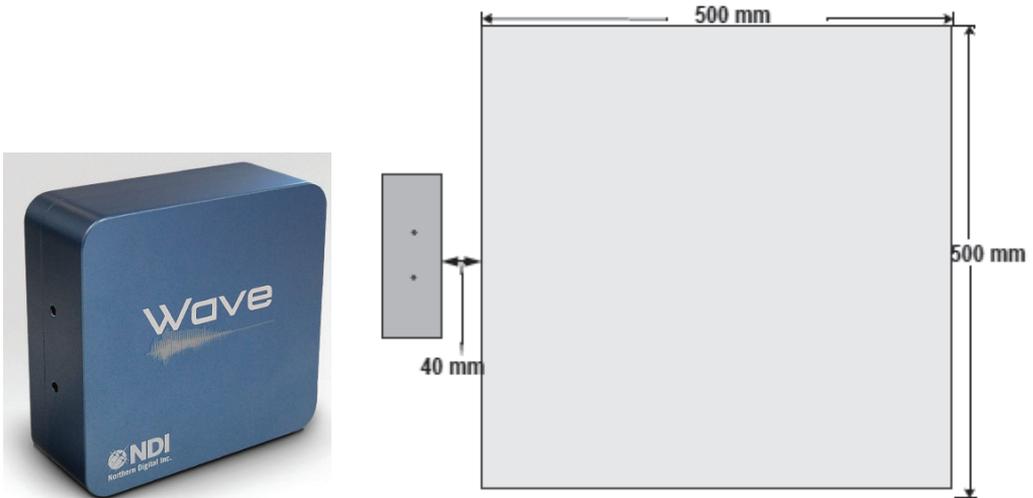
## 2 O articulógrafo da NDI WAVE

O equipamento WAVE da empresa canadense NDI funciona semelhantemente ao equipamento da Carstens, no entanto, com um custo-benefício melhor. De acordo com informações constantes no site da empresa e em seu manual de utilização (<http://www.ndigital.com/msci/products/wave-speech-research/>), o articulógrafo eletromagnético WAVE mede o posicionamento tridimensional de microsensores (Figura 2) fixados na língua, palato, lábios e face para rastreamento de movimentos orofaciais. As medidas são capturadas em tempo real, sem a necessidade de outros aparatos tecnológicos. Caso o laboratório possua dois geradores de campo (*field generator*), há ainda a possibilidade de se gravar um diálogo. As altas precisão temporal e resolução espacial permitem que os movimentos mais rápidos e mais sutis sejam gravados. A resolução temporal dos equipamentos vendidos vai de 100 Hz (nosso equipamento) até 400 Hz. Ressalto que 100 Hz já é uma boa resolução para captura de dados articulatórios, possibilitando até a captura de sons do português como a vibrante simples [r].



**Figura 2** Exemplo de um microsensor utilizado na captura de dados articulatórios.

Segundo o manual do WAVE, o gerador de campo produz campos eletromagnéticos que cobrem um volume fixo para medições de dados articutórios, correspondente a um cubo de 500 mm. O volume é projetado para fora da parte frontal do gerador de campo, a partir de 40 mm do gerador de campo, conforme pode ser visto na Figura 3.



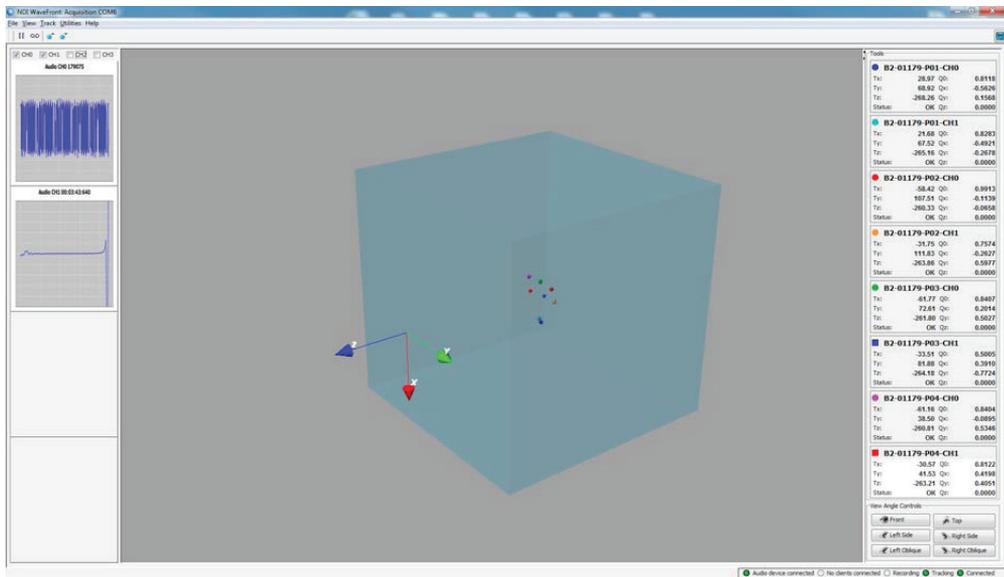
**Figura 3** Gerador de campo WAVE (esquerda) e campo magnético (direita).

É possível utilizar até 16 sensores com um único gerador de campo. No entanto, quanto mais sensores utilizamos na captura de dados, mais desconfortável é para o falante produzir os enunciados, devido à quantidade de fios que saem de sua boca. Por exemplo, em um de nossos experimentos, utilizamos 4 sensores dentro da boca (ponta de língua, corpo de língua, dorso de língua, dentes incisivos inferiores), 2 sensores de referência (orelha direita, orelha esquerda) e 2 sensores nos lábios (lábio superior, lábio inferior). Com isso, temos muitos fios grudados na boca do informante, semelhantemente à configuração de fios vista na Figura 1.

O programa WaveFront da NDI é utilizado para a captura simultânea de dados articutórios e acústicos e também para *playback* (Figura 4). A janela principal mostra o cubo virtual de 500 mm x 500 mm. A localização e o *status* de cada sensor conectado são apresentados em tempo real com cores e formatos diferentes. Cada sensor é identificado com sua porta e canal específicos no sistema, bem como suas coordenadas x, y e z, relacionadas à origem do sistema de coordenadas ou à referência global em que foram configuradas. Os dados articutórios são salvos com a extensão *.wco* (*wave collection*, i.e., coleção de dados necessários para o *playback* dos dados articutórios) e *.wav*. Também é possível especificar a saída de dados com um terceiro grupo de dados com a

O uso do magnetômetro (EMA) na análise de dados articulatórios da prosódia da fala

extensão .tsv (planilha de dados com informações de movimentação espacial dos sensores). Esse arquivo .tsv é utilizado pelo programa MAVIS para a análise quantitativa dos dados articulatórios.



**Figura 4** Exemplo da janela de captura do programa WaveFront com seu eixo tridimensional e, nesse caso, oito sensores sendo gravados.

### 3 Análises articulatório-quantitativas no programa MAVIS<sup>2</sup>

Os dados articulatórios obtidos com o uso do magnetômetro não estão prontos para análise. Para isso, deveríamos convertê-los, através de rotinas no programa Matlab ([www.mathworks.com](http://www.mathworks.com)). Antes desta conversão, porém, a fim de facilitar nossa análise, elaboramos ainda um script da linguagem de programação Perl (apêndice A.4), o qual renomeia os arquivos brutos originais. Após este processo, obtivemos os arquivos prontos para análise. A seguir, com o intuito de elucidar os métodos de análise utilizados, faremos uma descrição sucinta do MAVIS.

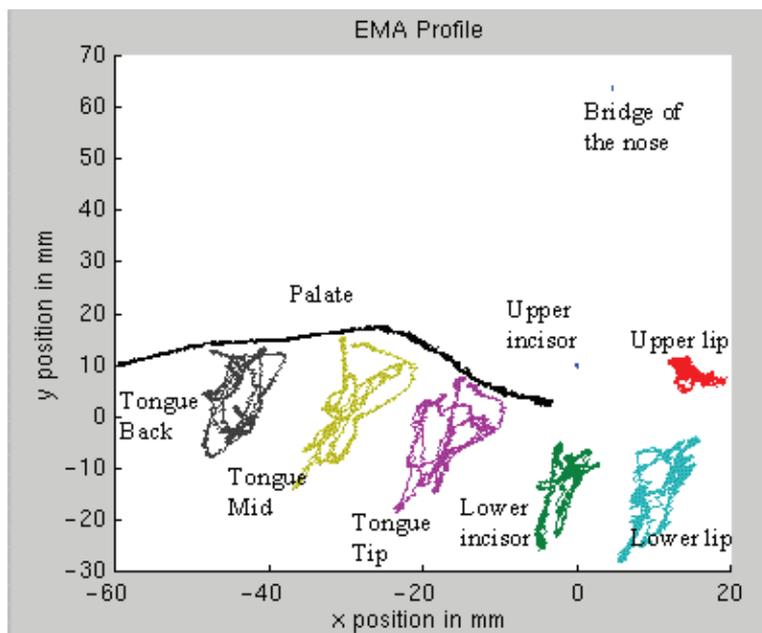
O programa MAVIS (TIEDE et al., 1999) foi elaborado pelo Dr. Mark Tiede (atual cientista sênior do Haskins Laboratories), no ATR Research Laboratories,

2 O texto aqui apresentado é uma adaptação atualizada das informações apresentadas para análise articulatória de dados prosódicos de minha tese de doutorado (MEIRELES, 2007).

em Kioto, Japão. Como o próprio nome diz, MAVIS (*Multiple Articulator Visualizer*) é um visualizador de múltiplos articuladores em simultaneidade. Em outras palavras, tem a função de visualizar vários articuladores ao mesmo tempo, como ponta de língua, dorso de língua, mandíbula, entre outros.

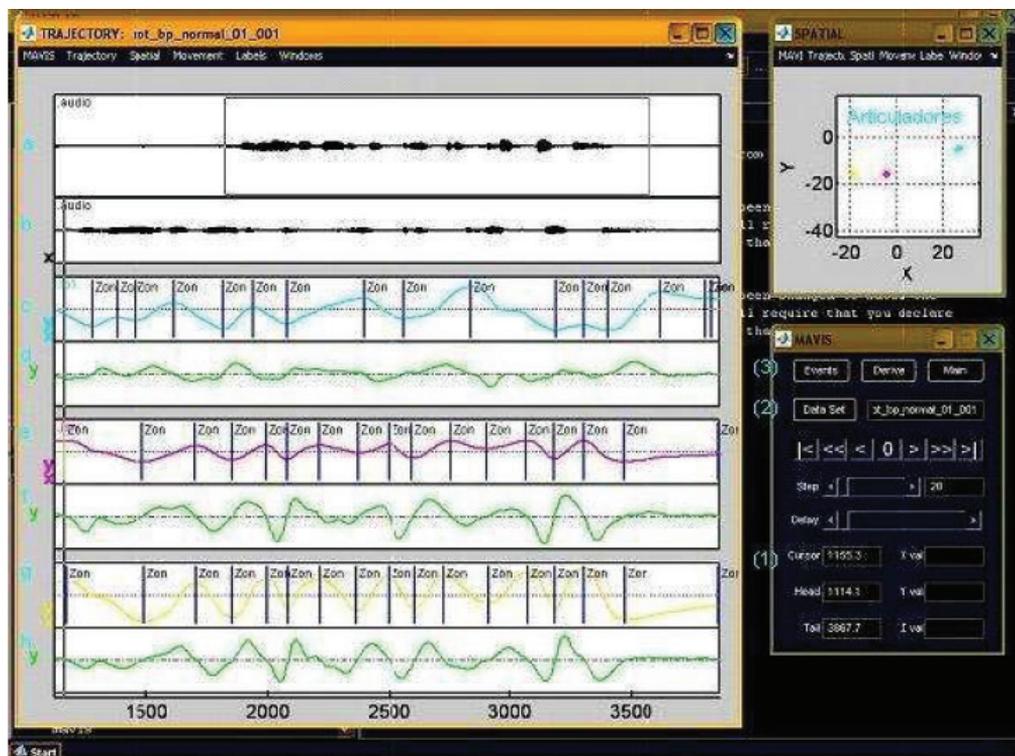
Como consta nas informações do programa, MAVIS é um conjunto de funções integradas do Matlab úteis para visualizar dados de movimentos sincronizados, como os gerados por métodos de rastreamento da fonte de um ponto, como o magnetômetro (EMA), OPTOTRAK, ou sistema de microfios de raios-x. Provê apresentação simultânea de trajetórias variantes no tempo e sua distribuição espacial em um contrabalanceamento temporal, que pode ser incrementado automaticamente para vários efeitos de animação.

MAVIS também permite a apresentação de informações estáticas como o traçado do palato ou dados de imagem, e fornece suporte para procedimentos tanto internos quanto requeridos pelo usuário (Figura 5).



**Figura 5** Representação de vários traçados possíveis com o magnetômetro: referência do nariz, palato, dentes incisivos superiores e inferiores, lábios inferiores e superiores, ponta de língua, parte central da língua ([www.phon.ox.ac.uk/~paula/emahome.html](http://www.phon.ox.ac.uk/~paula/emahome.html)).

Para uma apresentação sucinta das funcionalidades do MAVIS, vejamos a análise da Figura 6, obtida quando digitamos “mavis” dentro do programa MATLAB. Observamos que, o programa MAVIS, por não ser nativo ao MATLAB, deve estar em dos PATHs do MATLAB para que tal comando funcione.



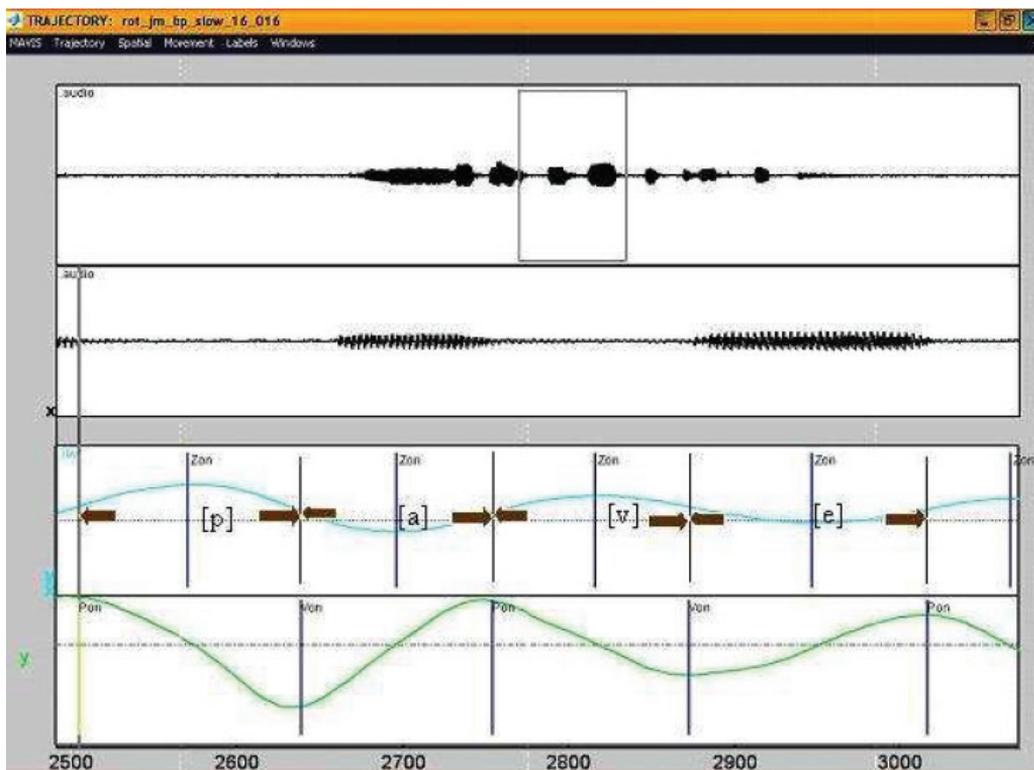
**Figura 6** Janela do MAVIS com todas suas funcionalidades no MATLAB.

Na Figura 6, temos, de cima para baixo: a) onda sonora completa; b) onda sonora selecionada pelo usuário; c) movimento articulatório do corpo da língua; d) velocidade (eixo vertical) de deslocamento do corpo da língua; e) movimento articulatório da mandíbula; f) velocidade (eixo vertical) de deslocamento da mandíbula; g) movimento articulatório dos lábios inferiores; h) velocidade (eixo vertical) de deslocamento dos lábios inferiores. Nas duas janelas à direita, temos: 1) parte superior (janela SPATIAL): disposição cartesiana dos três articuladores empregados no experimento 1: a) lábios (amarelo); b) mandíbula (roxo); c) corpo da língua (azul claro); 2) parte inferior (janela MAVIS): informações temporais de posicionamento dos gestos articulatórios (1 na Figura 6), nome do arquivo aberto (2 na Figura 6), e recursos adicionais (3 na Figura 6), como “Events”, “Derive”<sup>3</sup>.

A fim de compreendermos como são delimitados os gestos articulatórios, conforme especificado pela Fonologia Articulatória (BROWMAN e GOLDSTEIN, 1992), podemos notar que para delimitarmos articulatoriamente a palavra “pavê”

3 Para um detalhamento mais aprofundado dos procedimentos de análise no MAVIS vide o capítulo 4 de Meireles (2007).

da Figura 7, por exemplo, precisamos ter informações não somente da velocidade zero (correspondente, aproximadamente, ao ponto de máxima excursão do gesto), mas também de picos (velocidade máxima) e vales (velocidade mínima) dos mesmos. Com essas informações podemos delimitar o gesto<sup>4</sup>, como fizemos na Figura 7.



**Figura 7** Delimitação dos gestos articulatórios para a palavra “pavê” medidas entre intervalos de velocidade no eixo vertical (y, traçado verde), conforme podemos ver entre as setas marrons.

Por fim, após delimitarmos todos os gestos articulatórios do corpus, salvávamos seus respectivos valores e obtínhamos como resultado um arquivo .tsv pronto para análise em planilhas eletrônicas, como o Excel. Um exemplo desse arquivo é apresentado na tabela 1.

4 Aqui o gesto articulatório foi delimitado entre Pon (*peak velocity onset*) e Von (*valley velocity onset*). Outra forma de delimitar o gesto articulatório na literatura é entre Zon (*zero velocity onset*) e Zof (*zero velocity offset*).

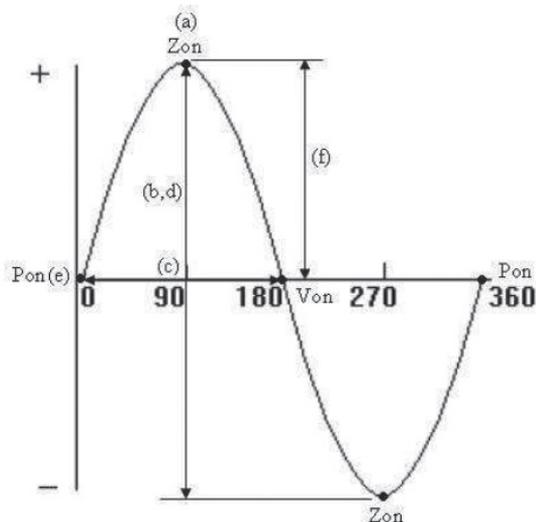
**Tabela 1** Informações temporais (zeros da velocidade no eixo y) e espaciais (informações do deslocamento no eixo vertical) do gesto mandíbula, juntamente com a referência segmental.

<b>SIGNAL 3:</b>	<b>rot-jm-bp-slow-16-016.jaw.VEL(:,2)</b>
Zero (tempo)	Info (deslocamento no eixo vertical)
1590.39	-17.0645
2005.53	-26.298
2285.625	-17.7774
2410.667	-22.4168
2570.721	-17.2356 <b>p</b>
2695.763	-21.5551 <b>a</b>
2815.803	-18.2515 <b>v</b>
2945.847	-20.7704 <b>e</b>
3065.888	-18.533
3180.926	-22.2437

## 4 Medidas articulatório-quantitativas de dados prosódicos

Descreveremos aqui algumas das medidas articulatórias utilizadas em nossas pesquisas sobre o ritmo da fala (MEIRELES; BARBOSA, 2008, 2009, 2014), as quais foram baseadas na pesquisa sobre o gesto prosódico ( *$\pi$ -gesture*) da Dra. Dani Byrd (BYRD; SALTZMAN, 2003), a fim de exemplificar quantitativamente algumas das possíveis medidas obtidas pela técnica EMA.

As seguintes variáveis articulatórias foram medidas em minha pesquisa de doutorado: (a) deslocamento da mandíbula: medido no cruzamento em zeros da velocidade-y; (b) deslocamento da constrição da mandíbula: medido como a diferença entre os cruzamentos em zero no início e no extremo da constrição; (c) duração (relacionada à acústica) do gesto mandíbula: medida pelo intervalo entre velocidades de máxima (positiva) e mínima (negativa); (d) duração do gesto mandíbula: medida entre os cruzamentos em zero da velocidade-y entre o início e o máximo da constrição; (e) velocidades de pico e vale da constrição do gesto mandíbula; (f) duração da aceleração da mandíbula; (g) tempo-para-velocidade-de-pico proporcional: medido do quociente entre duração da aceleração da mandíbula (eixo y) e duração da formação da constrição. Na Figura 8, temos uma descrição visual das medidas articulatórias aqui relatadas (MEIRELES, 2007, 2009).

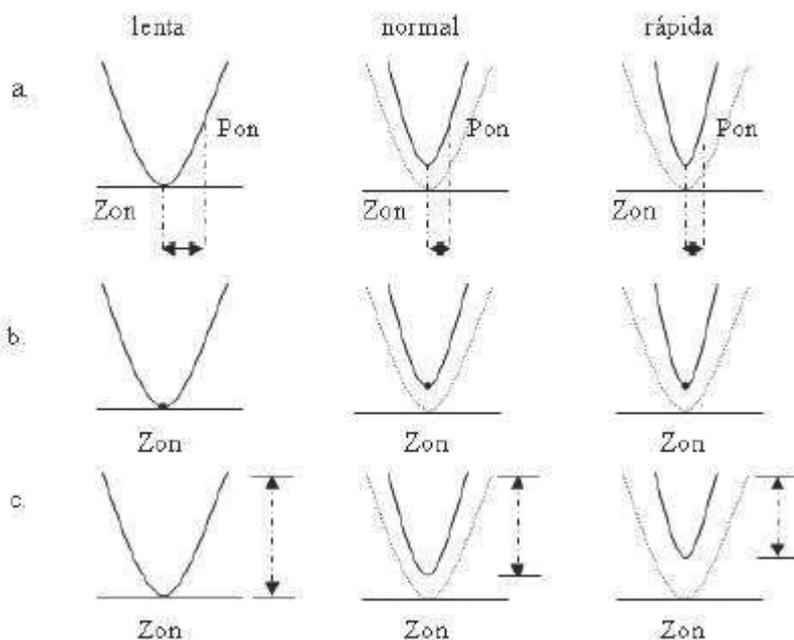


**Figura 8** Representação abstrata das medidas articulatórias realizadas para o experimento acústico prosódico. Von representa, onset do vale de velocidade (velocidade negativa); Pon, onset do pico de velocidade (velocidade positiva); e Zon, onset dos zeros de velocidade. O intervalo de 0 a 90 graus representa um trecho do movimento de fechamento do gesto até a constrição máxima (a), considerando-se uma consoante. O intervalo de 90 a 270 graus representa a abertura de um gesto consonantal. Temos aqui representados: (a) deslocamento da mandíbula, medido no cruzamento em zeros da velocidade y (valor pontual); (b) deslocamento da constrição da mandíbula, medido como a diferença entre os cruzamentos em zero no início (a) e no extremo (270 graus) da constrição, considerando-se um gesto vocálico; (c) duração (relacionada ao acústico) do gesto mandíbula, medida pelo intervalo entre velocidades máxima (Pon) e mínima (Von); (d) duração do gesto mandíbula, medida entre os cruzamentos em zero da velocidade y entre o início e o máximo da constrição; (e) velocidades de pico (Pon) e vale (Von) da constrição do gesto mandíbula (valor pontual); (f) duração da aceleração da mandíbula: positiva (270 a 360 graus) e negativa ((a) a 180 graus). O tempo-para-velocidade-de-pico proporcional é medido do quociente entre a duração da aceleração da mandíbula (f) e duração da formação da constrição (d) (MEIRELES, 2007, p.188).

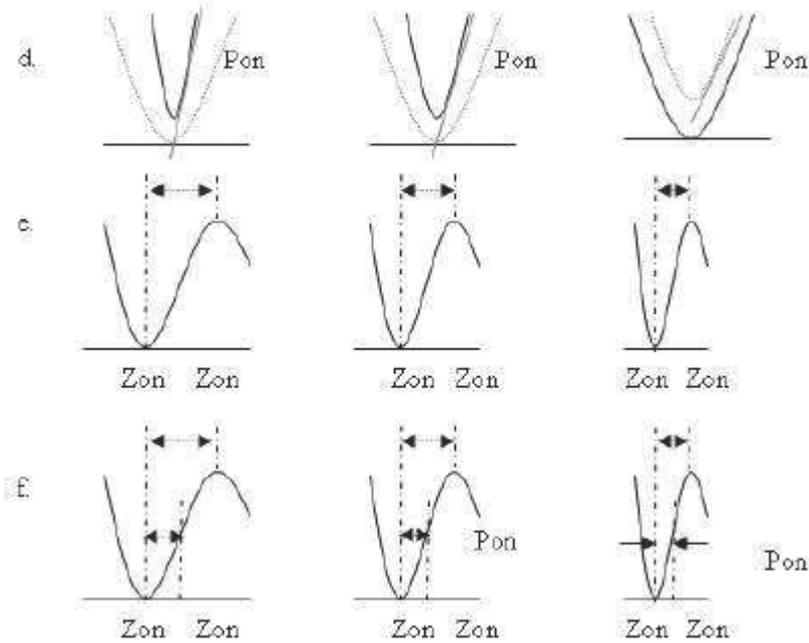
Com base nos resultados articulatórios apresentados acima, propomos as seguintes fontes cinemáticas de reestruturações rítmicas com o aumento da taxa de elocução (cf. Figura 9, MEIRELES, 2007, p. 202): a) diminuição da duração da aceleração, i.e., aumento da rigidez; b) diminuição do deslocamento vertical máximo; c) diminuição do deslocamento do gesto de constrição; d) diminuição (em módulo) das velocidade de pico e vale; e) diminuição da duração articulatória; e f) tempo-para-velocidade-de-pico proporcional permanece constante. Outra fonte cinemática de reestruturação rítmica (não relacionada à mudança da taxa de elocução) é a diminuição de tempo-para-velocidade-de-pico proporcional de vogais tônicas comparadas com as átonas.

De acordo com essas fontes, espera-se uma diminuição geral de parâmetros articulatórios da taxa lenta à rápida. A Figura 9a representa um encurtamento

articulatório causado por uma mudança na rigidez (cf. SALTZMAN; MUNHAL, 1989), que é indicado por uma duração da aceleração mais curta. A Figura 9b exibe deslocamento vertical máximos menores causados por uma posição menos aberta da mandíbula (vogais baixas) ou menos fechada (vogais altas ou meio-altas), que podem ser entendidas como um *undershoot* do gesto. Relacionada com a Figura 9b está a Figura 9c que representa uma menor diferença de deslocamento-y nas posições inicial e final do gesto. Se se espera menores deslocamento vertical máximos, também se espera menores deslocamentos da constrição nas taxas rápidas. A Figura 9d mostra uma diminuição (em módulo) da velocidade de pico e/ou vale da taxa lenta à rápida. Esta hipótese pode ser explicada pelo fato de que, para se alcançar uma grande velocidade de pico e depois uma velocidade zero na mudança do movimento (a mandíbula sobe e desce), é necessária uma maior distância, que pode ser encontrada em taxas lentas, não rápidas. A Figura 9e é meramente uma consequência das fontes cinemáticas anteriores, i.e., a diminuição da duração articulatória (tanto AT quanto medida entre velocidades de picos e vales). Por fim, a Figura 9f mostra que o tempo-para-velocidade-de-pico proporcional permanece constante com o aumento da TE.



**Figura 9** Fontes cinemáticas de reestruturas rítmicas com o aumento da TE: a. diminuição da duração da aceleração, i.e., aumento da rigidez; b. diminuição do deslocamento vertical máximo; c. diminuição do deslocamento do gesto de constrição; d. diminuição (em módulo) da velocidade de pico; e. diminuição da duração articulatória (tanto medida entre velocidades de pico quanto entre zeros); f. tempo-para-velocidade-de-pico proporcional permanece constante (Meireles, 2007:202). (continua)



**Figura 9** Fontes cinemáticas de reestruturações rítmicas com o aumento da TE: a. diminuição da duração da aceleração, i.e., aumento da rigidez; b. diminuição do deslocamento vertical máximo; c. diminuição do deslocamento do gesto de constrição; d. diminuição (em módulo) da velocidade de pico; e. diminuição da duração articulatória (tanto medida entre velocidades de pico quanto entre zeros); f. tempo-para-velocidade-de-pico proporcional permanece constante (Meireles, 2007:202). (continuação)

## 5 Considerações finais

Apresentamos neste capítulo uma introdução geral para a utilização da técnica EMA para a análise de dados prosódicos. Como podemos notar, este texto não tem a intenção de ser uma discussão exaustiva do tema e sim de ser uma primeira aproximação com esse tipo de pesquisa. Concluindo, gostaríamos de apresentar um detalhe metodológico importante na aquisição dos dados articulatórios para depois apresentar alguns textos importantes para um aprofundamento do tema.

Diferentemente da aquisição de dados acústicos, a aquisição de dados articulatórios requer um tratamento muito mais cuidadoso, pois, do contrário, não conseguimos delimitar os gestos articulatórios e, assim, não conseguiremos fazer as medidas articulatórias necessárias. Como pudemos perceber nas explicações do gesto articulatório apresentadas nas figuras 6-9, a delimitação deste depende de uma forma de onda quase-senoidal, pois, se não houver uma movimentação articulatória ampla, o algoritmo de detecção automática de softwares como MAVIS e

MVIEW<sup>5</sup> não consegue delimitá-lo. Por esse motivo, não é qualquer sentença que pode ser utilizada na detecção dos dados articulatórios com EMA e também com a técnica de ressonância magnética (MRI). Como exemplo de uma sentença gravada com o equipamento WAVE que não tinha uma alternância ampla no espaço articulatório oral, observemos a Figura 10.

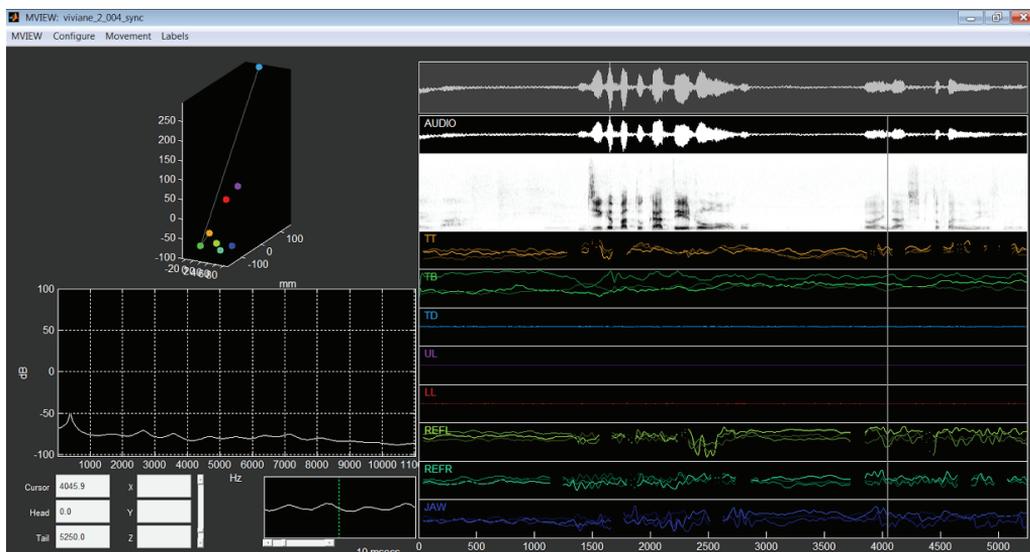
Na Figura 10, abaixo do espectrograma há uma linha verde intitulada TB (*tongue body* [parte média da língua]). Como podemos observar na parte inferior deste articulador (verde claro) não há uma senoide muito bem definida. Apresenta-se praticamente como uma linha reta com algumas irregularidades. Sendo assim, as medidas articulatórias para esse parâmetro estão comprometidas, pois a detecção automática do MVIEW não funciona para delimitar os vários gestos articulatórios presentes nessa sentença. Lembramos que cada segmento fonético é um gesto articulatório. Esse problema para detecção dos gestos articulatórios, deve-se ao não-controle dos movimentos articulatórios por parte do pesquisador.

Como exemplo de desenho experimental, descrevo aqui os procedimentos adotados para coletar dados para minha pesquisa de doutorado (MEIRELES, 2007). Nessa pesquisa, pretendíamos utilizar neste experimento as mesmas sentenças utilizadas num estudo acústico anterior, porém percebemos que, devido à enorme variabilidade articulatória nesses dados, não poderíamos utilizar nem as mesmas sentenças nem as mesmas palavras usadas em estudos acústicos anteriores (publicados posteriormente como MEIRELES; BARBOSA, 2008, 2009, 2014).

Como exemplo desta impossibilidade ou impraticabilidade de gravação, cito a sentença “Há três tipos de abóbora em Belo Horizonte”. Observamos que, para esta sentença, temos, entre outras, as vogais [a, e, i, u, ɔ, o, ê, õ], ou seja [a] = vogal posterior faringal, [e, i] vogais anteriores, [u, ɔ, o] = vogais posteriores arredondadas, [ê, õ] = vogais nasais. Para trabalhar com essa sentença, teríamos, pois, no mínimo, três sensores para a língua: um na parte posterior-faringal, outro na parte posterior, e outro na parte anterior da língua. Além disso, precisaríamos de um sensor no véu palatino, local problemático devido à ocorrência de ânsia de vômito.

---

5 Um conjunto de *scripts* do Matlab desenvolvido por Mark Tiede no Haskins Labs. Este pesquisador foi o mesmo que desenvolveu o MAVIS, que, serviu de base para o MVIEW. Para uma descrição de utilização do MVIEW consultem o manual disponível em [http://www.haskins.yale.edu/staff/gafos\\_downloads/ArtA3DEMA.pdf](http://www.haskins.yale.edu/staff/gafos_downloads/ArtA3DEMA.pdf).



**Figura 10** Exemplo de uma sentença não controlada articulatoriamente visualizada no MVIEW.

Tendo em vista a necessidade de se ter um corpus bem controlado articulatoriamente, a fim de se evitar ruídos articulatorios posteriores, elaboramos um desenho experimental em que as sentenças possuíam um padrão alternante de mandíbula aberta e fechada durante todo o enunciado. Por essa razão utilizamos uma sequência de vogais altas (ou média-altas) e vogais baixas do começo ao fim da sentença e também consoantes bilabiais nas fronteiras frasais previstas (MEIRELES, 2007). Esse procedimento de alternância de vogais altas e baixas deve ser seguido sempre que formos adquirir dados articulatorios, a fim de não prejudicarmos nossas análises articulatorias posteriores. Descrevemos sucintamente alguns artigos relevantes para a prosódia da fala que utilizaram dados de EMA e que podem ser gratuitamente encontrados na *internet*.

Byrd e Saltzman (1998) utilizaram dados de magnetômetro de 3 sujeitos, a fim de determinar se múltiplos níveis de fronteiras prosódicas podiam ser distinguidos no padrão espaço-temporal da articulação e se esses padrões seriam consistentes com um abaixamento da rigidez (*stiffness*) gestual. A modelagem empregada sugere que, embora o abaixamento da rigidez gestual seja a principal fonte de alongamento, outro parâmetro – o tempo de subida para a ativação gestual – também é necessário.

O artigo de Byrd et al. (2000) examinou o padrão espacial e temporal de gestos articulatorios orais em função da estrutura frasal para sequências de duas consoantes nasais em Tamil. Os resultados mostram que a estrutura prosódica é manifestada nos detalhes da articulação. A natureza precisa, contudo, dos efeitos frasais tanto na coordenação temporal inter-gestos quanto na magnitude e duração gestual permanecia um questão aberta à época.

Aprofundando os estudos sobre os efeitos da prosódia na articulação, Byrd e Saltzman (2003) examinaram a relação entre estrutura frasal e o controle e coordenação da articulação em um modelo de sistemas dinâmicos na produção da fala. Nesse artigo é apresentada a noção de gesto prosódico, que atua em fronteiras prosódicas para esticar ou encolher temporalmente as trajetórias de ativação gestual. Segundo os autores, entender fronteiras frasais como uma distorção do tecido temporal de um enunciado representa um confluência promissora entre os campos da prosódia e da dinâmica da fala.

Além de artigos na área de coordenação temporal da fala, também encontramos na literatura trabalhos que procuram mostrar a influência da emoção na coordenação gestual. Lee e Narayanan (2012), por exemplo, investigaram os efeitos da emoção nos movimentos dos lábios inferiores na produção de palavras monossilábicas em dois diferentes contextos prosódicos: início e final de enunciado. Os resultados encontrados mostram que a rigidez dos gestos de abertura do lábio inferior são quase invariantes sob variações emocionais, implicando que o gesto lábio inferior é regulado por um oscilador harmônico simples com uma rigidez constante.

Por fim, gostaria de salientar que, devido à sua especificidade técnica, controle de articuladores supralaríngeos, a maioria dos estudos com EMA na literatura fonética são feitos com dados segmentais e, portanto, são poucos os trabalhos realizados em prosódia da fala. Assim, há um vasto campo de pesquisa ainda a ser explorado na utilização do magnetômetro em dados prosódicos.

## Referências

- BROWMAN, C.; GOLDSTEIN, L. Articulatory phonology: an overview. *Phonetica*, v. 49, p. 155-180, 1992.
- BYRD, D.; SALTZMAN, E. Intra-gestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, v. 26, p. 173-199, 1998.
- BYRD, D.; KAUN, A.; NARAYANAN, S.; SALTZMAN, E. Phrasal signatures in articulation. In: BROE, M. B.; PIERREHUMBERT, J. B. (Ed.). *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge, Reino Unido: Cambridge University Press, p. 70-87, 2000.
- BYRD, D.; SALTZMAN, E. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, v. 31, p. 149-180, 2003.
- LEE, S. e NARAYANAN, S. Effects of Emotion on the Lower Lip Movements at Phrase Boundaries. *Proceedings of Speech Prosody 2012*, Shanghai, 2012.

- MARINO, V. C. C. Avaliação instrumental da fala: articulografia. In: IX Jornada Fonoaudiológica - Prof. Mariza Ribeiro Feniman, 2002, Bauru. *IX Jornada Fonoaudiológica - Prof. Mariza Ribeiro Feniman*. [S.l.: s.n.], p. 24-24, 2002.
- MEIRELES, A. Reestruturações rítmicas da fala no português brasileiro. Tese de Doutorado, Universidade Estadual de Campinas, Brazil, 2007.
- MEIRELES, A. R. Self-organizing rhythms in Brazilian Portuguese: speech rate as a system perturbation. Alemanha: VDM Verlag, 2009.
- MEIRELES, A. R.; BARBOSA, Plínio Almeida . Lexical reorganization in Brazilian Portuguese: an articulatory study. *Speech Communication*, v. 50, p. 916-924, 2008.
- MEIRELES, A. R.; BARBOSA, Plínio Almeida. O papel da taxa de elocução nos processos dinâmicos de mudança linguística. *Revista (con) textos linguísticos (UFES)*, v. 3, p. 91-116, 2009.
- MEIRELES, A. R.; BARBOSA, Plínio Almeida . Articulatory Reorganizations of Speech Rhythm due to Speech Rate Increase in Brazilian Portuguese. In: *Speech Prosody 2014*, 2014, Dublin, Irlanda. Proceedings of Speech Prosody 2014, 2014.
- PERKELL, J., COHEN, M., SVIRSKY, M., MATTHIES, M., GARABIETA, I., JACKSON, M. Electromagnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, 3078–3096, 1992.
- TIEDE, M.K., VATIKIOTIS-BATESON, E., HOOLE, P., YEHIA, H. Technical Report 1999. Kyoto, Japan: ATR Human Information Processing Labs. *Magnetometer data acquisition and analysis software for speech production research*, 1999.

## Aspectos de produção e percepção de estilos de elocução profissionais e não profissionais em quatro línguas

Plínio A. Barbosa  
Unicamp

### Introdução

Ao fazer uma revisão dos trabalhos apresentados para o *Workshop on Integrating Speech and Natural Language*, Llisterri (1992) apontava para a diversidade de termos associados a estilo de elocução. Embora os estilos majoritariamente mencionados fossem (e ainda são) os da fala lida e da fala espontânea, o estilo de elocução profissional fora citado para se referir à fala usada por jornalistas e professores em seu trabalho (DELGADO; FREITAS, 1991).

Essa variação de terminologia levou Eskénazi (1993) a propor uma definição de estilo de elocução (*speaking style*, em inglês) fundamentada na observação de três dimensões de variação: grau de clareza, familiaridade e estrato social. A primeira dimensão vai do esforço mínimo à clareza máxima (de pouco a muito inteligível, como da “leitura para criança” à “locução de futebol”), a segunda depende do grau de familiaridade entre os interlocutores (de familiar a não familiar, como de “falando a um amigo próximo” a “falando a um estranho”), enquanto que a terceira depende da situação e da diferença de nível social entre os interlocutores (de baixo estrato a alto estrato, como de “conversa diária” a “conferência formal”).

Um estudo adequado de estilos de elocução pressupõe, assim, que se identifique com clareza o tipo de tarefa que realiza o locutor ao se dirigir a um único ou mais interlocutores (presumíveis ou presentes). E, portanto, a denominação restrita de “fala espontânea” é insuficiente quando não é descrita a partir de dimensões como as apontadas por Eskénazi ou por Barbosa (2012), esse último a partir de dois eixos de variação: de gênero textual e grau de controle do experimentador. O estilo de fala se dá, portanto, numa dimensão eminentemente de variação que

requer um estudo formal e controlado dos parâmetros acústicos que são modificados (BARRY, 1995; BARBOSA; MADUREIRA, 2016). A necessidade desse estudo é historicamente apontada por Eskénazi (1993), após uma revisão de 150 artigos sobre o tema que mostram que os estilos de elocução estão relacionados a mudanças em qualidade de voz, como soprosa, tensa, laringalizada; taxa de elocução e situações específicas de comunicação, como fala profissional, entrevista e transmissão de notícias e esportes via rádio e televisão.

Tanto parâmetros segmentais quanto prosódicos são modificados entre os diferentes estilos. O estilo de elocução é, assim, entendido como um aspecto da diferenciação do modo de falar (IRVINE, 2001), que perpassa a interface entre fonética e sociolinguística, já apontada por Argente (1992) e cunhada pelo estudo da sociofonética, conforme propuseram Foulkes e Docherty (2006) ao fazer uma retrospectiva de seu uso:

The term has a reasonably long history (tracing back at least to Deshaies-Lafontaine, 1974), but it has been neither widely nor consistently used. [...] It has had its widest currency among phoneticians, who have used it to refer especially to descriptive accounts of speech production across different dialects, speaker groups or speech styles (e.g., Esling, 1991; Henton & Bladon, 1988).

Essa diferença no modo de falar, entre estilos de elocução, pode ser estudada pelo exame das diferenças prosódicas dos enunciados produzidos em cada estilo. Pressupõe-se que os ouvintes conseguem discriminar e identificar essas diferenças a partir da prosódia, conforme demonstra a literatura mais recente.

Ao usar enunciados deslexicalizados<sup>1</sup> de quatro diferentes estilos profissionais em francês padrão, a saber, o comentário esportivo, o sermão religioso bem como os discursos político e jornalístico, Obin et al. (2011) mostraram que é possível identificar acima do aleatório esses quatro estilos em teste de classificação de escolha forçada. Os resultados apontaram que o comentário esportivo é o melhor identificado, seguido do discurso jornalístico. O discurso político e o sermão religioso foram pior identificados. Esses dois últimos são muito confundidos entre si, enquanto o jornalístico é em parte confundido com o discurso político. O comentário esportivo é o menos confundido dos quatro. Os autores apontam que parâmetros como taxa de elocução, intensidade, uso de pausa e estruturação

---

1 A deslexicalização é uma técnica que permite mascarar ou eliminar a informação segmental impossibilitando a compreensão do conteúdo e preservando a informação prosódica do enunciado.

de proeminências prosódicas são os principais parâmetros que guiam os ouvintes na classificação.

São também características prosódicas que distinguem os estilos estudados por Higuchi et al. (1997) em japonês, a saber, enunciados naturais atuados com raiva, gentileza e de forma apressada. Os autores apontaram que o valor mínimo da frequência fundamental ( $F_0$ ) se manteve alto para a raiva com contornos melódicos monótonos. A variação melódica na fala com gentileza foi a maior entre os três estilos, enquanto que o estilo apressado exibiu contornos melódicos com variação intermediária.

Barbosa e Silva (2012) mostraram que os ouvintes discriminam fala lida da narrada em português brasileiro também com a fala deslexicalizada e que as taxas de elocução, de sílabas acentuadas e o uso de pausas são os principais parâmetros empregados para discriminar esses dois estilos. Recentemente, mostramos que parâmetros temporais têm papel mais relevante, embora não exclusivo, que parâmetros melódicos para realizar a discriminação entre falas narrada e lida (BARBOSA, 2015).

Assim, como mostramos nesses trabalhos e em pesquisas anteriores (cf. BARBOSA, 2006), embora a organização da duração de unidades silábicas ao longo dos enunciados tenha um papel fundamental para a estruturação rítmica da fala e, portanto, para a diferenciação entre estilos como a narração e a leitura, é possível que, com a inclusão de outros estilos, especialmente os profissionais, a melodia tenha um papel que tenha sido até então negligenciado. É esse o intuito deste trabalho, ao comparar diferenças prosódicas acústicas e percebidas em quatro línguas e quatro estilos de elocução, sendo dois profissionais (discursos político e jornalístico) e dois não profissionais (leitura e narração).

## 2 Metodologia

### 2.1 Corpus

O corpus CROSS-RHYTHM compreende gravações em português brasileiro (PB), português europeu (PE), francês padrão (FR) e alemão padrão (AL). Ele é composto de material de fala lida e narrada de 10 sujeitos (5 homens e 5 mulheres) entre 25 e 40 anos em cada língua, a que chamamos de fala não profissional, uma vez que nenhum desses sujeitos usa a voz profissionalmente. Para os discursos político e jornalístico, colhemos gravações de até 5 minutos do *YouTube* de 6 sujeitos (3 homens e 3 mulheres), também em cada língua, que denominamos estilo profissional. As gravações da fala não profissional foram feitas no Brasil por Sandra Madureira (PUC-SP) e por mim, enquanto que as gravações nas demais

línguas foram realizadas por colaboradores em Portugal (Catarina Oliveira, Univ. de Aveiro), França (Philippe Boula de Mareüil, LIMSI) e Alemanha (Hansjörg Mixdorff, Beuth Univ. of Berlin), todas usando um microfone unidirecional. Os áudios dos vídeos da seleção de fala profissional foram extraídos em formato WAV. Todos os enunciados foram amostrados ou reamostrados a 16 kHz, com 16 bits de quantização. A partir das gravações, extraíram-se trechos para análise acústica e perceptiva.

Todas as seleções de trecho para análise nos quatro estilos foram feitas para as quatro línguas, totalizando 517 trechos. Os trechos se estendem de 10 a 20 segundos, extensão considerada ideal para a análise perceptiva de diferenças no modo de falar, como mostrou estudo anterior (BARBOSA; SILVA, 2012).

O texto escolhido para a leitura foi “O Monge Desastrado”, um texto originalmente escrito em PE sobre a origem dos pastéis de Belém. O texto foi adaptado para o PB por Juva Batella e traduzido para o francês por Philippe de Mareüil e por mim, e, para o alemão, por Hansjörg Mixdorff (a partir de uma tradução para o inglês, feita por mim). As narrações foram feitas logo após as leituras pelos mesmos sujeitos em cada língua com gravações conduzidas pelos mesmos colaboradores mencionados acima.

## 2.2 Parâmetros acústicos

O script *ProsodyDescriptor2*, que roda em Praat (BOERSMA; WEENINK, 2016), que desenvolvemos para este trabalho, calcula oito parâmetros prosódicos a partir de um arquivo de áudio. Os parâmetros são mediana, desvio padrão e assimetria de  $F_0$  em semitons (re 1 Hz), mediana, desvio padrão e assimetria da taxa de  $F_0$  em semitons por segundo, taxa de picos da curva suavizada a 1,5 Hz de  $F_0$  e um parâmetro ligado ao esforço vocal, a ênfase espectral (TRAUNMÜLLER; ERIKSSON, 2000), medida obtida através do cálculo da diferença entre a intensidade acústica do sinal original e a intensidade do sinal após ser submetido a um filtro passa-baixas com frequência de corte em 400 Hz. Os seis primeiros parâmetros são estritamente melódicos, enquanto que a taxa suavizada de picos de  $F_0$  é um parâmetro rítmico, por envolver uma taxa de produção de proeminências. A ênfase espectral se relaciona ao esforço vocal: quanto maior seu valor, maior esse esforço.

Os parâmetros prosódico-acústicos mencionados foram obtidos pelo script para uma seleção de 218 trechos de fala lida e narrada dos 10 sujeitos gravados por língua. Os demais 299 trechos foram obtidos das falas profissionais. A partir dos valores dos parâmetros, usamos a estatística inferencial para avaliar as diferenças entre suas médias.

### 3 Análises acústicas dos estilos de elocução

#### 3.1 Estilos não profissionais: falas lida e narrada

O teste não paramétrico Kruskal-Wallis para  $\alpha = 5\%$ , seguido de análises post hoc de Bonferroni, revelaram que os parâmetros assinalados na Tabela 1 diferenciam fala lida e narrada em PB e em francês. Em PE e alemão, não há diferenciação dessas duas falas por nenhum dos oito parâmetros analisados.

**Tabela 1** Parâmetros que discriminaram entre fala lida (LE) e narrada (NR) em PB e FR. Mediana da taxa de  $F_0$  (dF0med) em st/s; Assimetria de Pearson da taxa de  $F_0$  (dF0as); Taxa de picos de  $F_0$  (F0t) em picos/s. Valores de p para o teste de Kruskal-Wallis para  $\alpha = 5\%$ .

Língua	Parâmetro	p <	Valores médios
PB	dF0med	$10^5$	-0,02 (NR) -0,15 (LE)
	dF0as	$10^4$	-0,02 (NR) -0,39 (LE)
	F0t	0,01	0,89 (NR) 0,94 (LE)
FR	dF0med	0,0002	0,05 (NR) -0,05 (LE)

Para o parâmetro dF0med, não há interação entre sexo e estilo em PB, mas há diferença entre os sexos com as seguintes médias: -0,19 (LE, M), -0,067 (NR, M), -0,13 (LE, H) e 0,02 (NR, H). Quanto ao francês, no mesmo parâmetro, diferença entre os sexos apenas para o estilo de leitura com valores médios de -0,1 (M) e -0,02 (H).

Para o parâmetro dF0as, não há interação entre sexo e estilo em PB, mas há diferença entre os sexos com as seguintes médias: -0,55 (LE, M), -0,16 (NR, M), -0,23 (LE, H) e 0,12 (NR, H). Já para o parâmetro F0t em PB, há diferença significativa entre os sexos. Quanto aos estilos, há diferença média nesse parâmetro apenas nas mulheres: -1,01 (LE, M), 0,92 (NR, M), 0,88 (H, nos dois estilos).

Antes de comparar o uso dos parâmetros acústicos entre as línguas, convém explicar para que tipo de aspectos da melodia e do ritmo apontam esses resultados. O parâmetro dF0med está relacionado à media das taxas de variação de  $F_0$ , ou seja, das subidas e descidas melódicas. Para o PB, o valor médio negativo

nos dois estilos, indica maior proporção de descidas do que de subidas. Mas a proporção de subidas é maior na narração, relacionado a subidas mais rápidas e frequentes do que a leitura. Esse dispositivo é ainda maior na narração masculina, que tem proporção total de subidas maior que de descidas. O mesmo se dá em francês, mas apenas na leitura, como homens com maior proporção de subidas que as mulheres. Na narração, nessa língua, a proporção total de subidas é maior que de descidas nos dois sexos.

O parâmetro  $dF_0as$  está relacionado à cauda da distribuição das taxas de variação de  $F_0$ , com valores negativos indicando distribuição assimétrica à esquerda, correlato de uma maior tendência a descidas de  $F_0$ . Em PB, esse valor é quase nulo na narração, indicando que essa tendência da assimetria negativa existe com maior grau na leitura, pela maior proporção de instâncias de terminalidade. O parâmetro  $F_0t$ , por sua vez, é a taxa de produção de picos de  $F_0$  na curva melódica filtrada a 1,5 Hz (para preservar apenas os picos relevantes linguisticamente). Em PB, esse valor é significativamente menor na narração, indicando maior distância temporal entre as unidades salientes acusticamente, o que parece ser característico da narração, relacionado também à presença de pausas silenciosas mais longas do que na leitura.

Quanto às diferenças no uso dos parâmetros entre as línguas, relacionadas ou não a um estilo em particular, a Tabela 2 apresenta os resultados. O que se deduz do exame da tabela é que o PB se aproxima em leitura em francês para dois parâmetros acústicos ( $dF_0as$  e  $F_0t$ ), enquanto se aproxima, independentemente de estilo do PE para  $F_0med$  e  $F_0dp$ . Também se vê que o PB, independentemente do estilo, é distinto do alemão em todos os casos. Que PB e PE não sejam distintos na mediana de  $F_0$  pode indicar um amálgama de aspectos fisiológicos e linguísticos, que fazem que portugueses e brasileiros soem mais agudo.

**Tabela 2** Parâmetros que discriminam entre as línguas. Além dos referidos na Tab. 1, Mediana de  $F_0$  ( $F_0med$ ) em st; Desvio-padrão de  $F_0$  ( $F_0dp$ ) em st; Assim. de Pearson de  $F_0$  ( $F_0as$ ); Desvio-padrão da taxa de  $F_0$  ( $dF_0dp$ ) em st/s; Ênfase espec. (enf) em dB.

Parâmetro	peelo menos, $p <$	Valores médios
$F_0med$	0,01	89 (PB=PE) > 87 (FR=AL)
$F_0dp$	0,02	1,5 (FR) < 2,2 (PB=PE) < 3,0 (AL)

(continua)

**Tabela 2** Parâmetros que discriminam entre as línguas. Além dos referidos na Tab. 1, Mediana de  $F_0$  ( $F_{0med}$ ) em st; Desvio-padrão de  $F_0$  ( $F_{0dp}$ ) em st; Assim. de Pearson de  $F_0$  ( $F_{0as}$ ); Desvio-padrão da taxa de  $F_0$  ( $dF_{0dp}$ ) em st/s; Ênfase espec. (enf) em dB. (continuação)

Parâmetro	peelo menos, p <	Valores médios
$F_{0as}$	$10^{-5}$ (LE) 0,002 (NR)	LE: PB (-0,08) < FR = PE (0,11) < AL (0,28) NR: PB (0,02) < FR =AL= PE (0,22)
$dF_{0med}$	0,04 (LE)	LE apenas: -0,15 (PB) < FR=PE=AL (-0,05)
$dF_{0dp}$	0,05	FR (2,5) < PB (3,2) < PE (4,11) < AL (5,4)
$dF_{0as}$	0,02 (LE)	LE apenas PB = FR (-0,3) < PE = AL (0,2)
enf	$10^{-5}$	FR (0,1) < PE = AL (0,8) < PB (2,6)
$F_{0t}$	0,002	LE: PB = FR (0,9) > AL= PE (0,8) NR: PB (0,9) > AL (0,7)

A fala dos alemães é a mais variável melodicamente entre todos e a do francês a menos. Sobre o primeiro fato já apontamos essa diferença entre PB e alemão em estudo anterior (Barbosa, Mixdorff, Madureira, 2011). A assimetria de  $F_0$  revela em PB uma tendência na narração a uma frequência superior de valores de  $F_0$  mais baixos do que a média na leitura. Essa tendência a uma frequência maior de valores de  $F_0$  mais altos nas demais línguas vale para os dois estilos. O PB tende também a ter taxas de  $F_0$  ( $dF_0$ ) em média mais negativas, privilegiando assim descidas mais rápidas do que nas demais línguas com variabilidade ( $dF_{0dp}$ ) intermediária em realizar isso, com francês a menos variável nesse parâmetro. PB é também a língua que faz o maior esforço vocal ao falar (enf), além da maior taxa de picos de  $F_0$ , o que aponta para uma maior vivacidade (*liveliness* em inglês). Em relação ao PE, PB tem mesmo valor mediano de  $F_0$ , com diferença apenas para as narrativas femininas: 94 st (PB) > 91 st (PE).

### 3.2 Estilos profissionais: discurso jornalístico e político

A Tabela 3 aponta as diferenças entre os dois estilos profissionais nas línguas que os diferenciam do ponto de vista melódico e rítmico. Em alguns casos somente as mulheres exibem a diferença entre os estilos, mas não os homens.

**Tabela 3** Parâmetros que discriminam entre discurso jornalístico (JN) e político (PL) entre línguas pertinentes. Mediana de  $F_0$  (F0med) em st; Ênfase espec. (enf) em dB; Taxa de picos de  $F_0$  (F0t) em picos/s. Valores de p para o teste post hoc de Bonferroni  $\alpha = 5\%$ .

Língua	Parâmetro	p <	Valores médios
PE	F0med	0,006 (M)	92 (JN) < 99 (PL)
	enf	0,0005	2,0 (JN) < 4,2 (PL)
FR	F0t	0,0002 (M)	0,8 (JN) < 1,2 (PL)
	enf	0,05	0,8 (JN) < 1,6 (PL)
AL	F0med	0,008	88 (JN) < 95 (PL)
	enf	0,02 (M)	0,4 (JN) < 1,9 (PL)

A ausência do PB na tabela se deve ao fato de ser a única língua que não distingue os dois estilos para os parâmetros considerados. Para as demais, a ênfase espectral assinala esforço maior no discurso político, embora em alemão essa diferença esteja restrita às mulheres. A mediana de  $F_0$  é superior no discurso político feminino português e, independentemente de sexo, também superior em alemão. O francês distingue pela taxa de picos de  $F_0$ , indicador de vivacidade, com maior valor para discurso político feminino.

**Tabela 4** Parâmetros que discriminam o discurso jornalístico entre as línguas. Vide tabelas anteriores para indicação dos parâmetros. Valores de p para teste de Bonferroni para  $\alpha = 5\%$ .

Parâmetro	peelo menos, p <	Valores médios
F0med	0,04	88 (AL) < 91 (PB=PE=FR)
F0dp	0,03	2,0 (FR=AL) < 3,0 (PB=PE)
dF0dp	0,008	4,0 (FR=AL) < 6,2 (PB=PE)
enf	0,007	0,6 (FR=AL) < 3,4 (PB=PE)
F0t	0,02	somente 0,9 (PB) < FR (1,1)

Quanto a diferenças de parâmetros entre as línguas, somente a fala de jornalistas é distinta entre as línguas, por isso, a Tabela 4 só se refere a esse estilo, o que é um resultado muito interessante. Assim, o discurso político tem mesma média para todos os parâmetros examinados entre as línguas. Outro resultado interessante é que os parâmetros no discurso jornalístico têm mesmas médias entre PB e PE contra a mesma média para francês e alemão.

Além disso, a tabela 4 mostra que PB e PE empenham maior esforço vocal que francês e alemão no estilo jornalístico. Que a variabilidade tanto de  $F_0$  quando de sua taxa é superior em PB e PE contra francês e alemão, tornando o discurso jornalístico mais vivaz nas primeiras. O Alemão tem em média o menor valor de  $F_0$ , soando mais grave que as demais, portanto. Em compensação, quando comparado ao francês, o PB tem menor taxa de produção de picos de  $F_0$ .

### 3.3 Estilos profissionais e não profissionais

A Tabela 5 apresenta uma comparação entre os estilos profissionais e não profissionais para as duas variedades do português.

**Tabela 5** Parâmetros que discriminam entre os estilos profissionais e não profissionais em duas variedades do português. Para todos os casos,  $p < 0,05$  em um teste post hoc não paramétrico de Bonferroni.

Língua	Parâmetros	
PB	F0med	ns
	F0dp	$(PL=JN) > (LE=NR)$
	F0as	$(PL=JN) > NR > LE$
	dF0med	$NR > LE$
	dF0dp	$(PL=JN) > (LE=NR)$
	dF0as	$NR > LE$
	enf	ns
	F0t	$PL > LE, NR < LE$
PE	F0med	$PL > (JN = LE = NR)$
	F0dp	$JN > LE$
	F0as	$(PL=JN) < (LE=NR)$
	dF0med	ns
	dF0dp	$PL > (LE=NR), JN=PL$
	dF0as	ns
	enf	$JN > PL > (LE=NR)$
	F0t	$(PL=JN) > (LE=NR)$

Para o PB, depreende-se da Tabela 5 que os estilos profissionais se agrupam com valores superiores aos dos estilos não profissionais para F0dp, F0as e dF0dp, que se relacionam à variabilidade da curva de F<sub>0</sub>. No estilo político, a taxa de picos de F<sub>0</sub> é maior do que na leitura. Já no PE, com exceção da ausência de significância nas variáveis ligadas à taxa de variação de F<sub>0</sub>, dF0med e dF0as, para as demais os estilos político (F0med), jornalístico (F0dp e enf) ou ambos (F0as, dF0dp, F0t) têm valores superiores aos estilos de leitura (F0dp) ou leitura e narração. Nas duas variedades do português, a variação de F<sub>0</sub> e sua taxa (dF0), além da taxa de picos de F<sub>0</sub> é superior nos estilos profissionais.

**Tabela 6** Parâmetros que discriminam entre os estilos profissionais e não profissionais em francês (FR) e alemão (AL). Para todos os casos,  $p < 0,05$  em um teste post hoc não paramétrico de Bonferroni.

Língua	Parâmetros	
FR	F0med	(PL=JN) > (LE=NR)
	F0dp	(PL=JN) > (LE=NR)
	F0as	JN < NR, LE < NR
	dF0med	ns
	dF0dp	(PL=JN) > (LE=NR)
	dF0as	ns
	enf	(PL=JN) > (LE=NR)
	F0t	JN > (RE=PL=NR)
AL	F0med	PL > (JN =LE=NR)
	F0dp	JN < NR
	F0as	PL < JN < (LE=NR)
	dF0med	ns
	dF0dp	ns
	dF0as	ns
	enf	PL > (LE=NR=JN)
	F0t	(PL=JN) > (LE=NR)

A Tabela 6 apresenta uma comparação entre os estilos profissionais e não profissionais para o francês e o alemão. Para o francês, depreende-se da Tabela 6 que os estilos profissionais também tendem a se agrupar com valores superiores aos dos estilos não profissionais para F0med, F0dp, dF0dp e enf, que se relacionam ao valor mediano de F<sub>0</sub> (F0med), à variabilidade da curva de F<sub>0</sub> e de sua taxa (F0dp e dF0dp) bem como ao esforço vocal (enf). No estilo jornalístico apenas a taxa de picos de F<sub>0</sub> é maior do que nos demais estilos. Como no PE, as variáveis

dF0med e dF0as não diferem entre os quatro estilos. Já no alemão, com exceção da ausência de significância nas três variáveis ligadas à taxa de variação de  $F_0$ , para as demais, os estilos político (F0med e enf) ou os dois estilos profissionais (F0t) têm valores superiores aos estilos de leitura e narração.

Depreende-se das duas tabelas que, em geral, excetuando-se o papel menor dos descritores da taxa de variação de  $F_0$  para o PE, francês e alemão e a mediana de  $F_0$  para o PB, as demais variáveis apontam para valores superiores para os estilos profissionais, diferenciando-os dos não profissionais.

## 4 Testes de percepção

Foram aplicados um teste de classificação e três testes de discriminação. O teste de classificação com 94 estímulos foi feito para reconhecer um dos quatro estilos de elocução nas 4 línguas. De forma a impedir o reconhecimento imediato de um político ou jornalista brasileiro, aplicamos o algoritmo PURR (SONN-TAG; PORTELE, 1998) para deslexicalizar todos os estímulos e fazer com que o ouvinte prestasse atenção apenas na prosódia para a classificação.

Os estímulos separados para os testes de discriminação consideraram os dois sexos de forma igual com pares intra e intersexos, as duas ordens de apresentação entre os estilos na mesma língua e em línguas distintas e a ordem entre as línguas, todas comparadas entre si. Isso foi feito para que não fossem fatores que pudessem afetar os julgamentos. Foram montados três testes: o primeiro, com 56 estímulos não deslexicalizados para discriminação dos estilos leitura e narração, considerando todas as combinações entre línguas, ordens de apresentação e combinação dos sexos. O segundo, com 78 estímulos não deslexicalizados para discriminação do estilo jornalístico em relação aos estilos político e de narração, pois são os estilos em princípio que mais se aproximam de uma fala para um interlocutor imediato. Esse teste excluiu o PB, pois o reconhecimento do profissional enviesaria as respostas. Por conta disso, um terceiro teste de discriminação com áudio deslexicalizado para estímulos do PB foi montado com 66 estímulos (em testes comparando o estilo jornalístico com narração e discurso político, para pares de línguas em que uma delas é o PB num desses estilos). A pergunta ao ouvinte, quando o bloco continha um trecho de estilo jornalístico, foi qual dos dois estímulos separados por um tom puro, era discurso jornalístico. Quando o bloco continha um trecho de leitura contrastado à narração, foi qual dos dois estímulos era leitura. Assim, 294 estímulos foram considerados para os testes, realizados com pelo menos 10 ouvintes brasileiros universitários em cada caso.

Para o teste de classificação, o único estilo que teve reconhecimento indistinto do aleatório foi a leitura. Isto é, a leitura não foi um estilo reconhecido.

Quanto aos demais estilos, eles são todos três diferentes do acerto aleatório (25 %) com  $p$  pelo menos  $10^{-5}$ . Os percentuais de acerto foram: (PL = 47 %, JN = NR = 40 %), significativo ao considerar esses três estilos apenas com  $X^2 = 5,9$ , e  $p = 0,015$  e poder de teste de cerca de 67 %, isto é, a probabilidade de rejeitar uma  $H_0$  falsa. Assim, o estilo discurso político é o mais facilmente identificado. A coerência entre os ouvintes na classificação, embora significativa, é de apenas 0,08 pelo teste Kappa de Fleiss.

A matriz de confusão da Tabela 7 permite visualizar o número de respostas para cada estilo.

**Tabela 7** Matriz de confusão para o testes de classificação.

↓ Resp. \ Estímulos →	LE	NR	JN	PL
LE	84 (29%)	78 (28%)	42	47
NR	90	112 (40%)	61	54
JN	76	47	111 (40%)	45
PL	38	39	62	130 (47%)

Percebe-se na matriz que os estímulos de leitura, por ser leitura de história, provavelmente, são mais confundidos com os de narração e muito próximos aos do discurso jornalístico também. Embora a narração também seja confundida em parte com a leitura, a confusão é menor e os discursos profissionais ficam mais à parte. Quanto a esses, o discurso político é o que tem mais acertos e é confundido em menor grau de forma semelhante com os outros três. Já o estilo jornalístico é confundido mais com o discurso político e a narração. Esses resultados são semelhantes ao de Obin et al. (2011) para o discurso jornalístico. Quanto ao fato de o discurso político ser melhor identificado aqui, mas não no estudo com o francês, é porque esse último inclui o sermão religioso, mais confundido com o discurso político. Mas ambos foram discriminados do jornalístico, conforme apontam os autores.

Quanto às línguas, o teste de proporções para  $\alpha = 5\%$  apontou todas as proporções iguais estatisticamente a 39 % de acerto (significativamente distinto do acerto aleatório). Não há assim efeito de língua no teste de classificação. As percentagens de acerto baixas são devidas ao fato de que o teste usa fala deslexicalizada. Passemos agora a examinar os acertos dos ouvintes quanto à discriminação entre os estilos de elocução.

Para prever as categorias do teste de classificação a partir dos parâmetros acústicos, empregamos uma Análise Discriminante Linear (LDA, na sigla em inglês). Considerando os 517 trechos empregados para a análise acústica apresen-

tada anteriormente, obtém-se com um modelo LDA cuja predição de categorias se encontra na Tabela 8.

**Tabela 8** Matriz de confusão para do modelo LDA.

↓ Resp. \ Estímulos →	LE	NR	JN	PL
LE	165 (76%)	77 (44%)	29 (48%)	18 (28%)
NR	44 (20%)	91 (52%)	8 (13%)	18 (28%)
JN	7 (3,5%)	5 (3%)	14 (23%)	2 (3%)
PL	1 (0,5%)	2 (1%)	9 (15%)	27 (41%)

Observa-se, pela comparação das tabelas 7 (classificação via percepção) e 8 (classificação via acústica), que a discriminação entre os estilos é melhor feita pelo modelo LDA a partir dos parâmetros acústicos para leitura e narração com percentagens de acerto superiores ao feito pelos ouvintes para esses estilos. Mas, para esse mesmo modelo, o estilo narrativo é mais confundido com a leitura do que na percepção. No que toca ao estilo jornalístico, os ouvintes acertam 40% contra apenas 23% via LDA, onde ele é mais confundido com a leitura. O desempenho da classificação do estilo político é semelhantes na classificação automática e via ouvintes.

Entre as quatro línguas: (1) o estilo político é melhor classificado pelo LDA para o PE (em PB e FR ele é mais confundido com a narração), seguido do alemão com percentagem semelhante de confusão com a leitura.; (2) o estilo jornalístico é melhor classificado pelo LDA para o PE (em PB ele é confundido com o político e a leitura enquanto que em FR e AL com a leitura); (3) a leitura é melhor classificada pelo LDA em todas as línguas com destaque para o PB com melhor discriminação em relação aos demais; (4) a narração é melhor classificada pelo LDA em PE, AL e FR enquanto que em PB é confundida com a leitura, que foi a segunda percentagem de classificação de narrações pelos ouvintes para todas as línguas sem o léxico.

Os resultados do teste de discriminação, que têm resposta aleatória para 50% de acerto, uma vez que são feitos por pares de estímulos, revelaram que:

- (1) Entre leitura e narração, o acerto é superior a 90 % para a leitura, independentemente de língua;
- (2) Entre estilos jornalístico e político, o acerto é superior a 85 % para o primeiro estilo, independentemente de língua, excluído o PB;
- (3) Entre estilos jornalístico e narrativo, o acerto é superior a 85 % para o primeiro estilo, independentemente de língua, excluído o PB;

- (4) Entre estilos jornalístico e político, delexicalizado para incluir o PB, o acerto é superior a cerca de 65 % (sig. distinto de 50% para alfa = 1 %) para o primeiro estilo, independentemente de ser entre PB e outra língua ou somente o PB;
- (5) Entre estilos jornalístico e narrativo, delexicalizado para incluir o PB, o acerto é superior a cerca de 80 % para o primeiro estilo, independentemente ser entre PB e outra língua ou somente o PB.

## 5 Conclusão

Os resultados apresentados revelam que:

- (1) O conjunto de parâmetros acústicos extraídos, analisados e usados para classificação predizem em parte o comportamento dos ouvintes, que tendem a considerar a leitura apenas pela informação prosódica mais como narração do que a classificação automática, que a prediz a mais de 75 % corretamente. Quando pareada com a narração apenas, com o léxico, seu reconhecimento é superior a 90 %. O estilo jornalístico é melhor reconhecido pelo ouvinte do que ele é predito pela acústica. O estilo político é o melhor classificado entre os estilos profissionais via acústica ou percepção;
- (2) A coerência entre ouvintes no teste de classificação é baixa, uma vez também que o teste envolve itens deslexicalizados;
- (3) Os testes de discriminação revelam grande facilidade de discriminação entre dois estilos, com exceção dos estilos jornalístico e político, o que é um resultado distinto do teste de classificação pelos ouvintes, mas semelhante àquele via LDA;
- (4) Línguas distintas em geral não interferem na discriminação entre estilos, no entanto, narração e leitura são mais confundidas dentro do PB, apesar de a Tabela 1 mostrar parâmetros que os discriminam que, certamente, não são bem apreendidos pelos ouvintes do PB.
- (5) Os estilos profissionais são melhor discriminados em PE e em segundo lugar em alemão, o que está de acordo com parâmetros de F0med e enf na Tabela 3.
- (6) Embora haja parâmetros que discriminem a fala profissional da não profissional, como há um certo grau de confusão entre esses estilos em cada língua, conclui-se que esses parâmetros nem sempre estão claramente distintos em cada par de estímulos examinados ou são de difícil apreensão pelo ouvinte.

## Agradecimentos

Este trabalho foi desenvolvido no quadro do projeto Universal CNPq “Análise e Modelamento Estatístico Inter-linguístico do Vínculo entre Produ-

ção e Percepção do Ritmo da Fala em Diferentes Estilos de Elocução”, Processo 476358/2013-2. O autor também é financiado pela bolsa de produtividade em pesquisa do CNPq, Processo 302657/2015-0. Agradecemos os colaboradores do projeto, Sandra Madureira, Philippe Boula de Mareüil e Hansjörg Mixdorff, o aluno Robson R. Monteiro, que ajudou com a preparação de áudios para os testes de percepção, bem como nossos locutores das tarefas de leitura e narração e os ouvintes das tarefas de percepção.

## Referências

- Argente, J. A. (1992). From speech to speaking styles. *Speech Communication* 11, 325-335.
- Barbosa, P. A. (2006) *Incursões em torno do ritmo da fala*. Campinas: Pontes/Fapesp.
- Barbosa, P. A. (2012) Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos da Linguagem* (UFMG). 20 (1), p. 11-27.
- Barbosa, P. A. (2015). Temporal parameters discriminate better between read from narrated speech in Brazilian Portuguese. *Proc. of the 18th International Congress of Phonetic Sciences*, Glasgow, Reino Unido: The Scottish Consortium for ICPHS 2015.
- Barbosa, P.; Madureira, S. (2016). Elicitation techniques for cross-linguistic research on professional and non-professional speaking styles. *Proc. Speech Prosody 2016*, 503-507.
- Barbosa, P. A., Mixdorff, H.; Madureira, S. (2011) Applying the quantitative target approximation model (qTA) to German and Brazilian Portuguese. *Proc. of Interspeech 2011*. Florença, Itália. Pp. 2065-2068.
- Barbosa, P. A.; da Silva, Wellington. (2012). A New Methodology for Comparing Speech Rhythm Structure between Utterances: Beyond Typological Approaches. In: H. Caseli et al. (Eds.): PROPOR 2012, LNAI 7243, pp. 329-337. Springer, Heidelberg.
- Barry, W. J. (1995). Phonetics and Phonology of speaking styles. *Proceedings of the International Congress of Phonetic Sciences*, Estocolmo, 1995. v. 2, p. 4-10.

- Boersma, P.; Weenink, D. (2016). Praat: doing phonetics by computer. <http://www.praat.org/>.
- Delgado, M. R.- Freitas, M. J. (1991). Temporal structures of speech: 'Reading news on TV'. In *Proceedings of the ESCA Workshop Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*. Barcelona, 1991. p. 19.1 -19.5.
- Deshaiies-Lafontaine, D. (1974). *A socio-phonetic study of a Québec French community: Trois-Rivières*. Tese de Doutorado, Londres: University College London.
- Eskénazi, M. (1993). Trends in speaking styles research. *Proceedings of the Eurospeech 1993*. Berlim, p. 501-509.
- Esling, J. H. (1991). Sociophonetic variation in Vancouver. In: J. Cheshire (Ed.) *English around the world*. Cambridge: Cambridge University Press. p. 123-133.
- Foulkes, P.; Docherty, G. J. (2006). The social life of phonetics and phonology. *Journal of Phonetics*. 34, 409-438.
- Henton, C.; Bladon, A. (1988). Creak as a sociophonetic marker. In: L. Hyman, & C. N. Li (Eds.). *Language, speech and mind: Studies in honor of Victoria A. Fromkin*. Londres: Routledge. p. 3-29.
- Higuchi, N.; Hirai, T.; Sagisaka, Y. (1997). Effects of speaking style on parameters of fundamental contour. In: van Santen, J. P. H.; Sproat, R. W.; Oliver, J. P.; Hirschberg, J. (Ed.) *Progress in Speech Synthesis*. Nova York: Springer-Verlag, p. 417-428.
- Irvine, Judith, T. (2001). Style as distinctiveness: the culture and ideology of linguistic differentiation. In: Ecket, Penelope; Rickford, John. R. *Style and sociolinguistic variation*. Cambridge: Cambridge University Press. p. 21-43.
- Llisterri, J. (1992). Speaking styles in speech research. In *ELSNET/SALT/ESCA Workshop Integrating Speech and Natural Language*, p. 17--37, University College Dublin.
- Obin, N.; Lanchantin, P.; Lacheret, A.; Rodet, X. (2011). Discrete/Continuous Modelling of Speaking Style in HMM-based Speech Synthesis: Design and Evaluation. *Proc. of Interspeech 2011*, Florença, Itália.

Sonntag, G.P.; Portele, T. (1998). PURR - a method for prosody evaluation and investigation. *Journal of Computer Speech and Language*, 12 (4), 437-451.

Traunmüller, Hartmut; Anders Eriksson (2000). Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107: 3438-3451.



## **A interface entre prosódia e discurso no inglês como língua estrangeira**

Rodrigo Garcia Rosa  
Universidade de São Paulo

### **1 Introdução**

Neste capítulo, discutimos dados da língua inglesa sobre a relação entre prosódia e discurso a fim de oferecermos uma discussão introdutória a essa área da linguística no âmbito de uma língua estrangeira. Para tanto, iniciamos por uma discussão que tem como objetivo localizar a prosódia em um contexto linguístico maior. A decisão por esse caminho, isto é, o de enquadrarmos a prosódia no estudo geral da gramática, se deve ao entendimento de que poucos materiais direcionados ao ensino de gramática de língua estrangeira fazem referência a aspectos prosódicos como parte integrante de construções gramaticais. Posterior a isso, o capítulo propõe uma discussão das relações gerais entre prosódia e discurso antes de embarcar em uma análise mais pormenorizada de dados da língua inglesa sobre proeminência prosódica.

### **2 Prosódia e a arquitetura da gramática**

Uma tentativa de discussão acerca do papel da prosódia como fenômeno da linguagem humana pressupõe uma revisão da literatura produzida na área que foge, em largas escalas, aos objetivos introdutórios deste capítulo. Ademais, dada a ubiquidade em que tal aspecto da linguagem humana é sentido na comunicação de falantes, o objetivo de caracterizá-lo como um todo seria, mais do que ambicioso, demasiadamente ingênuo. Não obstante, se um escrutínio completo desse aspecto da linguagem mostra-se impraticável neste momento, faz-se especialmente necessário que localizemos a prosódia em relação a outros pilares da arquitetura da gramática, uma vez que trataremos, mais adiante, justamente da relação entre prosódia, gramática e discurso. Sendo assim, da mesma forma que se faz

imprescindível localizar a prosódia em relação aos demais aspectos da gramática, é necessário que explicitemos a visão de língua adotada que, neste caso, privilegia a visão geral da linguística cognitiva.

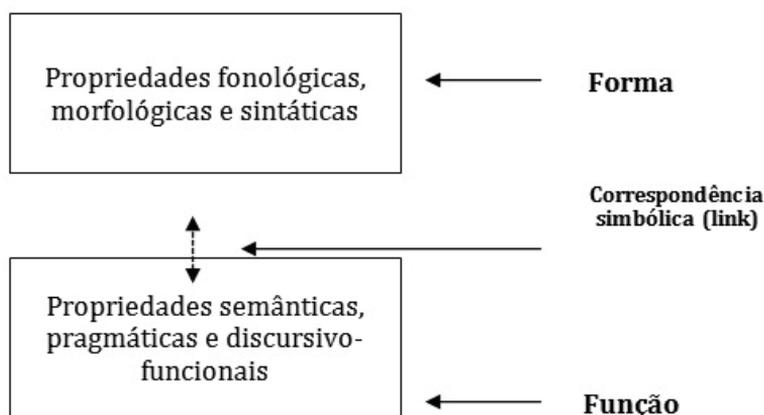
A linguística cognitiva caracteriza-se por ser uma área multifacetada que compreende uma série de subáreas com agendas e arcabouços próprios, como a Gramática Cognitiva langackeriana (LANGACKER, 1987, 2013), a teoria de frames e Modelos Cognitivos Idealizados (MCI) (LAKOFF, 1987), a teoria de Metáforas Conceituais (LAKOFF; JOHNSON, 1980), a Gramática Cognitiva de Construções (GOLDBERG, 1995, 2006), entre outros. Entretanto, há posições centrais a essas perspectivas variadas que as unificam sob o mesmo arcabouço e que, como tal, fazem delas teorias e modelos pertencentes a grande área da linguística cognitiva. Aqui mencionaremos duas dessas posições que se mostram mais relevantes para os objetivos de nossa exposição. A primeira delas refere-se ao compromisso cognitivo, isto é, ao compromisso de que o uso e o conhecimento linguísticos não são derivados de princípios cognitivos direcionados única e tão somente à língua, mas sim de princípios gerais da cognição humana como o raciocínio lógico, a percepção, a memória etc.

Ao discutir o compromisso cognitivo e contrastá-lo com a noção clássica de *competência linguística*, Paradis (2003) afirma que, se há espaço para a noção de competência linguística no arcabouço teórico da linguística cognitiva, essa competência deve ser entendida como a capacidade do falante de reunir uma série de competências cognitivas variadas, tais como a associação, a automatização, a esquematização e a categorização, por exemplo, de modo que o falante possa veicular e compreender significados por meio expressões linguísticas específicas e convencionais. Dito de outro modo, sob a perspectiva cognitiva, a língua não pode ser analisada como um fenômeno isolado de outras capacidades cognitivas mais gerais, isto é, para a linguística cognitiva, o conhecimento linguístico é simplesmente conhecimento (GOLDBERG, 1995).

A segunda característica fundamental refere-se ao fato de que, para a linguística cognitiva, a língua é vista como um conjunto de estruturas simbólicas de níveis variados de complexidade estrutural (de fonemas a construções argumentais) que se realizam através de uma função comunicativa em uma comunidade de fala específica. Essa característica faz com que as bases do conhecimento linguístico estabeleçam-se no uso linguístico que fazemos dos símbolos convencionais, isto é, das estruturas linguísticas pertencentes ao inventário de símbolos que compõem a gramática mental dos falantes.

Segundo Bybee (2006), teorias que propõem explicações *baseadas no uso*, termo cunhado por Ronald Langacker para exprimir a relevância do uso linguístico na construção e na manutenção do conhecimento linguístico (BOAS, 2013; BYBEE, 2006, 2013), tratam, resumidamente, do impacto que tem a

experiência linguística de um falante na representação cognitiva e na modelagem da gramática mental de sua língua. Isto é, o conhecimento de língua ou, mais especificamente, a gramática é vista como a organização cognitiva das experiências que um indivíduo tem com sua língua e como essa experiência foi conceptualizada e condensada na forma de símbolos. Símbolos, também chamados de construções, são definidos, por sua vez, como quaisquer padrões linguísticos em que algum aspecto de sua forma ou função não pode ser diretamente previsível a partir de suas partes constitutivas ou de outras construções reconhecidamente existentes (GOLDBERG, 1995, 2006). De maneira esquemática, podemos representar a estrutura de uma construção como na Figura 1.



**Figura 1** Estrutura interna das construções (adaptado de Croft e Cruse (2004, p. 247)).

A organização das construções pode ser exemplificada por uma relação verticalizada em que um polo formal se liga a um polo funcional por meio de uma relação de ligação simbólica interna à construção. Na parte formal das construções especificam-se as suas propriedades sintáticas, morfológicas e fonológicas que, por sua vez, são ligadas a uma interpretação semântica bem como a propriedades pragmático-discursivas próprias dessa construção (intenções dos interlocutores, registro (formalidade e informalidade), questões relativas ao gênero do discurso etc.). Quanto à prosódia, não há razões para acreditarmos que esta foge à regra de ligação simbólica existente entre os polos formais e funcionais, pois ela responde por uma gama de propriedades suprasegmentais pertencentes à especificação fonológica da construção. Além disso, é conhecimento amplamente difundido e aceito que traços prosódicos específicos podem evocar significados distintos em diferentes níveis de complexidade estrutural, isto é, do léxico ao discurso, conforme discutiremos na seção seguinte.

### 3 A interface entre prosódia e discurso

A prosódia, comumente usada como termo alternativo para *fonologia suprasegmental* (CRYSTAL, 2008), abarca uma série de fenômenos fonológicos na cadeia da fala como a duração (*duration*), a intensidade (*loudness*), a altura (*pitch*) e o ritmo. Entretanto, tais características não se restringem ao campo da linguística, podendo ser estudadas por áreas como a física acústica, a biologia e a musicologia (MOTTA MAIA, 1991). Em certas abordagens da fonologia suprasegmental, sobretudo aquelas dirigidas aos fenômenos de interface com o discurso, o termo *prosódia* é reservado para os traços prosódicos que se realizam no âmbito das sentenças, tais como entonação, padrões rítmicos, proeminência e tonicidade. A interface entre traços prosódicos específicos e padrões gramaticais torna a relação entre prosódia e discurso ainda mais evidente, uma vez que de modo geral, diferentes padrões prosódicos servem a propósitos comunicativos específicos e estão subjugados à intenção pragmático-discursiva daquele que comunica.

Na língua inglesa, embora Brazil (1980, 1985) tenha sido pioneiro no desenvolvimento de um arcabouço explanatório que equacionava prosódia e discurso, já havia, antes das referidas publicações, literatura que apontava para tal relação. Entre as principais publicações, destacamos o trabalho de Kenneth Pike (1945) que trata da prosódia do inglês americano fazendo especial menção ao modo como determinados padrões entoacionais são empregados na obtenção de nuances específicas de significado por parte dos falantes. Isto é, Pike descreve como a prosódia pode servir aos propósitos comunicativos dos falantes, que são capazes de alterar o significado pragmático de enunciados valendo-se da prosódia para dar forma a sua atitude e intenção. A obra do autor trata também, de modo especial, de como esses padrões podem ser utilizados no âmbito do ensino de inglês como língua estrangeira.

Além do trabalho de Pike, muitas outras publicações se seguiram a essa e trouxeram informações sobre como os falantes fazem uso de padrões entoacionais para exprimir significados diversos. Halliday (1967) e Chafe (1980) destacam como a prosódia afeta a distribuição informacional do discurso em informação velha e nova. Schiffrin (1987) pesquisa os marcadores discursivos do inglês *oh, well, now, then, you know* e *I mean*, bem como os conectivos *so, because, but* e *or*. A linguista propõe uma análise segundo a qual o significado linguístico dos marcadores é integrado com certas propriedades pragmáticas, dos quais fazem parte padrões entoacionais específicos, na obtenção de suas funções discursivas. No âmbito da análise da conversação, Cutler e Pearson (1986) analisam como padrões prosódicos podem ser utilizados como pistas na troca de turno entre interlocutores.

Na próxima seção trazemos dados da língua inglesa a fim de discutir uma questão de grande dificuldade para aprendizes de inglês como língua estrangeira, isto é, a proeminência.

## 4 Prosódia e discurso: proeminência e construções

De modo geral, proeminência pode ser definida como o nível de distinção de um segmento em relação aos demais em uma sequência (CRYSTAL, 2008). Deste modo, sílabas podem ser mais ou menos proeminentes em relação a outras, a depender de fatores já mencionados como duração, altura, intensidade e/ou sonoridade inerente<sup>1</sup> e como esses são empregados na definição da sílaba mais proeminente entre todas as presentes na palavra. No caso das sentenças, o mesmo processo de segmentação para a localização do elemento prosodicamente mais saliente se mantém, no entanto, ao passo que a palavra se segmenta em sílabas, os enunciados da fala devem ser segmentados, de acordo com a hierarquia prosódica de Nespor e Vogel (1986)<sup>2</sup>, no que os autores chamam de *frases entoacionais* (do inglês *intonational phrase*). Frases entoacionais, também chamadas de *thought groups* ou *intonation units* (Gilbert, 1983), caracterizam-se:

- I. Por ter uma estrutura interna semântica e gramaticalmente coerente;
- II. Pela delimitação da unidade a partir de pausas na fala;
- III. Por conter um elemento prosodicamente proeminente.

Por coerência semântica e gramatical, entende-se que as relações mantidas entre o significado (aspecto semântico) e uma sequência de palavras (sintaxe) encontram, na grande maioria das vezes, uma correspondência fonológica por meio de uma pausa. Isso não significa que todas as vezes que tivermos em adjacência duas unidades semântica e sintaticamente coerentes, elas devam ser separadas por uma pausa, mas o contrário tende a ser verdadeiro; isto é, a ocorrência da pausa

---

1 A sonoridade inerente está ligada a como a tonicidade é aplicada a certos itens por características próprias dos itens lexicais e não por regras fonológicas independentes. Segundo Celce-Murcia et al. (2010), além da classe gramatical a qual pertence a palavra, na língua inglesa a origem etimológica da palavra bem como sua constituição morfológica por prefixação ou sufixação são fatores determinantes na aplicação do acento primário. De modo geral, a aplicação do acento primário segue a regra da raiz anglo-saxã da palavra, como ocorre em NEV.er, FATH.er, MOTH.er etc. Casos de aplicação do acento primário de modo diferente ocorrem quando a palavra mantém a acentuação da língua original da qual emprestou a palavra (ba.ROQUE, bi.ZARRE do francês) ou quando a estrutura da palavra é alterada por afixação (emPLOYer vs employEE) (Celce-Murcia et al. 2010, p. 186).

2 Segundo a proposta de Nespor e Vogel (1986), os enunciados são o último nível de uma hierarquia prosódica que se estrutura da seguinte maneira: enunciados > frase entoacional > grupo clítico > palavra fonológica > pé > sílaba.

se dá prototipicamente entre elementos que possuem coerência semântica e sintática interna. Nas palavras de Roach (2009, p. 155):

É muito incomum encontrarmos uma fronteira de grupo entoacional em um local onde a única fronteira gramatical é a fronteira entre palavras. Soaria, por exemplo, excepcionalmente estranho se tivéssemos uma fronteira de grupo entoacional entre um artigo e um substantivo ou entre verbos auxiliares e principais quando estivessem adjacentes [...]³

Vejamos um exemplo, em (1), retirado de uma gravação de fala espontânea para ilustrar a discussão.

(1) and then nearer to the front || on the left | there's a bit of forest | coming down to the waterside || and then a bit of a bay (ROACH, 2009, p. 132)

As barras verticais duplas estão sendo usadas para caracterizar fronteiras entre frases entoacionais com uma pausa, ao passo que a barra única ilustra somente uma fronteira. Tomemos como exemplo de análise a unidade *on the left*. Como vemos, trata-se de um sintagma preposicional sintaticamente convencional [<sub>PP</sub> on [<sub>DP</sub> the [<sub>NP</sub> left]]] e que desempenha a função de adjunto adverbial do VP *there's a bit of forest*. Do ponto de vista semântico, o sintagma veicula parte da informação de localização espacial que se pretende narrar com a cena nucleada por *there's a bit of forest*. Portanto, fonologicamente, *on the left* pode ser caracterizado como uma frase entoacional, pois é delimitado por fronteiras semântico-sintática claras, mas não necessariamente por pausas.

Ainda sobre pausas que, aliás, compõem o segundo dos critérios elencados para a segmentação de frases entoacionais, acabamos de ver que critérios semântico-sintáticos podem ser utilizados como balizadores dos contextos em que elas ocorrerão. Entretanto, na grande maioria dos casos, a ocorrência de pausas dependerá de condições impostas pela intenção pragmático-discursiva dos falantes, ou seja, ainda que critérios formais possam ser usados na localização das pausas na cadeia da fala, a realização da pausa ainda dependerá fortemente das intenções comunicativas dos falantes. Por exemplo, observe o excerto de diálogo (2) e tente identificar onde poderiam ocorrer pausas.

3 Original: "It is very unusual to find a tone-unit boundary at a place where the only grammatical boundary is a boundary between words. It would, for example, sound distinctly odd to have a tone-unit boundary between an article and a following noun, or between auxiliary and main verbs if they are adjacent [...]"

(2) **Dave:** Do you remember when we used to stay up all night studying for the exams?

**Howard:** Do I ever! Finals week was so bad that we'd drink coffee by the gallon. (CELCE-MURCIA, 2010, p. 222)

Na fala de Dave, poderíamos ter de uma até três frases entoacionais a depender dos locais que podemos aplicar uma pausa. No caso de *uma* frase entoacional, o enunciado seria produzido de uma única vez sem que nenhuma pausa fosse colocada em sua estrutura interna. No caso de *três*, teríamos a seguinte segmentação:

(3) **Dave:** Do you remember (||) when we used to stay up all night (||) studying for the exams?

As barras com parênteses significam que apesar de poderem ocorrer nesses contextos, as pausas não são obrigatórias. Logo, o número de frases entoacionais e, conseqüentemente, de pausas dependerá do momento em que o enunciado for proferido e das intenções de quem o produzir. Já no caso da fala de Howard, a estrutura sintática do enunciado impõe restrições sobre a aplicação das pausas, havendo a necessidade de pausa em, pelo menos, dois locais:

(4) **Howard:** Do I ever! || Finals week was so bad || that we'd drink coffee (||) by the gallon.

Para notar a estranheza que causa a falta dessas pausas, tente pronunciar o enunciado de uma só vez e preste especial atenção na forma como essa realização atenta contra a interpretação de partes do enunciado. No primeiro caso, a ausência da pausa poderia causar uma estranheza semântica ao conduzir os participantes do diálogo, sobretudo o ouvinte, a uma tentativa de interpretação errônea do item *finals*. Não fosse pela presença do morfema *-s* do plural, o ouvinte poderia ser levado a interpretar o adjetivo *final* em um contexto de inovação em que a palavra estaria sendo empregada como um verbo. Essa interpretação é reforçada pela primeira sentença, que tem forma canônica de uma interrogativa. No segundo caso, temos o exemplo de uma subordinada adverbial consecutiva (*so* + adjective/adverb + *that*-complement) em que a oração encaixada é introduzida pelo complementizador opcional *that* que, se pronunciado, deve ser parte integrante da encaixada e não da oração matriz. Por fim, temos uma pausa opcional que pode segmentar o PP *by the gallon* como uma frase entoacional independente.

Como mostram os exemplos acima, em muitas ocasiões, a utilização da pausa depende dos propósitos comunicativos dos participantes do discurso. Roach (2009) exemplifica essa variabilidade inerente à pausa ao discutir o modo como

certos profissionais versados em oratória como, por exemplo, políticos e filósofos, podem se valer da pausa, até mesmo colocadas em contextos pouco prováveis, como estratégias comunicativas. Segundo o autor, a utilização de pausas em contextos improváveis abre menos possibilidades de que esses oradores sejam interrompidos em uma exposição. Celce-Murcia et al. (2010) também discutem como a pausa pode ser utilizada por políticos como uma estratégia comunicativa, mas nos exemplos dos autores, são mencionados os casos em que a pausa tem um objetivo de enfatizar as partes da mensagem de modo a torná-la mais clara e memorável. Esse é o caso da clássica frase do ex-presidente americano Barack Obama *Yes || we || can*.

Algumas outras utilizações da pausa por profissionais vão na direção contrária a daquela usada por políticos. Esse é o caso de algumas frases usadas no fim de comerciais de rádio e TV que informam os consumidores sobre efeitos colaterais do produto, no caso de medicamentos, ou das condições de oferecimento de promoções etc. Esse último é o caso do exemplo a seguir, retirado de uma campanha publicitária do McDonald's na divulgação do sanduíche McGriddle.

(5) || *McGriddle sandwich at participating McDonald's* ||

Em (5), as barras antes e depois da sentença estão sendo usadas para demonstrar que internamente não houve nenhuma pausa. Essa estratégia é usada de modo a poupar o tempo da publicidade com informações que não têm um propósito verdadeiramente publicitário.

Por fim, não podemos deixar de mencionar os casos em que a utilização da pausa se dá como estratégia de comunicação que afeta o nível semântico a ponto de alterar a mensagem pretendida de modo bastante drástico. Esse é o caso do excelente exemplo de Roach (2009, p. 160):

(6) i) four plus six divided by two equals five  
 ii) four plus six divided by two equals seven

Nesses casos, o local específico da pausa altera o resultado da equação. No exemplo (i), poderíamos representar a equação do seguinte modo  $(4 + 6) \div 2 = 5$ , ao passo que em (ii) teríamos  $4 + (6 \div 2) = 7$ . Ao representar as sentenças com as pausas teríamos o seguinte:

(7) four plus six || divided by two || equals five  
 four || plus six divided by two || equals seven

Dos critérios levantados na identificação das frases entoacionais, discutimos os dois primeiros, isto é, a coerência semântica e sintática de unidades para a delimitação das frases entoacionais e o papel da pausa nessa tarefa. O último critério, ou seja, “conter um elemento prosodicamente proeminente” pode soar como uma matrioska conceitual, uma vez que estamos justamente tentando definir o que é proeminência, mas como dissemos, a proeminência em nível discursivo depende de elementos hierarquicamente mais elementares.

Como já mencionado, o termo *proeminência* pode ser largamente utilizado para caracterizar sílabas prosodicamente mais salientes em relação a outras em uma palavra. No âmbito dos enunciados, a proeminência depende basicamente de três fatores para governar a sua ocorrência, isto é:

- i) a distribuição informacional nos enunciados;
- ii) partes informacionais enfáticas;
- iii) ênfase para efeito de contraste.

Além dos fatores acima, para os quais utilizaremos a discussão contida em Celce-Murcia et al. (2010), discutiremos brevemente alguns casos de construções específicas em que a parte formal das construções, isto é, sua especificação fonológica, exige uma descrição de sua prosódia. Em outras palavras, discutiremos algumas construções da língua inglesa que têm contornos prosódicos (do inglês *prosodic contour*) específicos.

#### 4.1 Proeminência e a estrutura informacional

Como já mencionado diversas vezes até aqui, a relação entre prosódia e discurso é uma relação íntima e quase interdependente em alguns casos. No que se refere à estrutura informacional das sentenças, Chafe (1980 *apud* CELCE-MURCIA et al. 2010) traça uma correlação importante entre o status prosódico de certas frases entoacionais e sua estrutura informacional que não podemos deixar de mencionar. O linguista americano chama atenção para o fato de que palavras que carregam informações velhas ou que são semanticamente previsíveis, também chamadas de **tema** ou **tópico**, são marcadas por alguns traços prosódicos específicos, como a ausência de acento, e são frequentemente pronunciadas com um baixo *pitch* (altura). Por outro lado, informações novas e semanticamente imprevisíveis são marcadas pela presença de um **acento marcado** (do inglês *strong stress*) e por uma realização mais vigorosa do *pitch*. A informação nova normalmente recebe o nome de **rema** ou **foco**. Para ilustrar essa relação, observemos excertos de diálogos, mas como proposto em outras circunstâncias, antes de discutirmos os dados, tente realizar a distribuição de informações velhas e novas neste diálogo.

(8) A: Julia made us tacos.

B: Chicken tacos?

A: That's right. Chicken tacos, with green peppers and onions. Red onions

Em (8), a análise dos enunciados nos mostra que o falante A oferece a B a informação, nova nesse caso, de que Julia preparou *tacos* para ambos. A simples menção do nome Julia, sem nenhuma explicação mais específica, nos faz entender que essa pessoa é alguém conhecida por ambos os participantes do diálogo. Logo, a sua identidade é previsível e já conhecida. A informação nova, nessa caso, recai sobre o predicado da sentença representado por *made us tacos*. Em termos prosódicos, a primeira sílaba de *tacos* receberá a proeminência dessa frase entoacional. Seguindo com o diálogo, B faz uma pergunta sobre o tipo de *taco* que Julia preparou. Ao dizer *chicken tacos?*, o falante dá à palavra *tacos*, que no turno anterior havia sido analisada como a porção nova de informação, um status de informação velha, já partilhada, e o acento, nesse caso, recai sobre o ingrediente de que os *tacos* são feitos, isto é, *chicken*. Por fim, temos uma descrição mais detalhada por parte de A a respeito do prato feito por Julia. Nesse momento, B já está ciente de que Julia preparou algo, que esse algo é *taco* e, com a expressão *that's right* de A, o ingrediente principal do *taco* é esclarecido, isto é, *chicken*. Nessa parte do diálogo, as informações novas ficam por conta dos ingredientes que acompanham o *chicken*, ou seja, *green peppers* e *onions*. O interessante é notarmos que o próprio A transforma uma informação recentemente oferecida como nova em informação velha no instante em que especifica o tipo de *onion* utilizada no prato. Nesse momento *onion* acaba de se transformar em informação já compartilhada e *red*, o tipo de cebola usada, passa a ser a porção nova de informação. Desse modo, o diálogo terá a seguinte realização prosódica (as partes em caixa alta exemplificam as partes proeminentes):

(9) A: Julia made us TAcos.

B: CHICKEn tacos?

A: That's right. Chicken tacos, with GREEN PEPPERS and ONIONS. RED onions

Discutidos os casos em que a estrutura informacional e a prosódia operam conjuntamente nos enunciados, passamos a seguir aos casos de partes informacionais enfáticas e contrastivas.

## 4.2 Acento enfático e contrastivo

Embora não tenham exatamente as mesmas razões de ser ou aplicações no discurso, trataremos do **acento enfático** (do inglês *emphatic stress*) e **contrastivo** (do inglês *contrastive stress*) conjuntamente por entendermos se tratar de fenômenos de natureza correlata.

O acento enfático, como o próprio nome diz, é utilizado como uma estratégia de ênfase que o falante pode utilizar para atribuir a dada parte informacional de seu enunciado. Segundo Celce-Murcia et al. (2010), há uma certa confluência entre a realização de acentos enfáticos e remas, isto é, na maioria dos casos em que o falante atribui ênfase prosódica a uma parte da sentença, essa traz informação nova. Porém, em casos de acento enfático propriamente ditos, a força de ênfase e proeminência pode ser maior do que aquela exibida em partes que simplesmente possuem informação nova. Vejamos um exemplo em (10).

- (10) A: What do you think of this dress?  
B: It fits you.

A depender da parte enfatizada por B, diferentes informações podem ser veiculadas. Suponhamos que B tivesse enfatizado *fits*. Nesse caso, B talvez estivesse querendo enaltecer o quão apropriado o corte do vestido era para A. Dito de outro modo, B enfatiza que o vestido acomoda-se de maneira adequada às formas do corpo de A. Se porventura B decidisse enfatizar *you*, talvez quisesse comunicar que o vestido parece ter sido feito exclusivamente para A. Com essa ênfase, entendemos que o vestido poderia não ser apropriado para ninguém além da pessoa de A. Assim, poderíamos ter a seguinte configuração:

- (11) A: What do you think of this dress?  
B: It FITS you / It fits YOU.

É interessante notarmos que o português e o inglês podem usar estratégias diferentes de comunicação no que se refere à atribuição de ênfase. Veja o exemplo em que dois interlocutores trocam agradecimentos.

- (12) A: THANK you.  
B: Thank YOU.

A troca de agradecimentos acima sugere que B é, na verdade, aquele que se sente favorecido por uma ação desempenhada por A de modo que ele é aquele que deve agradecer. Em português, temos alternativas diferentes, em (13) e (14).

- (13) A: Obrigado.  
B: Obrigado voCÊ.

- (14) A: Obrigado.  
B: EU que agradeço.

É interessante notarmos que, em português, essa troca de agradecimentos também se vale de estratégias prosódicas para atribuir ênfase, no entanto, diferentemente do inglês, o português também faz uso de outras estratégias lexicais, com a inclusão da palavra *voçê*, ou até mesmo diferentes estratégias fraseológicas, com o uso de um fraseologismo próprio desse contexto comunicativo, isto é, *eu que agradeço*.<sup>4</sup>

O segundo tipo de acento de que trataremos é o contrastivo. Como explicitado pelo nome, trata-se de uma estratégia prosódica que o falante usa para contrastar dois ou mais pontos informacionais no enunciado, como vemos em (15):

- (15) Was it the RED or the YELLOW pepper that she used?

Na sentença de (15), o falante atribui proeminência por meio do acento em duas partes do enunciado a fim de contrastá-los e de obter a informação desejada. Nesse caso, as partes em contraste estão explícitas no enunciado, mas nem sempre o acento contrastivo ocorre dessa forma, como podemos ver em (16):

- (16) I promise I didn't tell Mary you ate the last piece of chocolate cake.

Nessa oração, há uma série de possibilidades de ênfase por meio do acento contrastivo, mas em nenhum dos casos, o elemento com o qual a parte proeminente se contrasta está explícita no enunciado. Vejamos algumas das possibilidades.

- (17) i. I promise I didn't tell MARY you ate the last piece of chocolate cake.  
ii. I promise I didn't tell Mary YOU ate the last piece of chocolate cake.  
iii. I promise I didn't tell Mary you ate the LAST piece of chocolate cake.

Em (17i) a ênfase recai sobre o objeto de *tell*, isto é, *Mary*. Com essa proeminência, o contraste sugere que o enunciador jura não ter contado para *Mary*, mas não garante que ninguém além dela tenha sido informado. No caso de (17ii), o enunciador coloca o elemento *you* como proeminente, o que sugere que alguém

4 Há a possibilidade de utilização de uma estratégia sintática de clivagem do sujeito para efeitos enfáticos como em *sou eu que agradeço*.

tenha comido o bolo, mas que a identidade de quem o fez tenha sido mantida em segredo. Por fim, em (17iii) o enunciador sugere, com a proeminência posta em *last*, que ele contou a Mary que seu interlocutor comeu o bolo, mas nenhuma referência foi feita ao fato de se tratar da última fatia. Em todos os casos, os elementos proeminentes estão em contraste com um elemento que deve ser inferido pelo interlocutor. Desse modo, diferentemente do primeiro tipo de contraste, aquele em que os elementos em contraste estão explícitos no enunciado, nos casos de (17i) a (17iii), temos um contraste implícito.

Como dissemos no início desta seção, tínhamos o objetivo de discutir os três fenômenos que governam a atribuição de proeminência, isto é, a distribuição informacional em informação velha (tema) e nova (rema), a utilização de acento enfático e a atribuição de acento contrastivo. Para terminarmos, mencionaremos brevemente dois tipos de construções em inglês em que a proeminência parece ser parte da estrutura formal dessas construções.

### 4.3 Construções exclamativas e clivagem

O inglês, assim como todas as línguas, dispõe de construções gramaticais em que contornos prosódicos específicos são parte formal das construções; entretanto, embora tais características prosódicas possam ser evidentes para falantes proficientes da língua, o mesmo não pode ser dito de aprendizes de uma língua estrangeira, principalmente se pensarmos que na grande maioria das vezes, os materiais destinados ao ensino de gramática de língua estrangeira nem sempre fazem menção explícita a questões relativas à prosódia de certas construções. É com base nessa ideia que expomos abaixo duas construções da língua inglesa em que a prosódia é marcadamente saliente. A primeira delas, chamada de *Antitopic Exclamative* por Michaelis e Lambrecht (1996), é exemplificada em (18).

- (18) GOD it's hot.  
      MAN I'm tired.  
      DAMN you're good.

Nas construções exclamativas em (18), uma interjeição localizada no início da oração, que denota uma espécie de avaliação da parte do enunciador, recebe o acento focal de proeminência ao passo que o resto da sentença que segue é caracterizada com menor força prosódica. Assim como nas exclamativas, em que uma posição sintática está associada ao foco prosódico, nas construções clivadas (do inglês *cleft constructions*), há elementos sintáticos que devem obrigatoriamente ser realizados com um acento focal específico. Vejamos alguns exemplos.

(19) It is the wife who decides.

What I want is a gin and tonic. (HILPERT, 2014, p. 113)

No primeiro tipo de clivada, que Hilpert (2014) chama de *It-Cleft Construction*, o argumento predicativo de *it's*, isto é, *wife* é clivado entre a cópula *is* e a oração relativa introduzida por *who*. A posição em que *wife* se encontra é prototipicamente uma posição de foco e como tal deve receber proeminência prosódica por meio de um acento mais marcado em relação ao resto da sentença. Sendo assim, poderíamos representar a sentença prosodicamente como *It is the WIFE who decides*.

No segundo tipo de clivagem, a *Wh-Cleft Construction*, a organização da informação se dá de forma a interpretarmos a oração relativa *What I want* como a parte da oração que expressa a pressuposição pragmática *o falante quer algo*. A relativa é seguida de uma cópula *is* que serve aos propósitos de ligar a *Wh-cleft* ao seu argumento predicativo, ou seja, *a gin and tonic*. A esse argumento, chamado por Hilpert de *focus phrase*, atribui-se a proeminência prosódica da sentença. Portanto, a representação dessa construção com a proeminência prosódica seria *what I want is a GIN and TONIC*.

Nos casos muito brevemente discutidos acima, bem como em uma gama de outros contextos construcionais, a prosódia pode ser analisada como uma parte integrante do aspecto formal das estruturas.

## 5 Conclusão

Neste capítulo introdutório sobre a interface entre prosódia e discurso na língua inglesa, iniciamos a nossa discussão com o objetivo de localizar o lugar da prosódia na arquitetura geral da gramática. Com base no arcabouço geral da linguística cognitiva, perspectiva da linguagem fortemente enraizada no uso linguístico, entendemos que não há razões para acreditarmos que a prosódia não possa ser vista como um aspecto formal que tem fortes relações com a função que certas construções desempenham no discurso. Em seguida, realizamos uma discussão sobre a relação entre prosódia e discurso a fim de criar um contexto para a análise de alguns dados da língua inglesa que são relevantes para tal interface. Nesse âmbito, analisamos principalmente os contextos em que ocorre proeminência prosódica por meio de uma discussão mais detalhada da relação entre prosódia e a distribuição da informação em temas e remas, assim como os tipos de acento, enfático e contrastivo, e os contextos em que eles ocorrem. Por fim, oferecemos uma breve discussão a respeito de duas construções, exclamativas e clivadas, em que a proeminência prosódica deve ser analisada como parte integrante da interface formal e funcional dessas estruturas.

## Referências

- Boas, H. C. “Cognitive Construction Grammar.” In: *The Oxford Handbook of Construction Grammar*, por T., TROUSDALE, G. HOFFMANN, 233 - 252. Nova York: Oxford University Press, 2013.
- Brazil, D. *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press, 1985.
- Brazil, D., Coulthard, M. e Johns, C. *The Communicative Value of Intonation in English*. London: Longman, 1980.
- Bybee, J. L. “From usage to Grammar: The Mind’s Response to Repetition.” *Language*, n. 82 (2006): 711 - 733.
- Bybee, J. L. “Usage-based Theory and Exemplar Representations of constructions.” In: *The Oxford Handbook of Construction Grammar*, por T., TROUSDALE, G. HOFFMANN, 49 - 69. Nova York: Oxford University Press, 2013.
- Celce-Murcia, M.; Brinton, D. M.; Goodwin, J. M. *Teaching Pronunciation – A course book and reference guide*. New York: Cambridge, 2010.
- Chafe, W. *The pear stories: cognitive, cultural and linguistics aspects of narrative production*. Norwood (NJ): Ablex, 1980.
- Croft, W.; Cruse, A. *Cognitive Linguistics*. Cambridge: Cambridge University Press, 2004.
- Crystal, D. *A Dictionary of Linguistics and Phonetics*. 6 ed. Oxford: Blackwell, 2008.
- Cutler, A., & Pearson, M. On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (Ed.), *Intonation in discourse* (pp. 139–155). London: Croom Helm, 1986.
- Gilbert, J. B. *Pronunciation and listening comprehension*. *Cross Currents*, 10(1), 53-61, 1983.
- Goldberg, A. E. *A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press, 1995.

- Goldberg, A. E. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press, 2006.
- Halliday, M.A.K. *Intonation and Grammar in British English*. The Hague: Mouton. 1967
- Hilpert, M. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press, 2014.
- Langacker, R. W. *Essentials of Cognitive Grammar*. New York: Oxford University Press, 2013.
- Langacker, R. W. *Foundations of Cognitive Grammar, Volume I*. Stanford, CA: Stanford University Press, 1987.
- Lakoff, G. *Woman, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press, 1987.
- Lakoff, G. & Johnson, M. *Metaphors we live by*. Chicago: Chicago University Press, 1980.
- Maia, Eleonora Motta. *No reino da fala: a linguagem e seus sons*. 3<sup>a</sup> ed. São Paulo: Ática, 1991.
- Michaelis, L. A.; Lambrecht, K. Toward a Construction-Based Model of Language Function: The Case of Nominal Extraposition. *Language* 72:215-247. 1996.
- Nespor, M.; Vogel, I. *Prosodic Phonology*, Foris, Dordrecht, 1986.
- Paradis, C. "Is the notion of linguistic competence relevant in Cognitive Linguistics?" *In Annual Review of Cognitive Linguistics* (Amsterdam: Benjamin), 2003: 247 - 271.
- Pike, K. *The intonation of American English*. University of Michigan Publications in Linguistics, 1945.
- Roach, P. *English Phonetics and Phonology – a practical course*. 4 ed. Cambridge: Cambridge University Press, 2009.
- Schiffrin, D. *Discourse markers*. Cambridge: Cambridge University Press, 1987.

## O ensino de prosódia nas aulas de espanhol como língua estrangeira

Cristiane Conceição Silva  
Universidade Federal da Bahia

### 1 Introdução

Normalmente, o sonho de quase todo aprendiz de uma língua estrangeira é dominar, como se fosse um nativo, a pronúncia (tanto no nível segmental quanto prosódico) da língua que está aprendendo. Esse aprendiz, porém, quando começa a participar em situações reais de conversação com falantes nativos fora da sala de aula logo percebe que se conseguir se comunicar e se, cada vez menos, os interlocutores pedirem que ele repita palavras ou frases já terá, pelo menos, avançado um grande passo.

Quando se trata de brasileiros aprendendo espanhol como língua estrangeira, a situação não é muito diferente. Porém, talvez por causa de uma falsa crença de que a pronúncia do espanhol é fácil (POCH OLIVÉ, 2004), este campo tenha recebido pouca atenção no ensino como língua estrangeira. Uma evidência dessa deficiência pode ser comprovada, por exemplo, pela simples observação de materiais didáticos utilizados no Brasil para o ensino de espanhol como língua estrangeira (doravante espanhol/LE). Neles, há poucas seções dedicadas ao ensino de pronúncia dos segmentos e praticamente não há referências ao ensino da prosódia.

Por isso, podemos dizer que o ensino de pronúncia do espanhol/LE ainda é uma *asignatura pendiente* no Brasil, já que quando ele acontece, geralmente se ensina somente a pronúncia dos sons dessa língua. Poch Olivé (2004) ressalta que esse tipo de método de ensino que se concentra apenas na pronúncia dos sons da língua estrangeira desconsiderando a prosódia é muito simplista, pois assume uma falsa ideia de que os sons são produzidos isoladamente quando na verdade são produzidos de forma concatenada. Sabemos que a fala concatenada está intrinsecamente sujeita à coarticulação e, além disso, nela está integrada a prosódia.

Com isso, a prosódia é colocada em segundo plano e pouco ou quase nada se ensina sobre ela. Assume-se que sua aprendizagem acontecerá naturalmente somente a partir da escuta dos modelos de fala apresentados em sala de aula. Por isso, o

aprendiz brasileiro de espanhol costuma acreditar que, se praticar bastante a pronúncia dos sons do espanhol, então, dominará a pronúncia dessa língua estrangeira como um todo, e somente quando é colocado em situação de imersão, começa a perceber o quanto a pronúncia em uma língua estrangeira vai além da produção segmental. Nesse contexto, as dificuldades com a prosódia se tornam mais evidentes e podem, no pior dos casos, afetar a comunicação (um ouvinte espanhol não entender que um brasileiro está fazendo uma pergunta, por exemplo (SILVA, 2009)), contribuir diretamente para a percepção de maior grau de sotaque estrangeiro (SILVA, 2016), ou ainda, em casos menos graves, colaborar na formação de estereótipos (os brasileiros falam de forma melosa, estão sempre animados e alegres, etc.).

Este capítulo tem por objetivo ajudar a preencher esta lacuna ao propor algumas atividades que possam ser aplicadas pelo professor de espanhol/LE em sala de aula já que, apesar do espanhol ser a segunda/terceira língua mais aprendida no mundo<sup>1</sup>, a oferta de materiais didáticos para ensino da prosódia do espanhol ainda é muito pequena. Nesse sentido, os trabalhos de Poch Olivé (1999), Gil Fernández (2007), Cortés (2002), Lahoz (2012) e Rico (2012) são inovadores pois, além de abordarem a temática do ensino de prosódia do espanhol/LE, também apresentam exercícios.

Como os materiais anteriores estão direcionados, principalmente, para o professor de espanhol que é também falante nativo dessa língua, neste capítulo, reproduzimos alguns desses exercícios propostos pelos autores, porém, procuramos, sempre que possível, adaptá-los para o contexto brasileiro em que boa parte dos professores de espanhol/LE são também brasileiros e, portanto, não falantes nativos de espanhol. Além disso, como os aprendizes têm o português brasileiro (doravante PB) como L1, realizamos também algumas adaptações a fim de tornar os exercícios mais efetivos para este público específico.

Assim, neste capítulo, primeiramente, definimos o termo prosódia e apresentamos os fenômenos prosódicos abordados neste trabalho. Em seguida, fazemos uma pequena discussão teórica sobre cada um desses fenômenos e, finalmente, apresentamos algumas propostas voltadas para o ensino.

## 2 Definição de prosódia

O professor de espanhol/LE pode não estar muito familiarizado com o termo prosódia, já que, em muitos trabalhos, o termo é utilizado como sinônimo de

---

1 De acordo com primeiro informe Berlitz publicado sobre o estudo do espanhol no mundo, elaborado em 2005, o espanhol seria a terceira língua mais aprendida no mundo. O informe publicado pelo Instituto Cervantes em 2016, afirma que ainda não estão disponíveis dados exaustivos que analisem, de forma comparativa, o número de estudantes das diversas línguas do mundo. Apesar disso, estima-se que o espanhol esteja disputando com o francês e com o mandarim a segunda posição de língua mais estudada no mundo. Disponível em: <http://www.cervantes.es/imagenes/File/prensa/EspanolLenguaViva16.pdf>

suprassegimento<sup>2</sup>. Por isso, antes de apresentar os aspectos da prosódia que serão discutidos neste capítulo, faz-se necessário defini-lo e determinar quais os aspectos dessa ampla área de estudo serão tratados aqui.

Couper-Kuhlen (1986) traça a trajetória histórica do termo desde sua criação até os dias atuais. Segundo a autora, o termo prosódia (προσωδία) foi criado pelos gregos e se referia ao acento tonal e/ou melódico das palavras lexicais. Posteriormente, foram introduzidas as prosódias, símbolos ortográficos que refletiam os acentos tonais. Assim, em grego antigo, uma sílaba que portava uma prosódia aguda era pronunciada com tom alto, por exemplo. Dessa forma, a prosódia estava associada a traços melódicos da língua falada.

Ao longo do tempo, o significado do termo sofreu modificações. No século II, passou a se referir a traços que não eram expressos na sucessão segmental de vogal e consoante, como o alongamento vocálico, por exemplo e, em algum momento, os acentos tonais foram substituídos pelo acento de palavra. Quando isso aconteceu, prosódia passou apenas a denotar distinções de acento de palavra. Com isso, a partir do século XV, prosódia adquiriu o significado de “versificação”.

Foi somente com a publicação do artigo de Firth (1948) que o significado original de prosódia relacionada à melodia foi retomado. Enquanto a maioria dos linguistas estavam preocupados exclusivamente com as análises dos segmentos individuais (análise fonêmica), Firth (1948) argumentou que os linguistas deveriam concentrar suas análises nos traços que se estendem sobre porções de um enunciado, o que ele denominou análise prosódica.

Atualmente, o termo é amplamente utilizado e recobre uma gama variada de fenômenos. Assim, de forma bastante simplificada, poderíamos definir prosódia como as propriedades que vão além da descrição de segmentos individuais, ou seja, trata-se do estudo<sup>3</sup> do sinal de fala usado para transmitir informações sobre a estrutura e significado de enunciados ou porções maiores que o enunciado. Sendo assim, estudos, por exemplo, que tratam da entoação, ritmo, acentuação, pausas, taxa de elocução, etc., pertencem à prosódia. Neste capítulo, discutimos a taxa de elocução, o ritmo e a entoação.

Primeiramente, apresentamos uma pequena definição teórica para cada fenômeno prosódico. Em seguida, apresentamos exercícios práticos que os professores de espanhol/LE podem incluir em sua prática cotidiana em sala de aula.

---

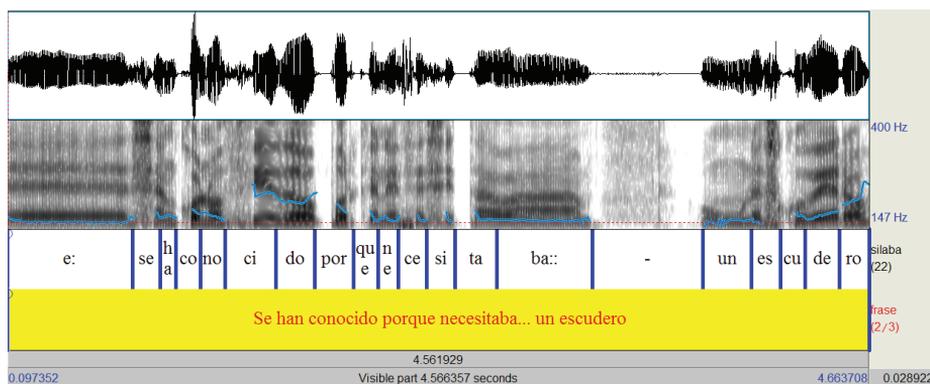
2 O termo foi criado pelo estruturalismo americano para designar traços distintivos que, diferentemente do fonema, não podem ser segmentados individualmente dos enunciados linguísticos. Assim, fenômenos de juntura, acento de palavra ou de frase, *pitch*, entoação, etc. são fenômenos suprassegmentais (BUSSMANN, 1996).

3 Do ponto de vista articulatório, acústico ou perceptivo.

### 3 Taxa de elocução

Em muitos trabalhos, **taxa de elocução** é denominada **velocidade de fala**. Preferimos utilizar aqui o termo taxa de elocução, pois velocidade de fala traz implícita a ideia de que os articuladores poderiam estar se movimentando com maior ou menor velocidade, o que não acontece concretamente. Sabemos que na produção da fala o que temos não é uma articulação mais ou menos veloz, mas uma maior ou menor sobreposição dos gestos articulatórios. Além disso, como assinala Gil Fernández (2007), a percepção da “velocidade” da fala está diretamente relacionada com a quantidade e a duração das pausas produzidas pelo falante. Sendo assim, quanto menor o número e a duração das pausas silenciosas ou preenchidas produzidas pelo falante, mais sua fala será percebida como rápida.<sup>4</sup>

Sendo assim, definimos a taxa de elocução pela divisão do número de unidades linguísticas<sup>5</sup> pronunciadas por um falante em uma determinada unidade de tempo<sup>6</sup> e a duração total desse intervalo. Em seu cálculo, tanto o conteúdo segmental, como as pausas (silenciosas<sup>7</sup> e preenchidas), são considerados (LAVIER, 1994). A Figura 1 ilustra uma das possibilidades de cálculo da taxa de elocução em um enunciado.<sup>8</sup>



**Figura 1** Forma de onda, espectrograma e segmentação em sílabas do enunciado em espanhol/LE *e: se han conocido porque necesitaba... un escudero*, pronunciado pelo informante brasileiro Br-1<sup>9</sup>.

- 4 Assim como demonstrado por Goldman-Esler (1961), quanto mais fluida é a fala, mais raras são as pausas, o que leva a percepção de que a fala é mais rápida. Por outro lado, quanto maior o número de pausas de hesitação, mais a fala é percebida como lenta.
- 5 A unidade também pode ser o segmento, a sílaba, a palavra, etc.
- 6 Normalmente, a medida adotada é o segundo.
- 7 Se no cálculo da taxa excluirmos as pausas, então, trata-se do cálculo da **taxa de articulação**.
- 8 Todas as figuras e análises acústicas deste trabalho foram realizadas no programa Praat. Esta ferramenta é bastante utilizada para análises acústicas da fala, mas também pode ser utilizada pelo professor de espanhol/LE tanto em sala de aula como fora dela. O programa pode ser baixado gratuitamente pelo site: [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html)
- 9 No momento da gravação, o informante residia em Madri há quatro anos e não havia estudado espanhol no Brasil antes de viajar para a Espanha.

Na Figura 1, temos o enunciado *e:: se han conocido porque necesitaba... un escudero* em espanhol/LE segmentado em sílabas. Se considerarmos a unidade de análise a sílaba, para calcular a taxa de elocução deste enunciado, basta contar o número total de sílabas pronunciadas (19) e dividir pela duração total do enunciado (4,566357 segundos), o que inclui tanto as pausas silenciosas quanto preenchidas. Dessa forma, a taxa de elocução do enunciado da Figura 1 será de 4,16 sílabas por segundo (4,16 sílabas/s).

Gut et al. (2007) chamam a atenção para o fato de que, normalmente, quando começamos a estudar uma língua estrangeira, temos a impressão de que os falantes nativos falam “rápido demais”. O interessante é que não temos essa mesma percepção com relação à nossa própria língua materna ou com relação às línguas estrangeiras que já dominamos. Geralmente, à medida que nosso nível de proficiência em uma língua estrangeira aumenta, tendemos a percebê-la como sendo enunciada cada vez mais “devagar”, se comparada a outras línguas com as quais temos pouca ou nenhuma familiaridade.

Para Osser e Peng (1964), isso acontece porque, quando escutamos uma língua estrangeira que não dominamos, tentamos prestar atenção não apenas à distribuição da fala, mas também das pausas ao longo do tempo. A dificuldade surge porque o ouvinte ainda não consegue perceber as pausas de menor duração e, por essa razão, tem a impressão de que a fala é enunciada de forma muito rápida. Por isso, à medida que o conhecimento da língua estrangeira aumenta, o aprendiz se familiariza mais com as unidades na fala e passa a avaliar adequadamente a taxa de elocução.

Estudos anteriores demonstram a importância da taxa de elocução para percepção da “velocidade” da fala, mas não é apenas do ponto de vista perceptivo que a taxa de elocução é importante. Do ponto de vista da produção, a taxa de elocução também pode ser utilizada como um indicador da proficiência ou fluência oral do falante. Lennon (1990), por exemplo, analisou quantitativamente as produções de quatro alemães aprendizes de inglês. Para isso, gravou em duas seções distintas a narrativa de uma história. A primeira gravação, realizada na 2ª semana após a chegada dos aprendizes alemães na Inglaterra, e a segunda gravação, realizada na 23ª semana, revelou que a taxa de elocução<sup>10</sup> aumentou significativamente entre a primeira e a segunda gravação. Além disso, o autor comprovou que a taxa de elocução é uma das variáveis temporais que está correlacionada com os julgamentos dos falantes nativos a respeito do grau de fluência.

Por tudo isso, é fundamental que o professor de espanhol/LE inclua em suas aulas exercícios que possam ajudar os aprendizes a melhorar tanto sua percepção como sua produção da taxa de elocução em espanhol.

---

10 O autor analisou também outras variáveis temporais como a duração das pausas silenciosas, preenchidas, etc.

### 3.1 Práticas – Taxa de elocução

Como vimos na seção anterior, a taxa de elocução tem um papel muito importante para a percepção e produção em língua estrangeira. O professor pode utilizá-la como uma ferramenta para quantificar o grau de fluência dos seus alunos, já que está correlacionada ao nível de proficiência avaliado pelos ouvintes nativos (GUT, 2007). Ademais, pode também considerá-la em atividades com fins didáticos. Tais atividades podem ser realizadas com o objetivo de afinar a percepção dos alunos para as diferenças de taxa de elocução em PB em comparação com às distintas variedades do espanhol. Além disso, pode também incluir atividades de produção oral voltadas para aumentar a taxa de elocução dos alunos ou ajustar a taxa de elocução às diferentes situações comunicativas<sup>11</sup> ou aos distintos estados afetivos.<sup>12</sup>

Como sugere Gil Fernández (2007), seria interessante que o professor propusesse atividades para os alunos, sempre que possível, seguindo a seguinte ordem: (1) atividades de percepção, (2) atividades de produção oral controlada e (3) atividades de produção oral livre. Assim, as propostas apresentadas aqui sempre irão partir da percepção em direção à produção livre.

#### **Atividade 1 ¿Quién habla más rápido?<sup>13</sup> (Percepção)**

<b>Material</b>	Gravações de áudio de personagens diferentes do rádio, da televisão ou da internet, identificados como A, B, C, etc.
<b>Requisitos</b>	Escolher gravações cujas diferenças de taxa de elocução sejam evidentes
<b>Público</b>	Estudantes de nível básico, intermediário ou avançado
<b>Procedimento</b>	<ol style="list-style-type: none"> <li>1. Escutar e identificar as personagens que falam mais rápido</li> <li>2. Ordenar as personagens das que falam menos e mais rápido</li> <li>3. Especular se as personagens que falam mais rápido têm menos pausas</li> </ol>

Como aponta a autora, o objetivo deste exercício é sensibilizar os alunos com relação às estruturas temporais da língua espanhola através da percepção das diferenças de taxa de elocução. Com relação às gravações utilizadas, nossa sugestão é que sejam selecionadas gravações de falantes diferentes em situações comunicativas diferentes, mas que pertençam à mesma variedade do espanhol.

11 No contexto familiar ou informal, a taxa de elocução normalmente é maior do que em situações formais de comunicação.

12 Normalmente, uma taxa de elocução maior está associada a sentimentos como a alegria ou a raiva, já taxa menores estão associadas com a tristeza, o desânimo, etc.

13 As duas atividades apresentadas nesta seção foram reproduzidas de Gil Fernández (2007).

Assim, o aprendiz poderá perceber que a taxa de elocução pode variar de acordo com o contexto conversacional.

O material utilizado nesta atividade pode ser, além das gravações de áudio, gravações de vídeo. Assim, o professor pode preparar uma primeira versão desta atividade com gravações de áudio e uma segunda versão com gravações de vídeo e, finalmente, discutir com os alunos em quais gravações a tarefa de identificação da taxa de elocução resulta mais fácil. Essa variação pode ser útil para verificar se os gestos que acompanham a fala podem servir também como pistas para a identificação da taxa de elocução.

Ainda com relação ao material utilizado, em outra versão do exercício são utilizadas amostras de fala pertencentes às diferentes variedades do espanhol, pertencentes ou não ao mesmo contexto comunicativo. Este tipo de gravação pode ser útil não só para ajudar na percepção da taxa de elocução nas distintas variedades do espanhol, mas também para sensibilizar os alunos de que a percepção pode estar condicionada não à maior ou menor “velocidade” da fala, mas ao maior ou menor contato que o aprendiz tem com a variedade apresentada. Por isso, como o aprendiz brasileiro normalmente tem mais contato com as variedades mexicana, argentina e centro norte peninsular espanhola, esta é uma excelente oportunidade para apresentar gravações não apenas dessas variedades, mas principalmente de outras variedades do espanhol com as quais os estudantes estão pouco familiarizados como a chilena, a cubana, a peruana ou a venezuelana, por exemplo.

Com relação à terceira etapa de aplicação da atividade (especular se as personagens que falam mais rápido têm menos pausas), se os alunos tiverem dificuldade em realizá-la, o professor pode entregar a transcrição das gravações de cada personagem sem qualquer marca de pontuação ou letra maiúscula. O estudante terá, então, que escutar com atenção as gravações e marcar com traço os lugares onde ele acha que houve pausas. Em seguida, os estudantes podem comparar entre si as marcações feitas em suas respectivas transcrições e discutir suas respostas. Depois da discussão, o professor poderá mostrar os locais onde realmente houve pausa nas gravações. Finalmente, os alunos podem discutir se realmente as personagens que falam mais rápido apresentam ou não menos pausas.

### **Atividade 2 *Soy una persona de muchos oficios*<sup>14</sup> (Percepção e prática controlada)**

<b>Material</b>	O professor seleciona textos de diferentes gêneros, tais como conferência científica, sermão religioso, transmissão esportiva, etc., e entrega um texto diferente para cada aluno.
<b>Público</b>	Estudantes de nível intermediário ou avançado

(continua)

14 O título original da atividade é *Velocidad de elocución y fonostilo*.

## Atividade 2 *Soy una persona de muchos oficios* (Percepção e prática controlada) (continuação)

### Procedimento

Cada aluno deve estudar em casa previamente o fragmento que lhe foi atribuído para depois repeti-lo em voz alta em sala de aula utilizando a taxa de elocução que lhe parecer a mais adequada dependendo do tipo de discurso. Outros fatores prosódicos estão presentes, mas o professor dará maior ênfase para a taxa de elocução.

Assim como o exercício anterior, este exercício também busca sensibilizar os alunos sobre a importância não só da taxa de elocução, mas também de outros aspectos prosódicos. Para praticar a percepção, antes de aplicar esta proposta de produção oral, o professor pode utilizar uma das duas atividades que apresentamos a seguir. A primeira possibilidade seria replicar os mesmos procedimentos da atividade 1 - *¿Quién habla más rápido?*, mas utilizando gravações em vídeo de textos pertencentes aos mesmos gêneros que serão trabalhados na atividade oral proposta. Dessa maneira, os alunos poderiam utilizar tais gravações como referência para ajustar sua taxa de elocução e prosódia assim como os gestos e movimentos corporais que caracterizam cada gênero. Outra possibilidade, seria pedir para os próprios alunos procurarem vídeos na internet e escolherem um vídeo como modelo, ou seja, como um apoio para a preparação da apresentação oral. Assim, no dia da apresentação, cada aluno apresentaria seu vídeo escolhido previamente para todo grupo e, logo em seguida, faria sua performance. No final, os alunos poderiam discutir quem conseguiu se aproximar mais dos modelos escolhidos.

## 4 Ritmo

Adotamos a definição de ritmo proposta por Barbosa (1999), que o define como a variação a longo termo da duração percebida. Da mesma forma, Gil Fernández (2007) também concebe o ritmo como um fenômeno essencialmente perceptivo, fundamentado na repetição.

Foi a partir dessa percepção do ritmo como um fenômeno de repetição que muitos linguistas passaram a defender a ideia da isocronia<sup>15</sup> na fala (ABER-CROMBIE, 1967). Com isso, passaram a classificar as línguas basicamente em dois grupos: o grupo das línguas de ritmo acentual (produção de sílabas acentuadas em intervalos de tempo regulares) e o grupo das línguas de ritmo silábico (regularidade determinada pelas próprias sílabas que apresentam aproximadamente a mesma duração). No primeiro grupo, foram incluídas as línguas como o inglês, o alemão, o PB e no segundo grupo línguas como espanhol e o francês.

15 Isocronia como a repetição em intervalos de tempo regulares.

Atualmente, os estudos experimentais têm demonstrado que esta tipologia rígida não funciona efetivamente para nenhuma língua, pois é possível encontrar uma tendência para um ou outro tipo de ritmo, porém, não da forma determinística que os linguistas inicialmente formularam.

Por isso, mais do que se preocupar com esta discussão teórica a respeito do ritmo das línguas, o que realmente interessa para o professor de língua estrangeira, como aponta Gut et al. (2007), é que o ritmo seja entendido a partir de uma perspectiva mais intuitiva, ou seja, como batidas regulares, fazendo referência direta à música ou à poesia. Com isso, é mantida a ideia de que cada língua apresenta potencialmente um ritmo diferente e, portanto, mais importante do que encontrar uma definição definitiva para o ritmo, o professor pode se concentrar nos fatores concretos que podem afetar diretamente o ritmo da língua materna dos seus alunos e da língua estrangeira que eles estão aprendendo para, então, desenvolver atividades que possam ajudá-los a produzir enunciados ritmicamente “corretos” e evitar enunciados ritmicamente “incorretos”.

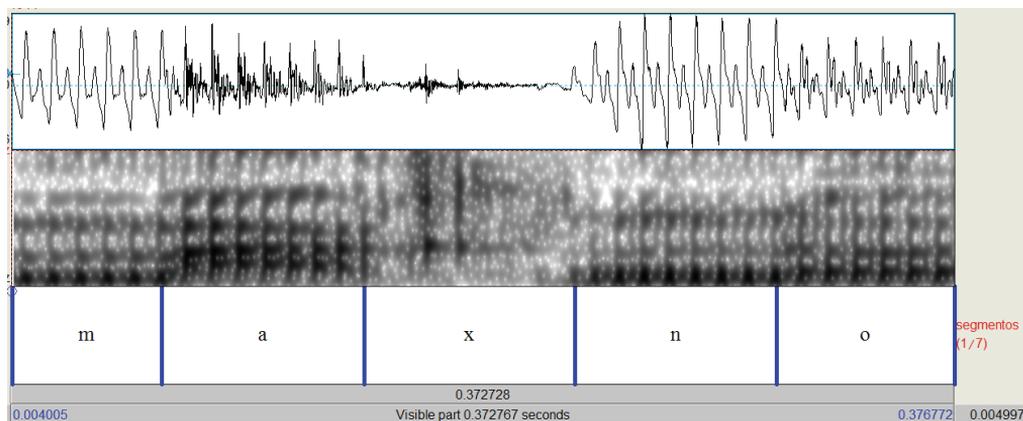
Gil Fernández (2007) e Lahoz (2012) apontam alguns dos fatores que podem afetar diretamente o ritmo de uma língua e sugerem que sejam sobre eles que o professor de espanhol/LE concentre sua atenção ao elaborar práticas que ajudem os alunos a melhorar sua percepção e produção do ritmo do espanhol/LE. O ritmo vai depender de diversos fatores, dentre eles: (1) a complexidade da estrutura silábica, (2) a posição mais comum do acento de palavra e o peso que a duração, a frequência e a intensidade recebem ao estabelecer a diferença entre sílabas tônicas e átonas, (3) a tendência ou não à redução vocálica e (4) a existência ou não de processos de ressilabificação. A seguir, discutimos cada um desses fatores ao contrastar o PB e o espanhol.

Com relação à complexidade da estrutura silábica, sabemos que as línguas geralmente diferem consideravelmente com relação ao tipo de sílaba que cada uma permite. No caso específico do espanhol e do PB, por serem línguas muito próximas, apresentam estruturas silábicas semelhantes. Sendo assim, tanto em espanhol quanto em português a sílaba canônica é a CV e ambas as línguas permitem estruturas silábicas como V, VV, CV, CVC, CVVV ou CCV, etc. As dificuldades do aprendiz brasileiro surgem, porém, com relação ao tipo de segmento que cada língua permite tanto na posição de ataque quanto na posição de coda silábica.<sup>16</sup>

---

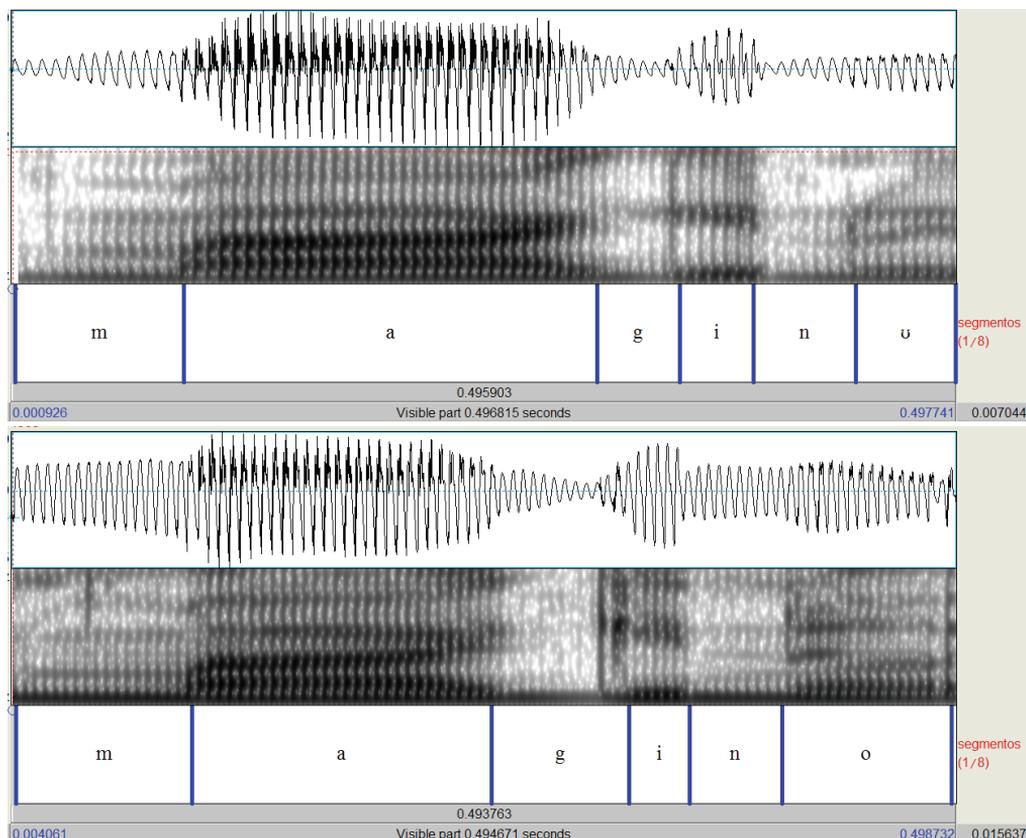
16 A sílaba pode ser dividida em unidades menores, sendo composta por um ataque (a(s) consoante(s) que precedem à vogal) e a rima, formada por uma vogal e qualquer consoante que venha depois dela. A vogal que pertence à rima é chamada de núcleo e a(s) consoante(s) que venham depois dela são definidas como coda. Dessa forma, uma sílaba como *trâns* da palavra *‘trânsito’*, seria composta pelas consoantes *‘tr’* na posição de ataque e a rima *‘ans’* seria composta pelo núcleo *‘a’* e as consoantes *‘ns’*, na coda.

Por exemplo, em posição de ataque silábico, as combinações de duas consoantes em que o segundo elemento é ‘s’ ou ‘t’ como na palavra ‘psicólogo’ são pronunciadas de maneiras diferentes nas duas línguas. Em espanhol, a tendência é a de não pronunciar o primeiro segmento [s]icólogo, enquanto em PB ocorre sempre a pronúncia de uma sílaba CV por meio da epêntese da vogal ‘i’. [pi]sicólogo, criando assim uma nova sílaba. Em posição de coda, ocorre um fenômeno semelhante. Na escrita, tanto o espanhol quanto o português apresentam sílabas CVC com oclusivas em posição de coda (Magno, apto, administrativo, etc.). Em espanhol, a pronúncia das oclusivas em coda sofre diversos processos de redução: pode se sonorizar, ser pronunciada como uma fricativa<sup>17</sup>, pode ser reduzida ou simplesmente não ser pronunciada; já em PB, porém, por ser uma língua que não permite outras consoantes em coda que não sejam exclusivamente ‘s’ e ‘r’, sempre cria uma nova sílaba por meio da epêntese da vogal ‘i’. Assim, podemos observar que, enquanto a pronúncia de um falante espanhol da consoante da coda na palavra *magno* foi realizada como uma fricativa (Figura 2), a mesma palavra pronunciada por uma informante brasileira tanto em PB quanto em espanhol/LE apresentou a vogal epentética ‘i’, resultando na pronúncia ma[gi]no nas duas línguas (Figura 3).



**Figura 2** Forma de onda, espectrograma, segmentação e transcrição fonética da palavra “Magno” lida na frase *¿Dónde están mis libros de Alejandro Magno y de Cid Campeador?* Pronunciada pelo informante espanhol Sp-1, de Madri.

17 Como acontece na variedade do espanhol falado na região de Madri.



**Figura 3** Forma de onda, espectrograma, segmentação e transcrição fonética da palavra *Magno* lida em português na frase *Onde estão meus livros do Alexandre Magno e do Cid Campeador?* (parte superior) e lida em espanhol na frase *¿Dónde están mis libros de Alejandro Magno y de Cid Campeador?* (parte inferior). Pronunciada pela informante brasileira Fer<sup>18</sup>.

Como podemos observar (Figura 3), normalmente, o aprendiz brasileiro de espanhol acaba transferindo as regras fonotáticas<sup>19</sup> do PB ao falar espanhol o que afeta diretamente o ritmo em espanhol/LE.

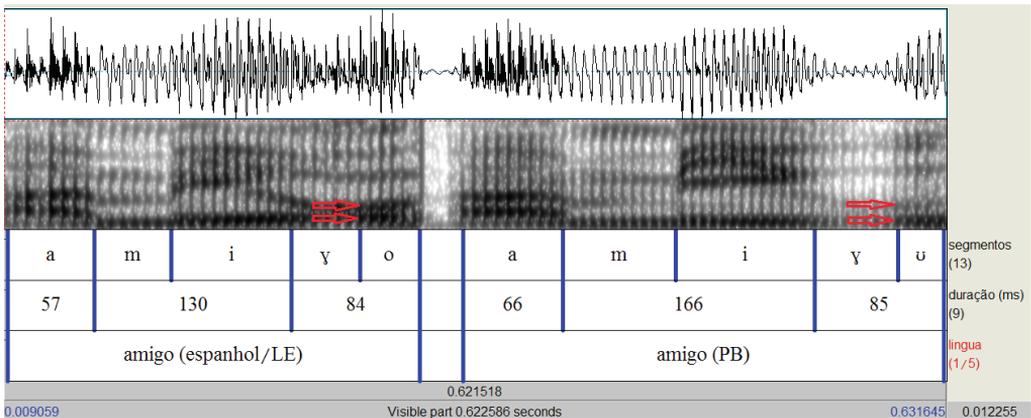
Quanto à posição do acento e seus parâmetros de marcação, o espanhol e o PB também divergem em alguns aspectos. A posição mais comum do acento de palavra, tanto em espanhol como em PB, é a penúltima sílaba da palavra. Além disso, ambas as línguas permitem que o acento recaia também na última ou na antepenúltima sílaba da palavra. A maior diferença, porém, entre as duas línguas,

18 No momento da gravação, a informante residia em Madri há um ano e havia estudado espanhol no Brasil por três anos antes de viajar para a Espanha.

19 Regras fonotáticas se referem ao conjunto de combinações de sons ou fonemas que uma determinada língua permite. (BUSSMAN, 1996).

está no peso atribuído a cada um dos parâmetros (frequência, duração e intensidade) na marcação do acento de palavra. Enquanto em espanhol, do ponto de vista perceptivo os parâmetros acústicos de frequência fundamental, juntamente com a duração e/ou intensidade são suficientes para a percepção do acento de palavra (LLISTERI et al. 2003b), em PB, os parâmetros acústicos mais importantes são, em primeiro lugar, a duração, seguida da intensidade. Somente quando as palavras estão em fronteira entoacional, a  $F_0$  também é importante para marcar o acento de palavra. (MORAES, 1987; MASSINI, 1991; BARBOSA, 1996; BARBOSA e ERIKSSON, 2013).

Outro aspecto que também afeta o ritmo é a redução vocálica que ocorre sistematicamente em posição postônica em PB, mas que em espanhol é muito rara. A redução vocálica também contribui para aumentar a diferença entre sílabas átonas e tônicas. Como consequência dessa diferença entre as duas línguas o aprendiz brasileiro tende a transferir o padrão acentual do PB para o espanhol o que afeta diretamente o ritmo do espanhol/LE, distanciando-o do espanhol falado como língua materna. Assim, para aperfeiçoar o ritmo do espanhol, o aprendiz brasileiro precisa utilizar, de forma mais sistemática, a variação da  $F_0$  para marcar o acento de palavra, precisa reduzir a duração das sílabas tônicas e, além disso, também evitar a redução vocálica nessa língua. Vejamos na Figura 4 um exemplo em que o falante brasileiro conseguiu diminuir a diferença de duração entre a sílaba tônica de uma palavra ao mesmo tempo em que conseguiu também diminuir o efeito da redução vocálica em espanhol/LE.



**Figura 4** Forma de onda, espectrograma, segmentação e transcrição fonética da palavra “amigo” lida em português na frase *Amigo Sancho, tens medo?* e lida em espanhol na frase *¿Amigo Sancho, tienes miedo?* Pronunciada pelo informante brasileiro Br-2.

Podemos observar na Figura 4 que, além da duração da sílaba tônica em PB ser maior que em espanhol/LE, a diferença entre a tônica, a pretônica e a postônica

também é maior em PB com relação à leitura da mesma palavra em espanhol/LE. Além disso, as flechas vermelhas indicam a localização do primeiro formante F1 e segundo formante F2<sup>20</sup> da vogal postônica nas duas línguas. Em espanhol/LE, a vogal apresenta F1 de 505 Hz e F2 de 1033 Hz enquanto que em PB as frequências são F1 de 349 Hz e F2 de 1045. Barbosa e Madureira (2015) apontam que os valores de F1 estão diretamente relacionados com a altura da mandíbula. Assim, quanto mais baixa está a mandíbula, mais altos serão os valores de F1 e quanto mais alta está a mandíbula menores serão os valores de F1. Com isso, podemos dizer, com base na observação da Figura 4 e no cálculo de F1, que a vogal [u] produzida em PB é mais reduzida que a mesma vogal [o] produzida em espanhol/LE.

Finalmente, o último fenômeno aqui tratado que também afeta o ritmo é a ressilabificação. Quando não há pausa entre duas palavras e a primeira palavra termina em consoante e a palavra seguinte começa com vogal, a consoante que pertencia à coda da sílaba anterior passa a ocupar a posição de ataque da sílaba seguinte que, de uma sílaba V, torna-se uma sílaba CV. Dessa forma, a frase: *En el mar oscuro veo las estrellas inolvidables*, seria pronunciada: e.nel.ma.ros.cu.ro.veo.la.ses.tre.lla.si.nol.vi.da.bles.<sup>21</sup>

Tanto em espanhol quanto em PB o fenômeno ocorre, mas o aprendiz brasileiro precisa aprender a aplicar sistematicamente o processo (inclusive quando a primeira palavra termina com ‘n’, contexto inexistente em PB). Para isso, precisa evitar inserir pausas entre palavras cuja fronteira prosódica é fraca, como acontece com os determinantes e seus nomes correspondentes.

Embora pareça simples, a ressilabificação é um fenômeno que precisa ser trabalhado em sala de aula, pois também influencia o ritmo da língua. Quando o aprendiz não reconhece os processos de ressilabificação da língua aprendida, ele terá dificuldade de compreensão e conseqüentemente terá a impressão de que a fala é muito rápida (como discutido na seção 3.1). Além disso, do ponto de vista da produção, se o aprendiz não for capaz de ressilabificar sistematicamente em espanhol, sua fala será marcada por um número excessivo de pausas em locais onde não deveriam estar presentes o que afetará diretamente o ritmo na língua estrangeira. Na próxima seção, apresentamos três propostas para praticar o ritmo nas aulas de espanhol/LE.

---

20 Na análise acústica, os formantes são as frequências de um som complexo reforçadas por um filtro acústico (o trato vocal). Para cada vogal, há infinitos formantes. No entanto, normalmente são analisados apenas os dois primeiros formantes, pois são eles responsáveis pelo timbre específico de cada vogal.

21 Marcamos as fronteiras entre as sílabas com um ponto.

## 4.1 Práticas – Ritmo

**Atividade 3 Escucha selectiva con los ojos cerrados – (PERCEPÇÃO).**

<b>Material</b>	Gravação em áudio da leitura de um conto, um pequeno diálogo de um filme, programa de rádio, entrevista, etc.
<b>Requisitos</b>	Estudantes de nível inicial, intermediário ou avançado
<b>Procedimento</b>	Pedir aos alunos que fechem os olhos e escutem a gravação tentando prestar atenção apenas no ritmo da fala.

O objetivo deste exercício é ajudar os alunos a desenvolverem uma escuta seletiva, para que possam tomar consciência do ritmo da língua mais ou menos como acontece com os músicos que desenvolvem a habilidade de perceber, em uma orquestra, o toque de apenas um instrumento (GIL FERNÁNDEZ, 2007).

Se este exercício resultar muito difícil para os estudantes, o professor pode acrescentar algumas etapas prévias antes de repeti-lo novamente em aula. Para isso, poderá, primeiramente, trazer uma gravação de uma música instrumental (de preferência que utilize instrumentos de percussão<sup>22</sup>). A principal função dos instrumentos de percussão é a de marcar o ritmo da música e, por essa razão, pode ajudá-los posteriormente na percepção do ritmo da fala. A tarefa aqui não se trata de exigir dos alunos que identifiquem o instrumento e o nomeiem, mas simplesmente de prestar atenção e tentar reproduzir o som do instrumento que, na música selecionada pelo professor, está marcando o ritmo. Para os estudantes brasileiros, esta tarefa normalmente não é muito difícil, pois há uma variedade de ritmos brasileiros que utilizam instrumentos de percussão com um ritmo bem marcado, tais como o samba, o maculelê, a capoeira, o coco, entre outros. Esse exercício prévio pode ajudar os alunos a se prepararem para a tarefa de percepção do ritmo da fala propriamente dita.

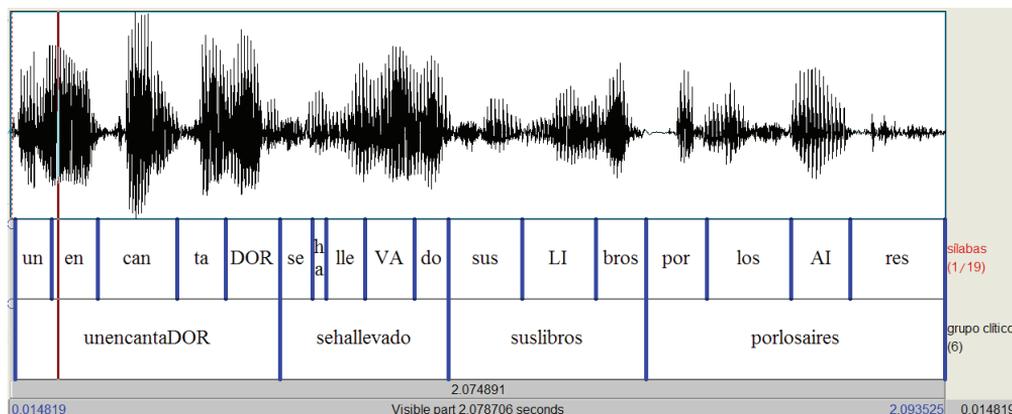
Ainda com objetivo de despertar a consciência dos alunos a respeito do ritmo da língua espanhola, Lahoz (2012) propõe um exercício de imitação que consiste em pedir para os alunos imitarem falantes nativos de espanhol falando PB como língua estrangeira. Essa tarefa normalmente é bastante divertida para os alunos, pois, atualmente há uma grande quantidade de jogadores argentinos, colombianos e de outros países latino-americanos em times de futebol brasileiros e, por essa razão, os alunos estão relativamente familiarizados com o sotaque desses jogadores em português.

22 Os tambores, tamborins, triângulos, berimbaus, agogôs, chocalhos são alguns instrumentos de percussão muito comuns.

Se se trata de grupos com alunos mais jovens, a tarefa pode ser bastante divertida, se o professor: (1) selecionar previamente pequenos trechos de gravações em vídeo de um grupo de famosos falantes nativos de espanhol, mas que estejam falando português, (2) anotar o nome desses famosos em papéis e sortear entre os alunos tais nomes, (3) explicar para os alunos que eles são os famosos indicados naqueles papéis e que foram convidados para dar uma pequena entrevista para um canal de televisão brasileiro. No final da performance de cada aluno, o grupo terá que adivinhar quem é este famoso. A tarefa final consistirá na comparação da produção de cada aluno com a gravação original do famoso correspondente trazida pelo professor. Nesse momento, o professor pode chamar a atenção dos alunos principalmente para o aspecto rítmico das produções dos nativos e evitar discutir as diferenças segmentais entre as duas línguas.

#### Atividade 4 – Karaoke – PERCEPÇÃO E PRODUÇÃO CONTROLADA

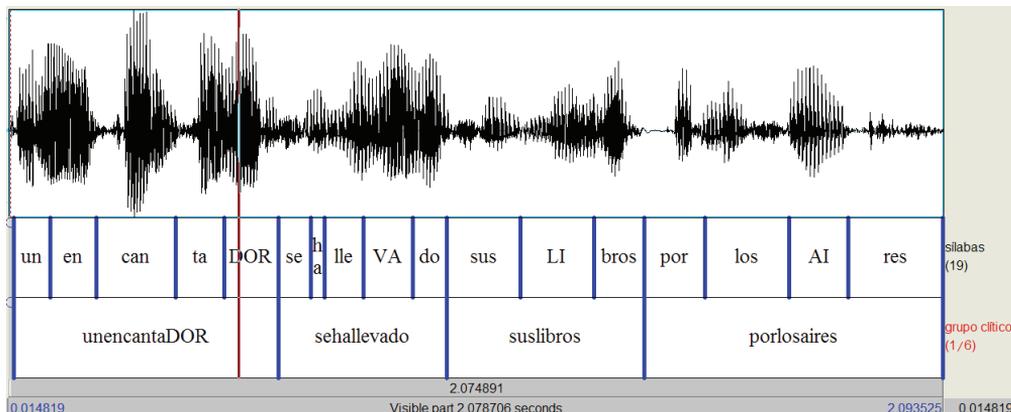
Outra proposta bastante interessante de Lahoz (2012) é a atividade que ele chama de karaokê. Nela, utilizando o programa Praat<sup>23</sup>, o professor pode segmentar e transcrever uma gravação qualquer, indicando as fronteiras das unidades que ele queira praticar com seus alunos, sejam elas sílabas, palavras, grupos clíticos, grupos entoacionais, etc. Na Figura 5, indicamos uma sugestão de segmentação em sílabas e grupos clíticos.<sup>24</sup>



**Figura 5** Forma de onda, segmentação em sílabas e grupo clítico do enunciado “Un encantador se ha llevado sus libros por los aires”. Pronunciada por uma informante espanhola de Madri. (continua)

23 Os professores que não estão familiarizados com programa, podem consultar o trabalho de Duarte (2014). O livro está disponível para download gratuito.

24 Seguindo Nespore e Vogel (1986), consideramos grupo clítico como a unidade prosódica que contém um ou mais clíticos e uma única palavra de conteúdo.



**Figura 5** Forma de onda, segmentação em sílabas e grupo clítico do enunciado “*Un encantador se ha llevado sus libros por los aires*”. Pronunciada por uma informante espanhola de Madri. (continuação)

Podemos observar na Figura 5 que, quando o som é reproduzido, o cursor vai avançando não só sobre a forma da onda, mas também sobre a segmentação realizada e, dessa maneira, o aluno tem tanto o apoio auditivo, quanto o visual. Além disso, pode observar também a duração dos segmentos ao longo do enunciado, a duração das sílabas tônicas em contraposição com as sílabas átonas e o ritmo como um todo. Decidimos não segmentar a gravação também em palavras para que o aluno possa, além disso, concentrar sua atenção exclusivamente no grupo clítico e poder perceber que, em um enunciado, as palavras não são produzidas isoladamente, mas de forma concatenada.

Depois que os alunos já tiverem escutado e observado visualmente as marcas do ritmo, devem desabilitar o som do computador e dar novamente play na gravação para apenas observar o movimento do cursor, mas sem o som. A tarefa, então, é a de tentar acompanhar o ritmo da fala produzida na gravação somente com o apoio visual do cursor. Finalmente, uma sugestão é que o professor peça para seus alunos se gravarem<sup>25</sup> durante o exercício, pois, dessa maneira poderão, utilizando o Praat, comparar a gravação original com sua própria gravação para verificar a duração total do enunciado, das sílabas tônicas e átonas e também dos grupos clíticos.

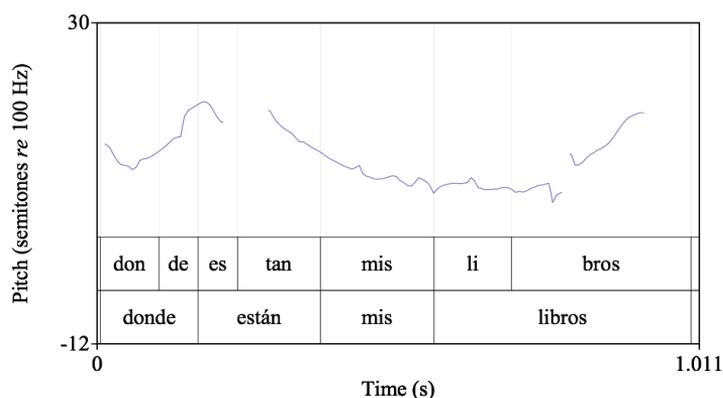
## 5 Entoação

Segundo Botinis et al. (2001), a entoação pode ser definida como a combinação de traços tonais que estão dentro de unidades estruturais maiores e que estão

25 Além do Praat, há diversos programas livres disponíveis atualmente para fazer gravações de áudio. Sugerimos, porém, o Audacity (<http://www.audacityteam.org/>) por ser gratuito e simples de usar.

associados com o parâmetro acústico da frequência fundamental ou  $F_0$  e suas variações no processo de fala.

A  $F_0$  é um parâmetro acústico correlato da frequência laríngea e definido pelo máximo divisor comum das frequências dos harmônicos que compõem a parte periódica do sinal de fala e é medido em Hertz (Hz). A produção da entoação, do ponto de vista articulatório, é dada pelo número de vezes por segundo que as pregas vocais completam um ciclo de vibração (abertura e fechamento). A percepção da entoação é definida pela sucessão de acentos tonais que são percebidos ao longo da enunciação como graves ou agudos. A Figura 6 ilustra uma curva de  $F_0$  (azul) de um enunciado em espanhol:



**Figura 6** Curva de  $F_0$  (azul) e segmentação em sílabas e palavras do enunciado *¿Dónde están mis libros?*, lido por uma informante espanhola, de Madri.

Na Figura 6, podemos observar que a curva de  $F_0$  não é contínua, ela é interrompida justamente onde há silêncio, como na oclusão para a produção de [t] e também durante a produção das fricativas desvozeadas [s]. Nosso ouvido, porém, não percebe essas interrupções na frequência fundamental no sinal de fala e, por isso, percebemos o *pitch* como se ele tivesse sido produzido de forma contínua. Além disso, observamos também que, em alguns momentos, a curva de  $F_0$  apresenta leves subidas ou descidas; essas diferenças, porém, também não são percebidas pelo ouvinte.<sup>26</sup>

26 Na literatura, tais variações são denominadas micromelódicas. A micromelodia pode ser causada por obstruções da passagem do fluxo de ar no trato vocal. É o que acontece, por exemplo, quando produzimos uma oclusiva sonora. Nesse caso, o fluxo de ar que passa pela glote diminui, o que reduz também a vibração das pregas. Acusticamente, o resultado é que observamos uma queda maior da  $F_0$  próxima às oclusivas. O falante não é capaz de controlar essa descida de  $F_0$  e, portanto, ela não é linguisticamente determinada.

A entoação é definida a partir do tipo de função que ela pode transmitir. As funções podem ser de natureza linguística, paralinguística e extralinguística. Do ponto de vista linguístico, as três principais funções da entoação estão relacionadas com as noções de marcação de proeminência, agrupamento e discurso, relacionados com os diversos níveis da gramática de uma língua (BOTINIS et al., 2001). A função paralinguística da entoação está relacionada com a transmissão das emoções, tais como a alegria, a tristeza, a raiva, o tédio etc. Por isso, sentimentos como a alegria ou a raiva costumam apresentar maior variação e valores mais altos de  $F_0$ , enquanto sentimentos como a tristeza ou o tédio estão associados com valores mais baixos de  $F_0$  e menor variação tonal. Finalmente, a função extralinguística da entoação é determinada pela fisiologia do falante, ou seja, depende diretamente do tamanho e da forma das suas pregas vocais. Assim, as crianças, por terem as pregas vocais menores, apresentam valores mais altos de  $F_0$  que as mulheres e os homens, respectivamente. Além de estar relacionada com a idade e o sexo, também a marcação de atitudes culturais, status social, etc. são consideradas funções extralinguísticas da entoação (BOTINIS et al., 2001). Esta última função da entoação, porém, não será discutida neste capítulo.

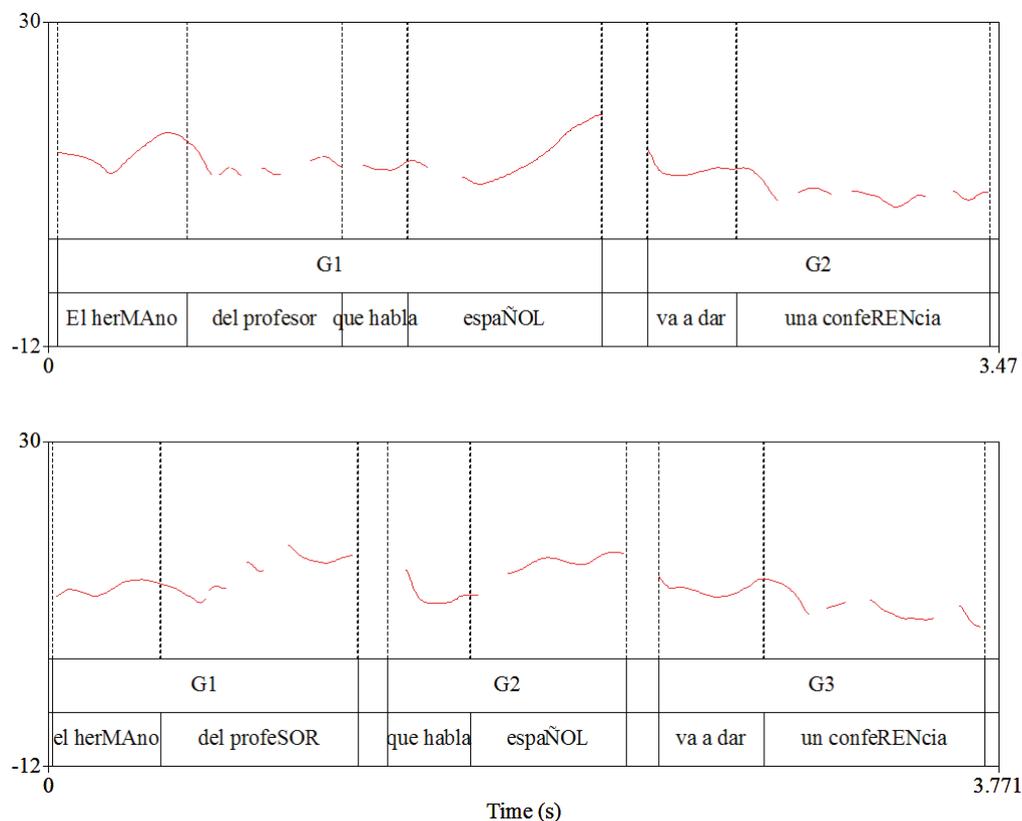
Para Lahoz (2012), uma boa forma de começar a familiarizar os alunos a respeito da entoação do espanhol/LE é iniciar o trabalho tratando da função paralinguística que o autor denomina de significado afetivo da entoação. Embora reconheça que há variação cultural na transmissão das emoções, elas apresentam um caráter icônico bastante comum<sup>27</sup> e, por isso, podem ser exploradas, principalmente, na fase inicial de aprendizagem da entoação.

Além da importância do ensino da função paralinguística da entoação, é fundamental também trabalhar a função linguística, já que está diretamente relacionada com a gramática de cada língua. Assim, a demarcação de fronteiras de constituintes, proeminência e modalidade de enunciado são alguns dos aspectos que podem ser incluídos nas aulas de espanhol/LE.

Com relação à função de marcação de fronteira dos constituintes, a entoação desempenha um papel fundamental não apenas em espanhol, mas em todas as línguas e, embora esteja diretamente relacionada com a sintaxe, não mantém com ela uma relação de isomorfia (cf. NESPOR; VOGEL, 1986). Lahoz (2012) observa que, em espanhol, quando um grupo entoacional termina com contorno ascendente, significa que existe alguma independência sintática com relação ao grupo seguinte, como podemos observar na Figura 7:

---

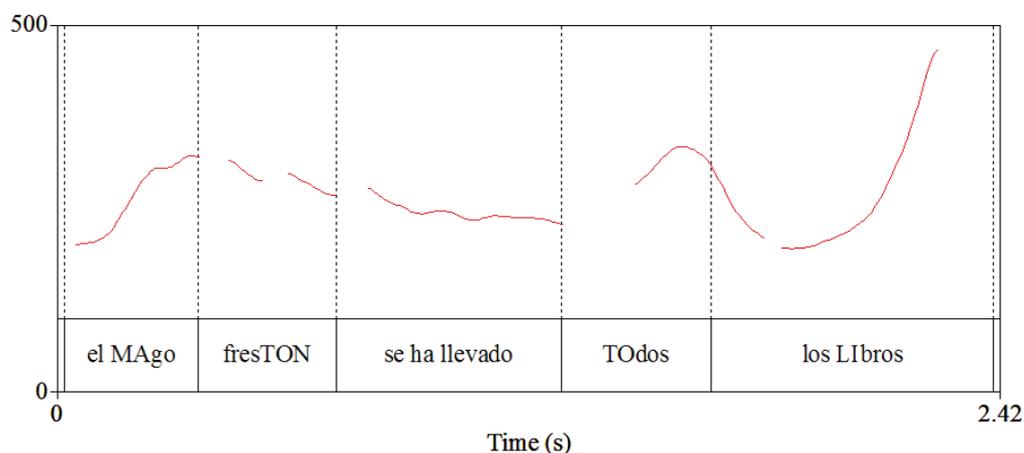
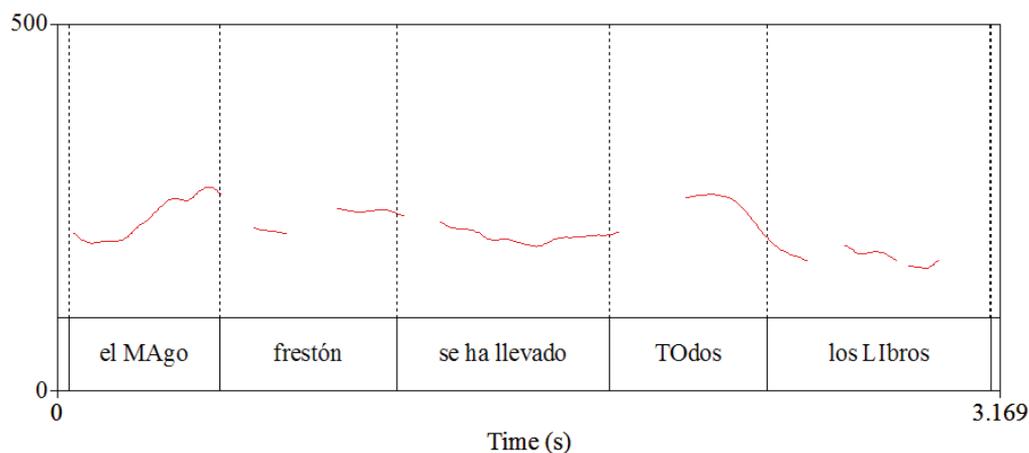
27 De certa forma, sua visão dialoga com o trabalho de Ekman (2011) a respeito da universalidade da transmissão das emoções.



**Figura 7** Curva de  $F_0$  suavizada em 10Hz (vermelha) e segmentação em palavras e grupos entoacionais dos enunciados: *El hermano del profesor que habla español va a dar una conferencia* e do enunciado *El hermano del profesor, que habla español, va a dar una conferencia*, pronunciadas por uma formante argentina, de Junín.

No primeiro enunciado, como o sujeito é muito longo, *el hermano del profesor que habla español*, é comum na fala que haja um contorno de  $F_0$  ascendente, justamente para indicar que o enunciado ainda não acabou e que está relacionado com toda a predicação que vem a seguir. Além do contorno ascendente, pode haver também pausa indicando esta fronteira. Já no segundo caso, o contorno ascendente indica que se trata de uma construção parentética. Além da função de marcação de fronteira prosódica desempenhada pela entoação, ainda na Figura 7, é possível observar outra função bastante comum, que também é transmitida pela entoação, a proeminência prosódica. No primeiro enunciado (Figura 7), apenas pela observação do contorno de  $F_0$ , é possível observar que as palavras *hermano* e *español* são claramente proeminentes (palavras que se destacam se comparadas com as demais palavras) enquanto que no segundo enunciado (abaixo) a proeminência é bem clara nas palavras *profesor* e *español*. Embora esses dois aspectos

sejam muito importantes para o entendimento da entoação das línguas, boa parte dos estudos sobre a entoação se concentram unicamente na função de marcação de modalidade<sup>28</sup> de enunciado. Na Figura 8, o que diferencia os dois enunciados é a modalidade. No enunciado de cima, trata-se de uma declarativa e no de baixo, de uma interrogativa total em espanhol/LE.



**Figura 8** Curva de  $F_0$  suavizada em 10Hz (vermelha) e segmentação em grupos clíticos dos enunciados: *El mago Freston se ha llevado todos los libros e ¿El mago Freston se ha llevado todos los libros?*, lidas por uma informante brasileira.<sup>29</sup>

28 Modalidade define o estatuto do enunciado, ou seja, se o enunciado é interrogativo, declarativo, exclamativo ou imperativo (DUBOIS et al., 1998)

29 No momento da gravação, a informante residia em Madri há dois anos e nunca havia estudado espanhol no Brasil antes de viajar para a Espanha.

É possível perceber, pelos múltiplos significados que são transmitidos pela entoação, que este ramo da prosódia ainda não foi suficientemente explorado. Os modelos linguísticos atuais se concentram, sobretudo, na análise dos significados linguísticos da entoação e a partir, principalmente, da análise de enunciados isolados. Por isso, ainda sabemos pouco sobre a entoação em porções maiores que o enunciado isolado e sobre os contextos discursivos mais amplos. Por essa razão, o professor de espanhol poderia se perguntar se, ainda assim, seria possível ensinar entoação nas aulas de espanhol/LE. A resposta é sim, mas antes de iniciar o trabalho com a entoação, o professor precisa ter consciência de que os contextos comunicativos mais comuns que ocorrem normalmente em sala de aula estão muito distantes dos contextos de comunicação reais, por isso sua tarefa primordial será tentar “recriar”, sempre que possível, esse ambiente para que os alunos se sensibilizem a respeito da importância da entoação para a comunicação em língua estrangeira. Dessa forma, os exercícios propostos na próxima seção estão direcionados para a sensibilização tanto do professor quanto dos seus alunos a respeito da importância da entoação em língua estrangeira.

## 5.1 Práticas – Entoação

### **Atividade 6 Aprendizices de Actores – (PERCEPÇÃO E PRODUÇÃO)**

<b>Material</b>	Cena de um filme escolhida pelos alunos
<b>Requisitos</b>	Estudantes de nível inicial, intermediário ou avançado
<b>Procedimento</b>	Cada aluno assume um papel (1ª fase) Apresenta-se a cena inteira (2ª fase) Apresenta-se a cena, fazendo pausas nas fronteiras de enunciados ou de sintagmas para que os alunos tenham tempo para memorizar as falas das personagens e repeti-las (3ª fase)

O objetivo deste exercício proposto por Cortés (2002) é o de ajudar os alunos a se concentrarem e prestarem atenção à entoação em um contexto comunicativo concreto. O exercício pode apresentar uma segunda etapa<sup>30</sup> que consiste em dublar a cena. Para isso, o professor entrega aos alunos a transcrição do diálogo e pede que, em duplas, ensaiem o diálogo até conseguir decorar suas falas. Em seguida, o professor apresenta novamente a cena do filme, porém sem som. A tarefa dos alunos será dublar a fala das personagens e, para isso, terão que adequar sua fala ao ritmo e à entoação da gravação original. Este é um exercício bastante lúdico e motivador, que poderá ser aplicado a estudantes de diferentes níveis de

30 Baseada na atividade intitulada pelo autor “Cinema Mudo”.

conhecimento de espanhol. É preciso, apenas, que o professor tenha cuidado ao selecionar as cenas dos filmes, pois o conteúdo segmental, taxa de elocução e duração precisam ser adequados ao nível de conhecimento de cada grupo de alunos.

### **Atividade 7 Significados de la entonación – (PERCEPÇÃO E PRODUÇÃO)**

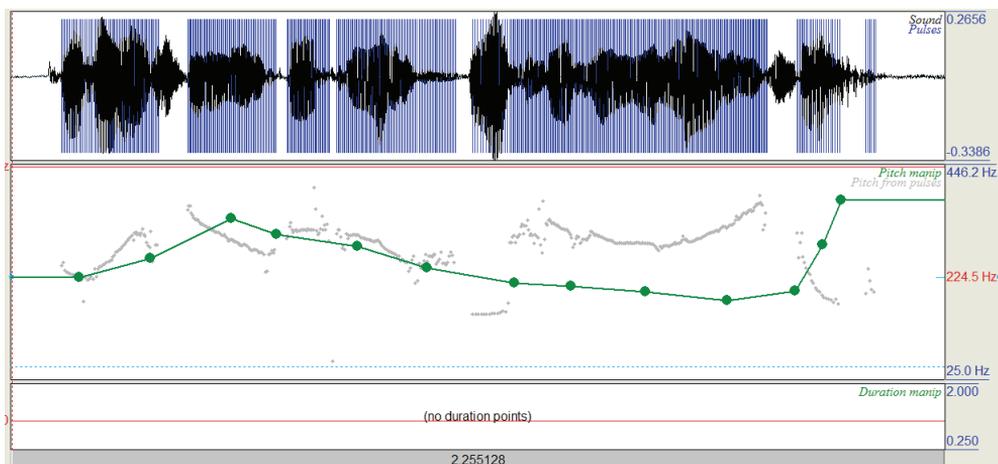
<b>Material</b>	Gravação de um falante nativo de espanhol ou a fala do próprio professor
<b>Requisitos</b>	Estudantes de nível inicial, intermediário ou avançado Gravação de uma mesma palavra ou expressão que transmitem diferentes significados comunicativos (linguísticos e/ou paralinguísticos)
<b>Procedimento</b>	Os estudantes devem indicar quais os significados que eles acreditam que estão sendo transmitidos por cada uma das emissões e se eles acham que coincidem, ou não, com os padrões entoacionais do PB.
<b>Objetivos</b>	Perceber a importância da entoação para a comunicação em L2 (assim como na L1)

Após o exercício de reconhecimento da entoação em espanhol, o professor pode utilizar o Praat como um apoio visual e auditivo das curvas entoacionais. Além disso, pode incluir uma atividade de produção na qual solicita aos alunos que escutem as frases e as repitam tentando imitar a entoação usada em cada uma delas. Assim, poderão comparar as curvas entoacionais produzidas com as do falante nativo ou professor.

Em muitos casos, os alunos têm dificuldade em repetir a frase seguindo o padrão entoacional do espanhol. Para solucionar este problema, Lahoz (2012) propõe a manipulação da própria voz do aluno. Para isso, na tela de objetos do Praat, selecionamos o arquivo de som que queremos manipular, escolhemos a função *Manipulate* e, em seguida, *To Manipulation*, podemos manter os valores *default* de mínimo e máximo de  $F_0$  e clicamos ok. Na tela de objetos do Praat, é criado um novo arquivo de tipo *Manipulation*. Clicamos em *View & Edit* e uma nova tela é apresentada com um contorno de  $F_0$  em verde. Escolhemos as opções *Pitch > Stylize Pitch (2 st)*. Assim, teremos uma curva de  $F_0$  estilizada, ou seja, bem mais simplificada, com poucos pontos. Ao clicar nos pontos da curva, é possível apagar ou criar novos pontos com as funções *Remove Pitch Point* e *Add Pitch Point*, respectivamente. É possível também arrastar os pontos, manipulando, assim, o contorno original para que ele se assemelhe cada vez mais ao formato da curva produzida pelo falante nativo.

Quando o aluno escuta sua própria voz com o novo contorno entoacional, geralmente é mais fácil que ele perceba a diferença entre o contorno da língua estrangeira e da sua língua materna e consequentemente consiga repetir a entoação, dessa vez seguindo o modelo do espanhol. Na Figura 9, temos um exemplo de manipulação da curva de uma interrogativa total em espanhol/LE, cuja curva de  $F_0$  é bastante diferente

da curva em espanhol (variedade centro norte peninsular). A curva em cinza corresponde ao contorno original e a curva em verde corresponde ao novo contorno com uma primeira manipulação. À medida que são feitas as mudanças no contorno, é possível ir ouvindo e verificando se as mudanças são perceptivamente aceitáveis ou não.



**Figura 9** Curva de  $F_0$  original (cinza) e manipulada (verde) do enunciado: *¿Queréis sentir la fuerza de mi lanza?*, lido pela informante brasileira Br-1.

Este exercício é muito eficaz pois, a partir da gravação do próprio aluno, é possível fazer as correções e ajustar sua pronúncia. A maior dificuldade, porém, está relacionada com o tempo necessário para fazer tal atividade. Como é necessário ouvir e testar a localização de cada ponto na curva a fim de conseguir um estímulo natural e que reflita a entoação do espanhol, gasta-se muito tempo. Além disso, para cada nova gravação feita pelo aluno, são necessários os mesmos procedimentos de manipulação. Até onde sabemos, ainda não existe nenhum programa gratuito que realize essa manipulação de forma automática ou semiautomática, mas esperamos que, em breve, tenhamos alguma ferramenta desse tipo, pois a manipulação do sinal de fala do próprio aluno é uma técnica que pode auxiliar o professor no trabalho de correção da produção oral do estudante.

## 6 Conclusão

O objetivo deste capítulo foi evidenciar a necessidade de se incluir nas aulas de espanhol/LE atividades que ensinem explicitamente a prosódia do espanhol. Para isso, apresentamos uma breve discussão teórica sobre os fenômenos prosódicos e propusemos atividades práticas para o ensino de prosódia que podem ser incorporadas às aulas de espanhol/LE.

Trata-se apenas de uma breve iniciação ao tema e, por essa razão, para que o professor de espanhol/LE possa realmente compor um conjunto amplo de atividades para o ensino de prosódia em suas aulas, recomendamos fortemente a leitura dos trabalhos de Porch Olivé (1999), Gil Fernández (2007), Lahoz (2012), Rico (2012) e Cortés (2002). Nesses livros, o professor encontrará uma ampla gama de atividades que poderá aplicar diretamente em suas aulas ou adaptar segundo as necessidades dos seus estudantes.

Mais do que apenas levantar uma discussão teórica sobre a prosódia, esperamos ter conseguido despertar no professor, sobretudo no professor brasileiro, a consciência da importância de se ensinar, a além da produção segmental, também a prosódia da língua espanhola. Como estamos diante de uma área muito ampla, ao professor de espanhol/LE que está iniciando seu estudo sobre a prosódia, sugerimos que não desanime e que busque, sempre que possível, entender as características da prosódia tanto do PB como do espanhol, pois a análise contrastiva poderá ajudá-lo a identificar e, muitas vezes, a prever as dificuldades de seus estudantes o que facilitará seu trabalho na preparação de novas atividades com objetivo de corrigir tais dificuldades.

Finalmente, muito longe de esgotar o tema, esperamos ter contribuído para conscientizar o professor do papel primordial da prosódia para comunicação, tanto na nossa própria língua, como também na língua estrangeira que estamos ensinando/aprendendo.

## Referências

- ABERCOMBIE, D. *Elements of general phonetics*. Aldine Pub. Company, 1967.
- BUSSMANN, H. *Routledge Dictionary of Language and Linguistics*. New York: Routledge, 1996.
- BARBOSA, P. A. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration. In: *First ESCA TRW on Speech Production Modeling: From Control Strategies to Acoustics; 4th Speech Production Seminar: Models and Data, 1996, Austrans: France. Resumos...* p. 85-88.
- BARBOSA, P. A. Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: E. M. SCARPA (Org.). *Estudos de prosódia*. Campinas: Editora da Unicamp, 1999. p. 21-52.
- BARBOSA, P. A.; ERIKSSON A.; AKESSON, J. Cross-linguistic similarities and differences of lexical stress realization in Swedish and Brazilian Portuguese. In

- E.L. Asu e P. Lippus (Eds) Nordic Prosody CONFERENCE, 11, 2012, Frankfurt am Main: Peter Lang, Tart. Resumos...2013, p. 97-106.
- BARBOSA, P. A.; MADUREIRA, S. *Manual de fonética acústica experimental: aplicações e dados do português*. São Paulo: Cortez, 2015, 591p.
- BOTINIS, A.; B. GRANSTRÖN; B. MÖBIUS. Developments and paradigms in intonation research. *Speech Communication*, v. 33, p. 263-296, 2001.
- CORTÉS, M. *Didáctica de la prosodia del español: la acentuación y la entonación*. Madrid: Edinumen, 2002, 158p.
- COUPER-KUHLEN, E. *An introduction to English prosody*. Edward Arnold, 1986.
- DUARTE, J. A. C. *Manual de análisis acústico del habla con Praat*. Series Minor (49). Instituto Cara y Cuervo. Imprenta Patriótica, Bogota, 2014. Disponível em: <<http://www.bibliodigitalcaroycuervo.gov.co/998/>>. Acesso em: 03 fev. 2017.
- DUBOIS, J. et. al. *Dicionário de Lingüística*. 10 ed. Cultrix: São Paulo, 1998.
- EKMAN, P. *A linguagem das emoções*. São Paulo: Lua de Papel, 2011.
- FIRTH, J. R. Sounds and prosodies. In: W. E. JONES e J. LAVER (Eds). *Phonetics in Linguistics: A Book of Readings*. Londres: Longman, 1948. p. 47-65.
- GIL FERNÁNDEZ, J. *Fonética para profesores de español: de la teoría a la práctica*. Madrid: Arco/Libros, 2007. 614p.
- GOLDMAN-EISLER, F. The significance of changes in the rate of articulation. *Language and Speech*, 4, 171, 1961.
- GUT, U.; TROUVAIN J.; BARRY, W. J. Bridging research on phonetic descriptions with knowledge from teaching practice: the case of prosody in non-native speech. In: J. TROUVAIN e U. GUT (Eds). *Non-Native Prosody: Phonetic Description and Teaching Practice*. Berlim: Walter de Gruyter, 2007. p. 3-21.
- LAHOZ, J. M. La enseñanza de la entonación, el ritmo y el tempo. In: GIL, J. (Ed). *Aproximación a la enseñanza de la pronunciación en el aula de español*. Madrid: Edinumen, 2012. p. 93-131.
- LAVÉ, J. *Principles of Phonetics*. Cambridge: Cambridge University Press, 1994.

- Lennon, P. Investigating fluency in EFL: a quantitative approach. *Language Learning* 40, 387-417, 1990.
- LLISTERI, J. et al. The perception of lexical stress in Spanish. In M. J. Solé, D. Recasens e J. Romero (Eds.). In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 15, vol. 39, 2003, Barcelona: Causal Productions. Resumos... p. 2023-2026. Disponível em: [http://liceu.uab.cat/~joaquim/publicacions/Llisteri\\_Machuca\\_Mota\\_Riera\\_Rios\\_03\\_Perception\\_Stress\\_Spanish.pdf](http://liceu.uab.cat/~joaquim/publicacions/Llisteri_Machuca_Mota_Riera_Rios_03_Perception_Stress_Spanish.pdf). Acesso em: 02 fev. 2017.
- MASSINI, J. *A Duração no estudo do acento e do ritmo em português*. 1991. Dissertação (Mestrado em Linguística) – Instituto de Estudo da Linguagem. Universidade Estadual de Campinas, Campinas, 1991.
- MORAES, J. A. Corrélats acoustiques de l'accent de mot en portugais brésilien. In: international congress of phonetic sciences, 11, vol. 3, 1987, Tallinn: Academy of Sciences of the Estonian S.S.R. Resumos... p. 313-316.
- NESPOR, M. e VOGEL, I. *Prosodic phonology*. Foris: Dordrecht, 1986.
- OSSER, H.; PENG, F. A cross-cultural study of speech rate. *Language & Speech*, v. 7, p. 120–125, 1964.
- POCH OLIVÉ, D. La pronunciación en la enseñanza del Español como Lengua Extranjera. *RedELE: Revista Electrónica de Didáctica ELE*, v. 1, 10, 2004.
- POCH OLIVÉ, D. *Fonética para aprender español: Pronunciación*. Madrid: Edinumen, 1999, 110p.
- RICO, J. R. El acento y la sílaba en la clase de ELE. In: GIL, J. (Ed). *Aproximación a la enseñanza de la pronunciación en el aula de español*. Madrid: Edinumen, 2012. p. 75-92.
- SILVA, C. C. Análisis melódico de declarativas e interrogativas absolutas en español/LE. *Phonica*, vol. 5, 2009.
- SILVA, C. C. *Análise fonético-experimental da entoação de declarativas e interrogativas em espanhol/LE*. 2016. 225p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, Campinas, 2016.

## Escrita, prosódia e leitura

Vera Pacheco  
Universidade Estadual do Sudoeste da Bahia

### 1 Apresentação

O ser humano, em algumas culturas, além da comunicação oral, ainda conta com o sistema de escrita, para fins comunicativos. Assim, a língua pode ser expressa tanto na sua forma oral, quanto na sua forma escrita; dois sistemas de signos que, de acordo com Saussure (1970), são diferentes entre si, mas que acabam se misturando intimamente.

Diferentemente da língua falada, a língua escrita não é uma capacidade inata, um instinto, mas uma invenção humana, um produto sócio-cultural, cuja origem remonta a um passado relativamente recente, se se considerar as dezenas de milhares de anos pelos quais se estende o progresso intelectual da humanidade (DIRINGER, 1969).

Da pré-história, com o homem das cavernas e suas pinturas rupestres, até os dias atuais, a humanidade conheceu um número razoável de sistemas de escrita. Pesquisas paleontológicas indicam que a escrita cuneiforme, desenvolvida pelos sumérios, por volta dos anos 4.000 a 3.1000 a.C., seja o primeiro sistema de escrita.

Independentemente do tipo de sistema de escrita, para Cagliari (2006, p. 1), todos os seus “criadores tiveram que buscar na observação da fala as diretrizes para a formação dos sistemas de escrita. Isso é particularmente claro no caso da criação do alfabeto”. Para esse autor, a invenção de um sistema de escrita, qualquer que seja ele, implica a existência de uma língua oral.

Nesse sentido, partindo da hipótese de que um sistema de escrita implica a existência de uma língua oral, podemos supor que a escrita possui recursos para representar aspectos da oralidade. Há, pois, uma estreita relação entre escrita e língua oral, relação essa que constitui o foco do presente capítulo.

Buscamos nas páginas que seguem refletir sobre a relação língua falada e língua escrita e de que maneira a escrita registra as variações entoacionais/melódicas, atitudes de falantes e emoção. Discutimos a importância dos marcadores prosódicos da escrita na construção do texto escrito e na leitura em voz alta.

## 2 Língua falada e língua escrita: pontos de desencontros e encontros

A fala e a escrita constituem dois sistemas de comunicação humana. Ao traçarmos um paralelo entre ambos sistemas, somos capazes de rapidamente estabelecermos diferenças entre si. Enquanto a fala é natural, inata e instintiva; a escrita não é inata, é, ao contrário, uma invenção humana, um produto sócio-cultural.

A fala, por ser instintiva, é desenvolvida sem intervenção formal e sistematizada. O desenvolvimento da escrita, por sua vez, requer, na grande maioria das vezes, a intervenção formal e sistematizada, ou seja, ela deve ser ensinada.

Ao desenvolver a fala, por exemplo, a criança não precisa ser ensinada a abrir a boca e deixar o ar sair livremente para a produção das vogais, bem como não precisa ser ensinada a fechar por completo a boca para fazer uma oclusiva. Isso ela já faz, instintivamente, na fase do balbúcio. Com o tempo, ela irá selecionar quais de suas produções sonoras são importantes na sua comunicação com o adulto e quais não são.

Mas, para representar graficamente os sons que a criança produz, precisamos ensiná-la a desenhar as letras que representam os sons de sua língua materna. Sem ser ensinada, dificilmente ela vai desenvolver a escrita. Sem que lhe seja ensinada a ortografia oficial de sua língua materna, a criança poderá ter uma escrita com muitas formas divergentes.

Não temos, dessa forma, como negar que fala e escrita são dispositivos de comunicação humana de natureza bem diversa. E, obviamente, não podemos negar que a escrita não é efetivamente o espelho da fala. Não há uma relação biunívoca entre os sons da fala e a sua representação gráfica por meio das letras.

O sistema ortográfico do Português está recheado de exemplos dessa não biunivocidade entre som e letra. Em alguns casos, temos uma única letra representando vários fonemas; em outros, temos um único fonema sendo representado por várias letras.

É só lembrarmos do nosso desespero em dizer se uma palavra é escrita com x, s ou com z. Só a letra x no nosso sistema ortográfico pode representar /z/, *exato*; /s/, *extrato*; /ʃ/ *lixo*; ou ainda uma sílaba inteira /kis/ como em *látex*. Em associação com a letra c pode também representar o /s/ como em *exceto*.

Há igualmente o outro lado da moeda: um único fonema sendo representado por diversas letras, como é o caso do /s/. Esse fonema pode ser representado pela letra s, *sapo*; pela letra x, *extrato*; pela letra c, *cela*; pelo ç, *caça*; e pelas combinações de ss, *assado*, ou xc, *exceto*.

Esses são alguns dos vários exemplos do pseudo caos que é o nosso sistema de ortografia. Sim, **pseudo** porque um sistema ortográfico tem motivações diversas e não só fonológica. Assim, uma determinada forma escrita pode resgatar questões de natureza etimológica ou filosófica. Um sistema de escrita pode ser essencialmente de “natureza fonológica”, como é o sistema da língua portuguesa, de acordo com Mateus (2006), mas não exclusivamente.

De forma resumida, podemos afirmar que a pouca relação entre letra e fonema que existe nos sistemas ortográficos não é uma evidência de falta de regularidade na ortografia de uma língua. Defendemos aqui que a ortografia de uma língua é satisfatoriamente coerente, a ponto de registrar aspectos prosódicos, tais como variação entoacional/melódica, atitudes de falantes e emoções. Há, nesse sentido, muitos pontos de encontros entre fala/língua e escrita que vão para além dos aspectos segmentais. Esses pontos de encontro são decorrentes dos diversos recursos gráficos que a escrita dispõe para representar aspectos que são recorrentes na língua falada, especificamente os aspectos prosódicos.

Os recursos que a escrita dispõe para representar os aspectos prosódicos da fala são os Marcadores Prosódicos-MP (CAGLIARI, 1989), aos quais dedicaremos o tópico seguinte.

### 3 Prosódia e escrita: os marcadores prosódicos

A organização sonora de uma língua conta com segmentos: vogais, consoantes, glides; que se organizam em unidades maiores, sílaba, palavra, frase, enunciado. A partir da sílaba, são organizados os constituintes prosódicos, que, grosso modo, estão relacionados ao acento, às variações entoacionais e melódicas.

No que se refere à constituição sonora, a prosódia está relacionada a) à intensidade (amplitude); b) à melodia (frequência fundamental –  $F_0$ ); c) à duração; d) às marcas de saliência (ársis/tésis); f) ao andamento; g) à velocidade de fala, h) ao registro e i) à qualidade de voz (CAGLIARI, 1981).

Do ponto de vista do significado, os eventos prosódicos da língua podem estar relacionados a) aos fatos sintáticos, à semelhança dos diferentes tipos de frases: exclamativa, interrogativa etc; b) à definição sócio-pragmática: nos diversos usos da linguagem; e c) às características discursivas, para-linguísticas e extra-linguísticas: na determinação de atitudes de falantes e expressão de emoção (CAGLIARI, 1981).

No ato comunicativo, as variações prosódicas são, portanto, determinantes para a continuidade e perfeita fluidez da fala dos interlocutores. Um padrão pro-

sódico inadequado pode colocar em risco a comunicação entre dois falantes. Pode inclusive levar a maus entendimentos. Por exemplo, se você não marcar adequadamente uma frase interrogativa, sua fala pode soar com uma ordem grosseira. Nesse sentido, as marcações prosódicas não são acessórios em nossa fala e por isso devem, sim, ser resgatadas na escrita.

De acordo com Cagliari (1998), um sistema de escrita como o nosso possui recursos que mostram ao leitor mais sutilezas e nuances da fala do que comumente se costuma acreditar. Para esse autor, esses recursos constituem os marcadores prosódicos. Segundo ele, um texto escrito, em especial a narração, possui marcas gráficas que têm como função principal indicar para o seu leitor como deverão ser as variações melódicas e entoacionais da passagem que estão sob escopo dessas marcas gráficas, que podem ser de natureza diversa e incluem desde formatação do texto à escolha lexical e uso de pontuação.

Esses MPs, de acordo com Pacheco (2006) com base em Cagliari (1989), podem ser de dois tipos: os marcadores prosódicos gráficos (MPG) e os marcadores prosódicos lexicais (MPL), como descrito a seguir.

### 3.1 Os marcadores prosódicos gráficos

A escrita alfabética pode possuir como parte constitutiva sinais gráficos de diferentes formas. Esses sinais gráficos são chamados de sinais de pontuação, com usos e funções bastante diversificados, como evidencia a fala de Catach (1998, p. 32):

La ponctuation constitue, auprès des signes alphabétiques, un système de renfort, d'ordre à la fois séparateur (des mots et groupés de mots) intonatif, syntaxique, discriminatoire et sémantique, destiné avant à une aide à la lisibilité. Ce système possède une sémiologie propre, et occupe depuis toujours (c'est là une découverte) dans l'écrit une place majeure. Il est destiné à éclairer le texte, à éviter les ambiguïtés de l'acte de lecture, dans tous les sens de ce mot, saisie matérielle et interprétation à la fois orale, visuelle et sémantique (...).

Essa multiplicidade de funções que os sinais de pontuação agregam hoje, conforme Catach (1998), não deixa de ser, em última análise, um reflexo das múltiplas funções que essas marcas gráficas assumiram ao longo da história da escrita, que, segundo Cagliari (1995), registra as formas primitivas do que hoje se considera como sinais de pontuação, já nos primeiros documentos escritos pelos sumérios. Ainda de acordo com esse último autor, “a primeira manifestação da escrita veio não só com os caracteres, mas também com marcas discursivas que hoje costumamos chamar de sinais de pontuação” (CALGIARI, 1995, p. 178).

Os primeiros textos escritos pelos sumérios, segundo Cagliari (1995), separavam os grupos sintáticos e semânticos com traços formando quadrados e retângulos. Já os egípcios, quando usavam a escrita hieroglífica, separavam os nomes próprios, colocando-os dentro de casulos (CAGLIARI, 1995).

Nesse sentido, como sugere Rocha (1997), os sinais de pontuação teriam sua origem juntamente com a origem da escrita, apresentando, ao longo da história, funções diversificadas, e não teriam tido uma aquisição tardia. Ainda para essa autora, a aquisição da escrita também teria sido lenta, ponto de vista que de fato faz sentido se se considerar que na história da escrita não há registro de uso de qualquer marca de pontuação nos documentos escritos durante a Antiguidade.

Para Cagliari (1995), há evidências, ao longo da história da escrita, que permitem crer que o sistema de pontuação que hoje nós usamos tenha se originado do sistema grego e definiu-se como tal no século XV, justamente com a invenção da imprensa. Apesar das regras impostas, no século XV, pelos gramáticos e pelos tratadistas de questões de ortografia, os sinais de pontuação nunca tiveram normas de usos rigorosamente estabelecidas.

Até mesmo nos dias atuais, em que esses sinais constituem parte integrante de qualquer texto escrito, no sistema de escrita, não há normas precisas quanto ao uso dessas marcas gráficas, que é bastante variável entre os autores de uma mesma língua e, naturalmente, entre as línguas, sendo quase uma questão de estilo. Não há, além disso, um acordo entre os autores quanto ao conjunto de marcas gráficas que devem ser consideradas como sinal de pontuação (CAGLIARI, 1995).

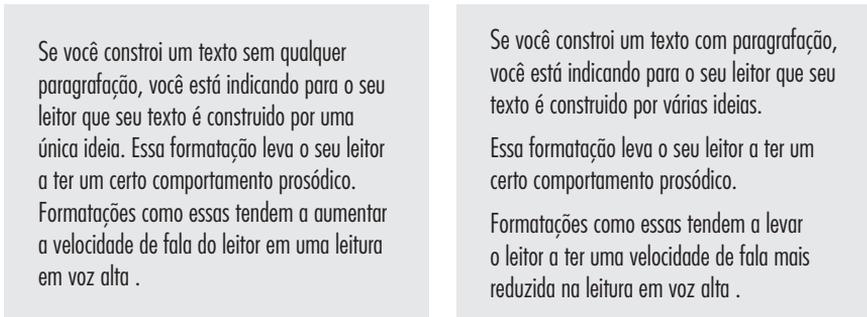
Nos dias atuais, é possível verificarmos nos textos narrativos e nos textos escritos nas redes sociais um uso, digamos “moderno”, dos sinais de pontuação. A quantidade e o uso combinado de alguns dos sinais de pontuação mais típicos da língua como exclamação, interrogação e reticências são usados com finalidade de expressão prosódica que vão para além de seus usos convencionais. Basta observarmos os diferentes sentidos assumidos pela frase “*Você não fez nada*” em função da pontuação que lhe segue:

- a) Você não fez nada.
- b) Você não fez nada....
- c) Você não fez nada!
- d) Você não fez nada!!
- e) Você não fez nada nada!!!!!!
- f) Você não fez nada!...
- h) Você não fez nada?
- i) Você não fez nada??
- j) Você não fez nada??????
- k) Você não fez nada?!....

Podemos verificar que o uso combinado e/ou em excesso dos sinais de pontuação incitam comportamento prosódico que indica uma certa irritabilidade, desencanto, raiva, etc. Esses recursos mostram que você não está diante de uma exclamação e uma pergunta típicas. Há uma informação prosódica adicional de extrema relevância que revela atitude do falante e seu estado emocional.

Além dos sinais de pontuação, outras marcas gráficas podem trazer informações prosódicas para o leitor, como a formatação do texto, tipo de letras, etc.

Por meio da formatação do texto, a presença ou ausência de parágrafos, por exemplo, pode levar o leitor a ter um ou outro comportamento prosódico. Vejamos nos exemplos da Figura 1:



**Figura 1** Diferença de formatação e comportamento prosódico.

O estilo de letra é um outro recurso gráfico que pode nos trazer pistas prosódicas. Quando, em um texto, queremos dar destaque a certas passagens, frequente usamos caixa alta, negrito, itálico. Ao ler esse texto, você alterará seu comportamento prosódico. Você pode aumentar ou abaixar seu tom de voz, você acelerar ou reduzir a sua velocidade de fala. Você pode AINDA, silabar uma palavra, o que le-va-rá o seu leitor a ler essa palavra de forma mais devagar. O **bom** leitor não ficará alheio a essas *marcas gráficas*.

Finalizamos por aqui os exemplos das diversas possibilidades de marcação prosódica gráfica. Podemos afirmar que temos à nossa disposição diversos recursos gráficos que podem expressar nossas intenções prosódicas e sinalizar para o nosso leitor como o texto que lhe é apresentado deverá ser lido. Tudo só depende de nossa criatividade. Nesse sentido, podemos trazer evidências robustas para os teóricos que defendem que a pontuação é um tipo de transcrição entoacional e, portanto, seria um análogo visual da prosódia (KONDO; MAZUKA, 1996; COHEN; DOUAIRE; ELSABBAGH, 2001, entre outros).

Além das marcas gráficas, os MPG, que nos orientam como deverá ser nosso comportamento prosódico, temos ainda os marcadores prosódicos lexicais, MPL, matéria da nossa próxima seção.

### 3.2 Os marcadores prosódicos lexicais

Os falantes têm, segundo Fónagy (1993), a entonação como um componente privilegiado para a atribuição de sua carga semântica. Assim, de acordo com o autor, “a entonação desempenha papel fundamental no ato de comunicação linguística através da manifestação de atitudes de falantes” (FÓNAGY, 1993, p. 223). Por meio da entonação, portanto, é possível aferir atitudes do falante, tais como a polidez, autoritarismo, arrogância, etc, sendo a atitude, diferentemente da emoção, um comportamento do falante que é controlado e determinado por si próprio, com implicações morais e intelectuais, como a ironia, a reprovação, a justificativa.

Além das atitudes dos falantes, as situações comunicativas diárias podem apresentar variações entoacionais que têm por objetivo a própria comunicação, como a exclamação, a interrogação; ou ainda, pode apresentar situações comunicativas cujas variações entoacionais não são controladas, pelo contrário, são inconscientes e produtos da tensão psíquica, como a cólera, a tristeza, a alegria. Segundo Fónagy (1993), tem-se, no primeiro caso, modalidade e, no segundo, emoção.

Essas variações melódicas presentes nas situações comunicativas, decorrentes da atitude do falante, da modalização e da emoção são variações prosódicas também passíveis de serem registradas na escrita por meio de expressões que, segundo Cagliari (2002, p. 7):

Caracterizam atitudes do falante, emoções e modos de dizer que fazem uma referência à prosódia da língua, uma vez que tais atitudes, emoções e modos de dizer precisam ser realizados foneticamente de uma determinada maneira e não de outra.

Assim, para Cagliari (1989, 2002), esses aspectos das situações comunicativas são representados na escrita por meio de expressões semânticas até por pequenas “descrições de situações em que a fala aparece bem definida em vários aspectos fonéticos, inclusive prosódicos” (CAGLIARI, 2002, p. 10). Esses marcadores constituem entradas lexicais no léxico mental dos falantes e, sendo uma entrada lexical, estão sujeitas ao funcionamento de qualquer unidade lexical, como, por exemplo, pertencer a uma classe gramatical.

Dessa forma, essas referências semântico-lexicais de cunho prosódico podem pertencer a diferentes categorias gramaticais, podendo ser um adjetivo, advérbio, expressões adverbiais que são usadas para qualificar prosodicamente o significado de um verbo que se refere ao ato de falar (CAGLIARI, 2002).

O que torna essas entradas lexicais especiais ao ponto de serem classificadas como marcador prosódico é justamente a sua carga semântica, que traz informações de cunho prosódico, podendo se referir “às atitudes do falante, seus

sentimentos, ao ritmo, ao volume, ao andamento, enfim, a qualquer um dos parâmetros constitutivos da prosódia como um processo geral” (CAGLIARI, 2002, p. 11).

Trata-se, pois, de um marcador prosódico do tipo lexical, podendo ser tratado, então, como Marcadores Prosódicos Lexicais (MPL) que possuem tanto informações da ordem da escrita, já que são palavras constituídas ortograficamente, quanto informações da ordem da fala, precisamente prosódica, já que sua carga semântica traz necessariamente informações que remetem a variações prosódicas.

Para Cagliari (2002), o Português possui várias palavras e expressões que podem ser usadas na escrita para fazer referência a elementos prosódicos de várias naturezas, como, por exemplo, ao acento. Na língua escrita, referências ao acento podem ser encontradas em expressões como: “falou acentuando bem as palavras”, “disse enfaticamente” (p.4)

Referências à velocidade de fala são feitas como em “falou devagar”, “falou destacando as palavras”, “disse bem devagar”. Expressões como “falaram os dois ao mesmo tempo”, “interrompeu fulano e disse”, “continuou”, “replicou”, etc. (p. 5) fazem referência à concatenação que é o modo de fala sem pausas.

A tessitura, que, de acordo com Cagliari (2002), refere-se à escala melódica usada na fala e está compreendida entre o tom mais grave e o mais agudo, pode ser resgatada na escrita por meio das palavras como “murmurou” (tessitura baixa), “berrou” (tessitura alta).

Já o volume, que é a intensidade com que o enunciado é dito, pode ser resgatado na escrita por meio de “gritou” (volume forte, geralmente classificado pelo ouvinte como som alto) e “falou baixo” (volume fraco, geralmente classificado pelo ouvinte como som baixo).

Com essas marcas, dentre outras apresentadas por Cagliari (1989, 2002), a escrita faz referências a aspectos estritamente prosódicos que são típicos da fala oral. Assim, é possível fazer referências às falas de personagens, indicando, dessa forma, as atitudes dos falantes, bem como suas emoções e sentimentos. Esses modos de dizer indicam ao leitor que as falas dessas personagens devem ser realizadas de um modo e não de outro, o que implica variações nos parâmetros prosódicos, como entoação, ritmo, acento, etc. (CAGLIARI, 2002).

Da mesma forma que uma frase pode ter seu sentido alterado em função do MPG que lhe acompanha, como vimos no item anterior, uma frase pode ter seu sentido alterado a depender do marcador prosódico lexical que a acompanha. Vejamos como uma frase como “*ela disse que ia embora*” pode expressar estados de espíritos diferentes de quem a pronuncia a depender do marcador prosódico lexical que a acompanha:

- a) Ela disse calmamente que ia embora.
- b) Ela disse gritando que ia embora.

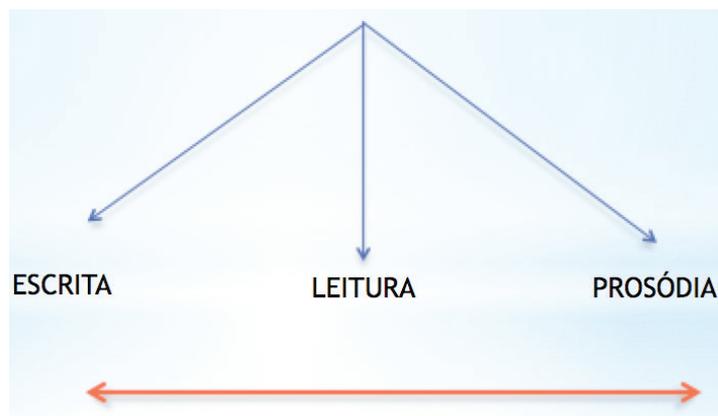
- c) Ela disse lentamente que ia embora.
- d) Ela disse desesperada que ia embora.
- f) Ela disse sussurrando que ia embora.
- g) Ela disse aos berros que ia embora.

É fácil verificar os efeitos de sentido que a alteração da palavra que segue o verbo *disse* acarreta à frase aqui analisada. Cada uma das palavras apresentadas remete a um comportamento prosódico específico, que implica em aumentar ou abaixar o tom e o volume da voz. Acelerar ou reduzir a velocidade de fala. Em cada uma dessas situações, estão expressas atitudes e emoções diversas do falante. O leitor não pode estar alheio a essas informações, sob risco de não compreender o que de fato o texto propõe.

## 4 O lugar dos marcadores prosódicos na leitura

Vimos até aqui que a escrita possui recursos para representar, no texto escrito, variações prosódicas. Há dessa forma, uma intenção por parte de quem escreve de registrar, de informar ao seu leitor o comportamento prosódico desejado, esperado e ideal.

Espera-se, por sua vez, que o leitor, ao ter contato com essas marcas, que são pistas prosódicas, seja capaz de recuperá-las. Dessa forma, podemos afirmar que a leitura dá voz a marcas prosódicas da escrita. A leitura, é, nesse sentido, o elo que existe entre a escrita e a prosódia, como pode ser visualizado na Figura 2.

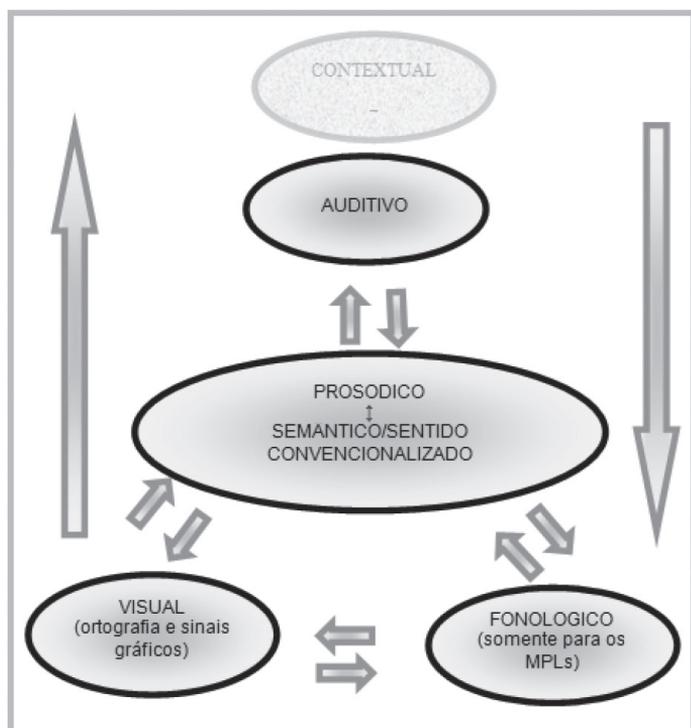


**Figura 2** Esquema representativo da relação escrita, leitura e prosódia.

A conversão entre informação prosódica visual do texto escrito em variação prosódica durante a leitura em voz alta tem implicações diretas para o processamento da leitura. De que forma fazemos essa conversão? Os modelos de processamento de leitura atuais dão conta somente da conversão entre ortografia e

fonema. Contudo, os MPG e os MPL possuem uma realidade que vão para além da ortografia. Eles possuem uma informação prosódica que deve ser acionada quando o estímulo visual - quer o sinal de pontuação, a formatação do texto ou a forma ortográfica do MPL - é apresentado ao leitor.

Para Pacheco e Oliveira (2014), somente um modelo interativo consegue dar conta do processamento dos marcadores prosódicos durante a leitura. Mas para essas autoras, é necessário também que se reveja os módulos envolvidos no processamento de leitura. Para elas o modelo deve ser como o que segue na Figura 3:



**Figura 3** Arquitetura proposta para o processo de reconhecimento dos marcadores prosódicos da escrita (PACHECO; OLIVEIRA, 2014, p. 209).

Assim para as autoras, é necessária a projeção de um modelo que seja capaz de prevê a conversão entre o semântico (para os MPL)/sentido convencionalizado (no caso dos MPG) em informação prosódica e vice-versa; bem como preveja um módulo AUDITIVO que permita que, em situações de escuta de leitura, sejamos capazes de, por exemplo, ouvir uma pergunta e representá-la como o sinal de interrogação.

É possível que um módulo CONTEXTO, conforme defendem as autoras, possa estar atuando no processamento dos marcadores prosódicos. Contudo, ain-

da não existem evidências experimentais controladas que permitam afirmar sobre a atuação desse módulo.

A proposta de um modelo que contemple o processamento dos marcadores prosódicos durante a leitura e a escuta de uma leitura torna mais evidente a estreita relação entre escrita e fala. É inegável as particularidades de cada um desses sistemas de comunicação humana, mas é igualmente inegável a interligação que ambos possuem.

## 5 Considerações finais

Nosso objetivo neste capítulo foi refletir sobre a relação língua falada e língua escrita. Mostramos que estamos diante de dois sistemas de comunicação humana que possuem natureza diversa, mas que guardam estreita relação entre si. Muitos aspectos da língua/fala são registrados eficazmente na escrita, em particular os aspectos prosódicos. Nesse sentido, a escrita dispõe dos marcadores prosódicos gráficos (sinais de pontuação, formatação de texto, estilos de letras) e dos marcadores prosódicos lexicais (palavras que indicam atitudes do falante e emoções). Assim, cumprimos outro objetivo nosso, a saber, mostrar de que maneira a escrita registra as variações entoacionais/melódicas, atitudes de falantes e emoção.

A presença de um outro sinal de pontuação, ou a maneira como ele é usado, ou o uso de uma ou outra palavra tem direcionamento prosódico específico e implicações de sentido para o texto. Com isso contemplamos o objetivo de discutir a importância dos marcadores prosódicos da escrita na construção do texto escrito e na leitura em voz alta, especificamente quando tratamos do lugar desses marcadores no processamento de leitura.

Entender e assumir que a escrita dispõe de muitos recursos para representar a prosódia da fala, e, com isso, assumir que a escrita possui uma prosódia visual, tem, no nosso entender, implicações importantíssimas para o ensino de escrita, leitura e compreensão de texto. Prosódia implica em sonorização, escuta da fala. Nesse sentido, o ensino de escrita e leitura não pode estar desvinculado de escuta e prática de leitura oral. O aluno precisa ouvir, sentir as variações prosódicas para ser capaz de registrá-las graficamente com eficácia e perceber o efeito de sentido que esses dispositivos acarretam para a compreensão do texto. O ensino de texto deve contemplar a tríade escrita leitura e prosódia de forma conjugada.

## Referências

CAGLIARI, Luiz Carlos. Fonética: uma entrevista com Luiz Carlos Cagliari. *Revista Virtual de Estudos da Linguagem - ReVEL*. Vol. 4, n. 7, agosto de 2006.

- CAGLIARI, Luiz Carlos. *A estrutura prosódica do romance A Moreninha*. Oxford: Estágio Pós-Doutoral, 2002. (Relatório)
- CAGLIARI, Luiz Carlos. Breve história dos sinais de pontuação. In: MASSINI-CAGLIARI, Gladis; CAGLIARI, Luiz Carlos. *Diante das letras: a escrita na alfabetização*. Campinas, SP: Mercado das Letras, 1999.
- CAGLIARI, L.C. Breve História da Pontuação. In: CONGRESSO BRASILEIRO DE LINGÜÍSTICA APLICADA, 4, 1995, Campinas. *Anais...* Campinas: Unicamp, 1995. p.177-183.
- CAGLIARI, Luiz Carlos. *Marcadores prosódicos na escrita*. In: SEMINÁRIO DO GRUPO DE ESTUDOS LINGÜÍSTICOS, 18, 1989, Lorena. *Anais...* Lorena: Grupo de Estudos Linguísticos de São Paulo, 1989. p. 195-203.
- CAGLIARI, Luiz Carlos. *Elementos de fonética do português brasileiro*. Campinas: Unicamp. Tese de (Livre-Docência), 1981.
- CATACH, N. La ponctuation estles systèmes d'écriture: dedans ou dehors? In: DEFAYS, J-M. ROSIER, L. TILKIN, F., (Orgs.). *À qui appartient la punctuation?* Paris: Duculot, 1998. cap. 2, p. 31-46.
- COHEN, H.; DOUAIRE, J.; ELSABBAGH, M. The role of prosody in Discourse. *Brain and Cognition*, v. 46, n. 1-2, p. 73-81, jun/jul. 2001.
- KONDO, T.; MAZUKA, R. Prosodic Planning While Reading Aloud: On-line Examination of Japanese Sentences. *Journal of Psycholinguistic Research*, v. 25, n. 2, p. 357- 381, 1996.
- DIRINGER, D. *A escrita*. LUIZ, A. (trad.). Lisboa: Gris Impressores, 1969. 245p.
- FÓNAGY, I. As funções modais da entonação. *Cadernos de Estudos Linguísticos*, n. 25, 26-66, 1993.
- MATEUS M. M. Sobre a natureza fonológica da ortografia portuguesa. In: PACHECO, V. e MASSINI-CAGLIARI, G. (Orgs.). *Questões de Fonética e Fologia: uma homenagem a Luiz Carlos Cagliari*. *Estudos da Língua(gem)*, v. 3, p. 159-180, 2006.

PACHECO, V. *O efeito dos estímulos auditivo e visual na percepção dos marcadores prosódicos lexicais e gráficos usados na escrita do português brasileiro*. 2006. 349p. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2006.

PACHECO, V.; OLIVEIRA, M.. *Reconhecimento dos marcadores prosódicos da escrita em situação de leitura e de oitiva: um processo interativo*. Revista da Anpoll, Florianópolis, n° 37, p. 199-212, jul./dez 2014.

ROCHA, I.L.V., O sistema de pontuação na escrita ocidental: uma retrospectiva. *DELTA*, v. 13, n.1, 1997

SAUSSURE, F. *Curso de lingüística geral*. 26. ed. São Paulo: Cultrix, 2006.



## Análise variacionista de pausas preenchidas em fronteiras de constituintes

Raquel Meister Ko. Freitag  
Bruno Felipe Marques Pinheiro

Lucas Santos Silva  
Universidade Federal de Sergipe

### 1 Introdução<sup>1</sup>

Truncamentos, falsos começos e alongamentos de segmentos são características da fala que costumam ser associadas à disfluência, e vistos como “sintomas” de uma patologia a ser tratada, como a gagueira, por exemplo. Por outro lado, estas características podem ser consideradas como pistas de reorganização do processo de produção da fala, como advoga Scarpa (1995). Uma destas pistas são as “pausas preenchidas”, ou pausas hesitativas, que ocorrem quando o falante preenche o tempo de seu turno com sons que não configuram itens lexicais de uma dada língua, por meio de alongamentos vocálicos não-enfáticos.

As pausas preenchidas se diferenciam das pausas silenciosas, ou pausas fluentes, em que há um silêncio durante a fala. Essa diferenciação dos tipos de pausa se baseia em aspectos prosódicos e sintáticos. Do ponto de vista sintático, as pausas silenciosas ocorrem em fronteiras de constituintes maiores, enquanto as pausas preenchidas costumam ocorrer em fronteiras de constituintes menores. Do ponto de vista prosódico, as pausas silenciosas apresentam *pitch* mais alto antes e depois, entre grupos acentuais, enquanto as pausas preenchidas ocorrem no interior de grupos acentuais (MERLO, 2006; MERLO; BARBOSA, 2012).

---

1 Este trabalho está vinculado ao projeto “Desenvolvimento de tecnologias sociais para formalização e ressignificação de práticas culturais em Aracaju/SE”, financiado pelo edital Desenvolvimento de Ciência e Tecnologia na Educação Básica (CTI-EB) CAPES/FA-PITEC/SE 05/2014.

Nas transcrições de entrevistas sociolinguísticas, este tipo de fenômeno costuma ser registrado como (hes) para a pausa preenchida, e reticências [...] marcam as pausas silenciosas, sem detalhamento quanto à qualidade acústica do preenchimento ou a sua duração. Em outros tipos de transcrição, as pausas preenchidas no português brasileiro costumam ser registradas como “éh , ah, ahn, mm”; já no português europeu, “aa, aam, mm” (ALMEIDA, 2009, p. 170).

As pistas prosódicas como as pausas preenchidas e as pausas silenciosas, vem sendo objeto de atenção dos estudos da Sociolinguística, seja na caracterização de estilos de fala na abordagem variacionista, já que sua recorrência pode auxiliar no desvelamento dos tipos de sequência discursiva e o grau de familiaridade com o tópico temático, contribuindo para a análise da dimensão estilística da variação (PODESVA, 2008; FREITAG, 2013a, 2015a), seja na abordagem sociofonética, ao lidar com parâmetros de velocidade e taxa de elocução de falantes (KENDALL, 2013), ou, ainda, as pausas preenchidas podem ser consideradas como variantes de uma variável linguística, como sugere a diferença na forma de registro da transcrição e estudos no inglês (TOTIE, 2011; FRUEHWALD, 2016).

Assumindo a perspectiva de que as pausas preenchidas podem ser consideradas variantes de uma mesma variável linguística, neste texto, apresentamos os resultados de uma análise variacionista das pausas preenchidas “ah” e “éh” em entrevistas sociolinguísticas pertencentes ao banco de dados Falares Sergipanos (FREITAG, 2013b). Nosso objetivo é verificar a atuação dos condicionamentos linguísticos (posição), estilísticos (tipo textual/sequência discursiva e tópico temático) e sociais que levam à ocorrência de uma ou outra forma.

## 2 Procedimentos metodológicos

Para este estudo, o *corpus* analisado foi constituído por 14 entrevistas sociolinguísticas que compõem o Banco de Dados *Falares Sergipanos*. (FREITAG, 2013b). A subamostra utilizada refere-se à entrevistas, de cerca de 60-80 minutos cada, realizadas com alunos de ensino médio de escolas públicas pertencentes ao projeto “Desenvolvimento de Tecnologias Sociais para Formalização e Ressignificação de Práticas Culturais em Aracaju/SE” (CTI-EB/FAPITEC/CA-PES). A gravação foi realizada em local sem tratamento acústico específico, em sala de estudos do Colégio Estadual Ministro Petrônio Portela, localizado no Conjunto Augusto Franco no bairro Farolândia, e Colégio Estadual Professor João Costa, localizado no bairro Getúlio Vargas, ambos de Aracaju/SE. O registro foi feito em um gravador de áudio Zoom H4n SP Digital Handy Recorder. As entrevistas seguem o roteiro definido pelas diretrizes de coleta do banco de dados, obedecendo à seguinte estrutura: autoconhecimento, moradia, globali-

dade, fatores sociais, lazer, educação, esporte, cultura e variação linguística. A padronização do roteiro permite o controle do tópico temático e do tipo textual/sequência discursiva das partes da entrevista, o que nos permite avaliar os efeitos estilísticos.

Os áudios das entrevistas foram transcritos no software ELAN (WITTENBURG et al., 2006). Após a transcrição, as 14 entrevistas foram examinadas para a identificação das pausas preenchidas: foram identificadas 499 ocorrências de pausas preenchidas ‘ah’ e ‘eh’, que, após cotejados às variáveis linguísticas, estilísticas e sociais, foram submetidos à análise estatística.

Posteriormente, para análise acústica, foram recortados dos áudios 60 excertos de contexto de pausas preenchidas (30 de falantes masculinos e 30 falantes femininos), e submetidos à análise no software Praat versão 5.1.31 (BOERSMA; WEENINK, 2009), em que as pausas preenchidas foram especificadas em forma de onda, espectograma, contorno da frequência fundamental com base nos pressupostos da Fonologia Prosódica e Entonacional.

## 3 Resultados e discussão

### 3.1 Análise prosódica

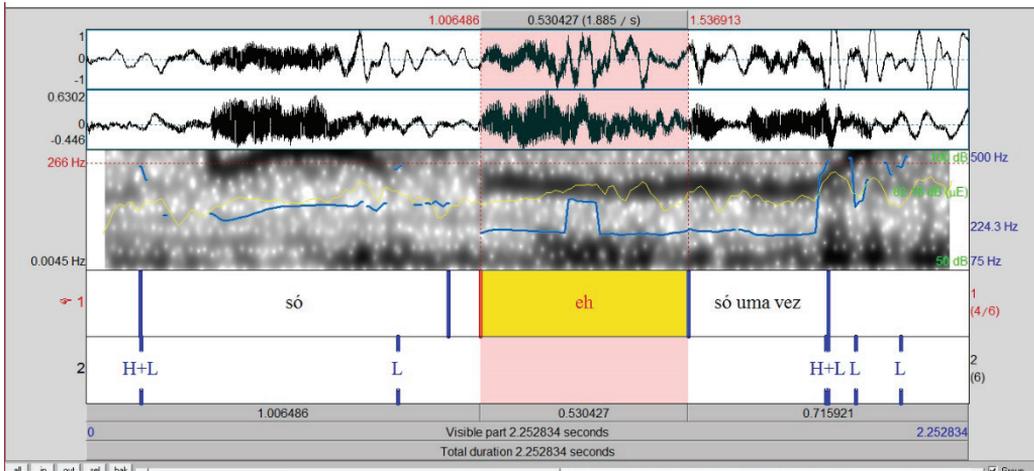
Segundo Barbosa (2012),

a prosódia está, no cenário da pesquisa atual, associada a fatores linguísticos como acento, fronteira de constituinte, ênfase, entoação e ritmo, a fatores paralinguísticos como marcadores discursivos (e.g., “né”, “entendo”, “an-han”) e atitudes proposicionais (e.g., “confiante” e “duvidoso”) e sociais (e.g., “hostil” e “solitário”), além de tratar de fatores extralinguísticos como as emoções. Todos esses fatores se combinam com aspectos sociais e biológicos indiciais como gênero, faixa etária, classe social, nível de escolaridade, entre outros. (BARBOSA, 2012, p. 13).

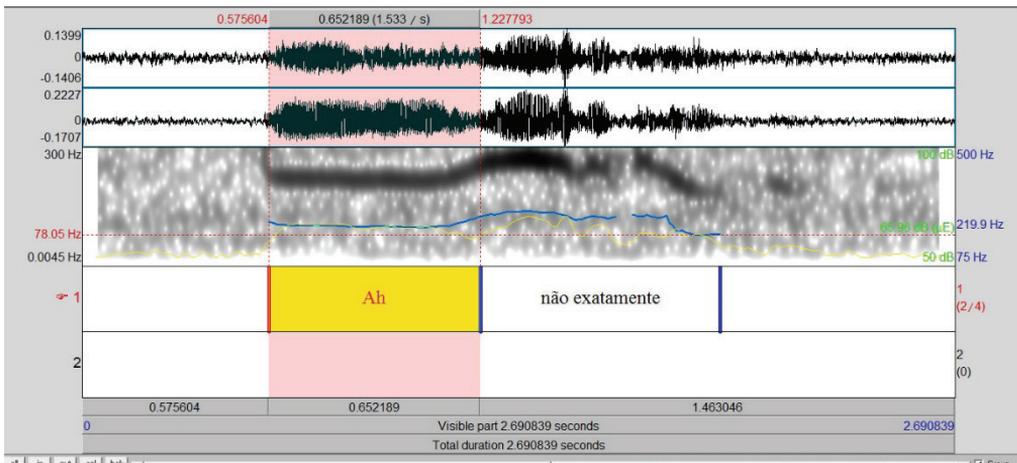
Assim, a análise prosódica contempla desde a análise de pequenas unidades fônicas, como é o caso das pausas preenchidas, até suas relações a partir das sílabas aos níveis gramaticais, mais altos, como o nível discursivo. Nossa análise considera o contorno entonacional dos alongamentos vocálicos “ah” e “eh” na função de pausas preenchidas.

Dos 60 excertos analisados, trazemos 4, sendo dois de pausas “ah” e dois de “eh” no programa *Praat*, a fim de ilustrarmos a forma de onda e espectograma dessas pausas preenchidas, com base no pressuposto de que a entoação tem uma

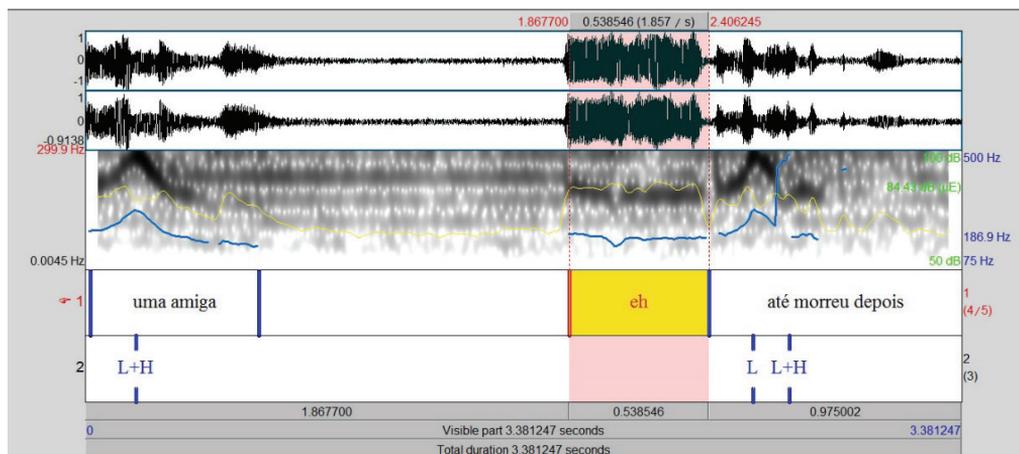
organização fonológica e sua relação com a sintaxe, pois na análise também se levou em consideração a posição sintática em que as pausas ocorrem (figuras 1 a 4).



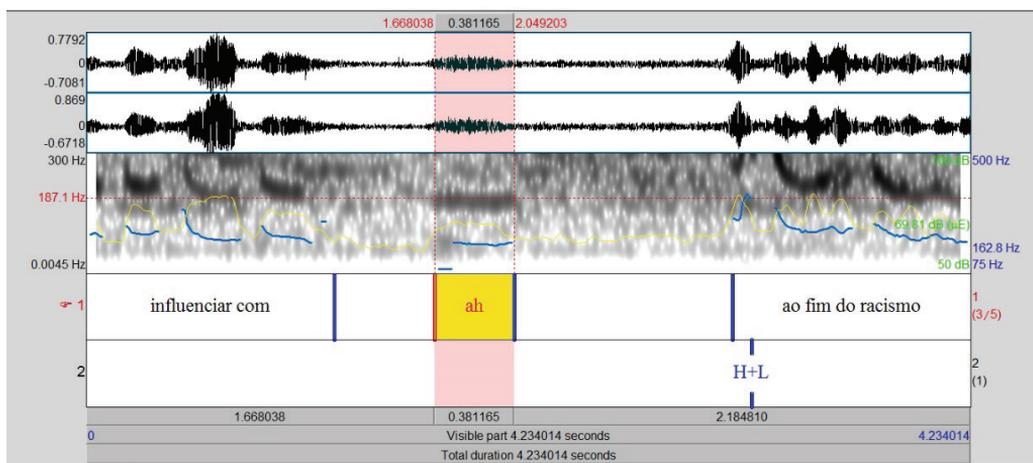
**Figura 1** Forma de onda, espectrograma, contorno de  $F_0$  e transcrição tonal do trecho “só eh uma vez” em posição/função repetições, correções e falsos começos (Gli\_Pet\_29\_09\_16).



**Figura 2** Forma de onda, espectrograma, contorno de  $F_0$  e transcrição tonal do trecho “ah não exatamente” em posição/função tomada/retomada de turno (and\_joã\_29\_09\_2016).



**Figura 3** Forma de onda, espectrograma, contorno de  $F_0$  e transcrição tonal do trecho “uma amiga eh até morreu depois” em posição/função sujeito e predicado (Ing\_Pet11\_10\_16).



**Figura 4** Forma de onda, espectrograma, contorno de  $F_0$  e transcrição tonal do trecho em posição/função “influenciar com ah o fim do racismo” verbo e complemento (lor\_joã\_04\_10\_2016).

Os contornos entonacionais não são somente afetados por aspectos gramaticais, mas também por fatores semânticos e pragmáticos. Características como desempenho do falante, velocidade da fala e estilo pessoal sugerem que, para uma definição das pausas preenchidas, também deve ser considerado o fator de variabilidade (NESPOR; VOGEL, 1986).

De acordo com o modelo de Fonologia Prosódica (NESPOR; VOGEL, 1986), o fluxo de fala é organizado hierarquicamente dentro de domínios prosódicos e

uma hierarquia prosódica que vai desde a sílaba até o enunciado, e que regras fonológicas se aplicam no interior de certos domínios. Logo, as pausas preenchidas ocupariam uma posição/função demarcativa ou de segmentação, constituindo uma fronteira prosódica.

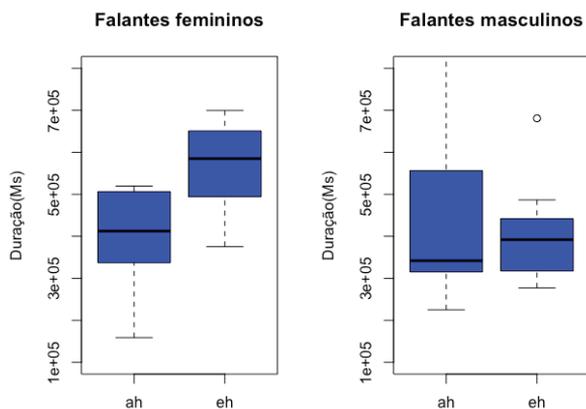
Segundo Scarpa (2006), as pausas ocorrem em três lugares no enunciado: (i) fronteiras de constituintes maiores (sujeito/predicado), (ii) antes de palavras de alto conteúdo lexical (determinante/cabeça de algum sintagma) e (iii) depois da primeira palavra de um grupo entonacional (repetições/ falsos começos).

Considerando que a estrutura prosódica é parcialmente determinada pela estrutura sintática, como evidenciado nas figuras 3 e 4. Nesses excertos, suas estruturas coincidem com a estrutura sintática, pois suas fronteiras de constituintes maiores quando ocorrem, acontecem na posição entre sujeito e predicado, verbo/complemento.

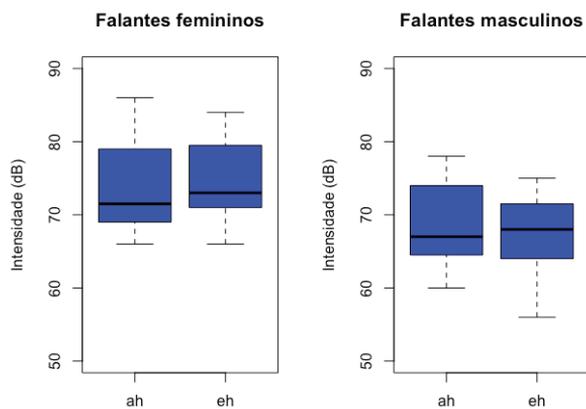
As pausas que ocorrem em fronteiras de constituintes maiores, como sujeito/predicado e verbo/complemento, são chamadas de pausas de respiração, que, por sua vez, não são sinônimos de disfluência, mas sim pausas fluentes pois costumam se situar em fronteiras sintáticas fortes. Há correlação entre o tipo de fronteira em que ocorre a pausa e a sua duração: nas figuras 3 (0,538546ms) e 4 (0,381135ms), a duração das pausas é muito curta, possivelmente porque os pontos das fronteiras de constituintes com duração maior envolvem um novo tópico, logo, as pausas tendem a ser mais curtas.

Já as figuras 1 e 2 referem-se a pausas que marcam uma nova posição sintática de repetições/correções/falsos começos e tomada/retomada do falante em relação à pergunta do entrevistador. Este tipo de pausa é marcado por uma duração maior: 0,0530427ms e 0,652189ms, respectivamente. Este tipo de ambiente em que ocorre a pausa está relacionado ao processamento, sinalizando um planejamento on-line para retomar o discurso, com o tempo preenchido com o que chamamos de “pausa preenchida”. Ao dividir o enunciado em constituintes prosódicos maiores, as pausas preenchidas acabam assumindo também uma função conversacional. Por isso, “esta é a razão por que a sintaxe nem sempre faz previsões corretas sobre a estrutura prosódica” (SCARPA, 2012, p. 30), e neste caso, o exemplo acima não corresponde a nenhuma teoria sintática estabelecida. As pausas preenchidas não necessariamente são disfluências, pois as pausas não são aleatoriamente distribuídas em domínios prosódicos, e não ocorrem em certos pontos do enunciado prosódico.

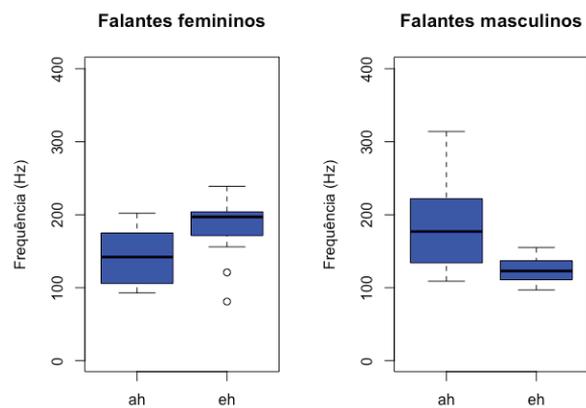
Constituímos uma subamostra de 60 excertos de pausas preenchidas, sendo 30 de “ah” e “eh”, com os falantes estratificados em função do sexo/gênero, que foram submetidos a tratamento acústico a fim de identificar os valores da intensidade, de *pitch* e duração. Os resultados são apresentados nas figuras 5 a 7.



**Figura 5** Duração das pausas preenchidas em função de sexo/gênero.



**Figura 6** Intensidade das pausas preenchidas em função de sexo/gênero.



**Figura 7** Frequência de pausas preenchidas em função de sexo/gênero.

Os boxplots apresentam  $F_0$  (frequência fundamental) dos contornos entonacionais das pausas preenchidas “ah” e “eh” medidas em Hz e o valor de *pitch*, da intensidade e a duração e também relacionados à máxima e mínima altura tonal em um ponto, tais parâmetros permitem observar as diferenças de comportamento entre os informantes da amostra quanto ao sexo/gênero.

As pausas preenchidas realizadas por falantes femininos são mais intensas do que as realizadas por falantes masculinos, havendo inclusive uma inversão: na fala masculina, a pausa “eh” apresenta resultado maior de intensidade do que a fala “ah”, ao contrário do que ocorre na fala feminina.

O valor de *pitch* das pausas preenchidas “ah” produzidas por falantes masculinos é maior em comparação com o valor de *pitch* das pausas dos falantes femininos. Já nas pausas preenchidas “eh”, o valor de *pitch* das falantes de sexo/gênero feminino se sobressai em relação aos falantes de sexo/gênero masculino; a intensidade é um parâmetro físico que controla diretamente a sensação de *pitch*, apontando para diferenças fisiológicas de sexo/gênero (FREITAG, 2015b; LEITE, 2012): as falantes tendem a ter uma  $F_0$  maior do que os falantes para as pausas preenchidas “eh”.

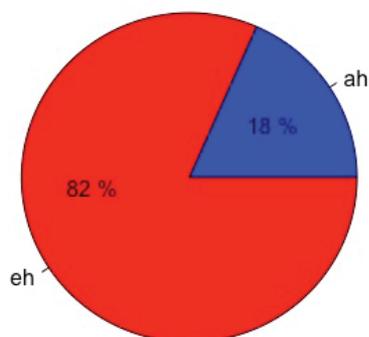
Assim, a duração, intensidade e o *pitch* que diferenciam as pausas preenchidas quanto ao sexo/gênero dos informantes. Outra característica é a duração: as pausas preenchidas “eh” têm uma maior duração das pausas preenchidas “ah”, essa duração é percebida pelas unidades medidas em milisegundos e “é o parâmetro de controle por excelência” (BARBOSA, 2012, p. 20).

A duração está associada às fronteiras de domínios prosódicos e que indicam a posição das pausas preenchidas. As pausas preenchidas “ah” “eh” ocupam diferentes posições sintáticas nos enunciados e, por isso, assumem diferentes funções, vejamos como a distribuição dos comportamentos se dá em uma análise variacionista.

### 3.2 Análise sociolinguística

A análise quantitativa das pausas preenchidas aponta que, das 499 ocorrências identificadas no corpus, 92 referem-se a pausa “ah”, e 407 referem-se à pausa “eh” (Figura 8). Por isso, elegemos como valor de aplicação a pausa “eh”, em função da qual os resultados serão apresentados a seguir.

**Distribuição dos tipos de pausa  
(n = 499)**



**Figura 8** Distribuição das pausas preenchidas “ah” e “eh”.

As pausas “eh” e “ah” foram correlacionadas aos fatores: posição da pausa, tipo de texto/sequência discursiva, tópico temático e sexo/gênero do informante.

As posições em que as pausas preenchidas ocorrem são as seguintes:

**Entre verbo e complemento:**

- (1) “foi eh melhor do que está agora né?” (Adr\_Pet-29\_09\_16)
- (2) “então acho ah os alunos daqui” (Ing\_Pet\_11\_10\_16)

**Enumeração:**

- (3) “crônica eh romance contos poemas depende”(ali\_joã\_29\_09\_2016)
- (4) “ah alguns jogos de ação RP” (ali\_joã\_29\_09\_2016)

**Entre determinante e sintagma:**

- (5) “eu acho a escola tipo muito eh organizada” (Adr\_Pet-29\_09\_16)
- (6) “a disputa ah ah por conta” (and\_joã\_29\_09\_2016)

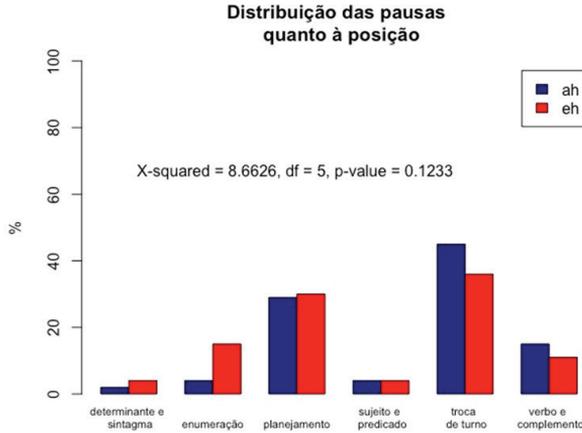
**Entre sujeito e predicado:**

- (7) “o trabalho eh é algo que a gente vai ter que acostumar”(and\_joã\_29\_09\_2016)
- (8) “mulheres ah a se dar o valor” (reb\_joã\_04\_10\_2016)

**Funções interacionais:**

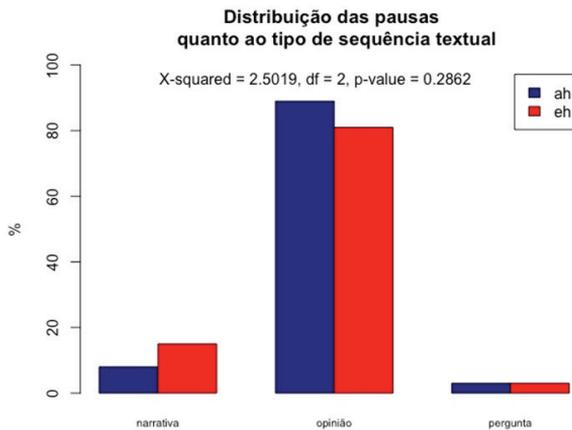
- (9) “F1: qual? F2: eh .. Eminem” (jos\_joã\_28\_09\_2016)
- (10) “F1: qual foi a sua melhor? F2: ah a melhor o quê, festa?”(ali\_joã\_29\_09\_2016)

Os resultados da variável posição/função da pausa preenchida se distribuem em duas direções: nas fronteiras de constituintes maiores (posição sujeito/predicado e funções interacionais), em que a pausa preenchida “eh” tende a ser mais recorrente, e em fronteiras de constituintes menores (determinante/sintagma, enumerações), a pausa preenchida “eh” tende a ser menos recorrentes (Figura 9).



**Figura 9** Resultados da variável posição quanto à distribuição das pausas preenchidas.

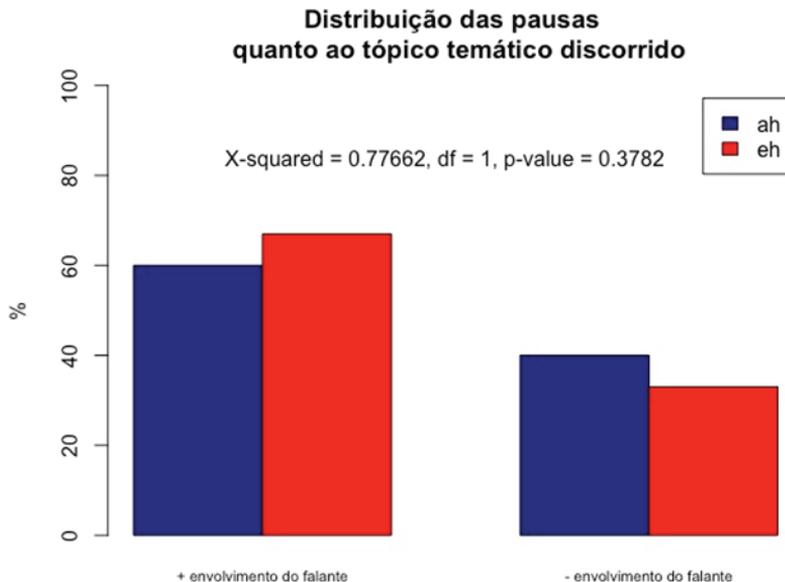
As pausas utilizadas em fronteiras de constituintes maiores são chamadas de pausas de respiração e não indicam uma disfluência. Já as pausas que ocorrem em funções interacionais e de constituintes menores geralmente são chamadas de “fenômenos de hesitação”. Parece haver uma tendência de a pausa “ah” ser associada à não disfluência e a pausa “eh”, à disfluência. As diferenças identificadas na análise prosódica são corroboradas pela análise variacionista.



**Figura 10** Resultados da variável tipo textual/sequência discursiva quanto à distribuição das pausas preenchidas.

Os tipos de sequência discursiva e o grau de familiaridade com o tópico temático costumam ser associados à dimensão estilística da variação (PODESVA, 2008; FREITAG, 2013a, 2015a). Moniz (2006) aponta uma correlação entre o uso das pausas e o tipo textual/sequência discursiva, pois em textos considerados mais formais se faz necessário uma busca mais elaborada de itens lexicais, favorecendo a um planejamento discursivo maior para construir períodos. Seria esperado que as pausas que denotam maior tempo de processamento fossem mais produtivas em sequências argumentativas, que envolvem maior tempo de planejamento, como podemos constatar na distribuição em relação aos tipos de sequência opinativa. A recorrência da forma “ah” está associada aos contextos de opinião, enquanto a forma “eh” mostra-se mais recorrente na narrativa, como podemos ver na Figura 10.

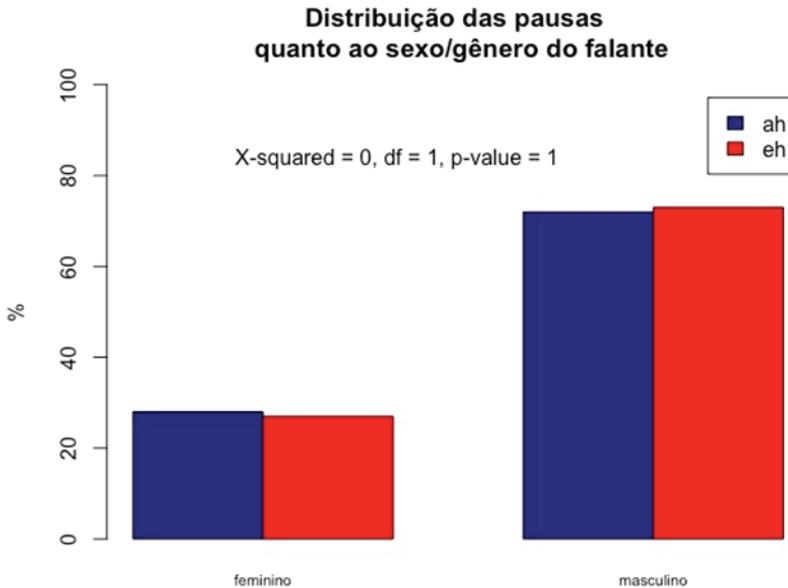
Assim como o tipo textual/sequência discursiva, o assunto sobre o qual está se discorrendo poderia influenciar a ocorrência de pausas preenchidas, denotando uma relação de maior ou menor subjetividade e, em consequência, diferentes demandas de processamento. Observamos correlação entre a pausa “ah” e tópicos com menor envolvimento pessoal do falante, e a pausa “eh”, a tópicos com maior envolvimento do falante, ou seja, existe uma tendência à especialização da forma quanto à sinalização da fluência/disfluência (Figura 11).



**Figura 11** Resultados da variável tópico temático quanto à distribuição das pausas preenchidas.

As diferenças fisiológicas, que refletem no *pitch* e na intensidade, do ponto de vista variacionista, se evidenciam em um maior uso deste recurso por parte dos

falantes do sexo/gênero masculino; este resultado requer um exame mais acurado, com a correlação do uso de outras formas para o desempenho das funções relacionadas à fluência, como marcadores discursivos (Figura 12).



**Figura 12** Resultados da variável sexo/gênero quanto à distribuição das pausas preenchidas.

Em suma, os resultados obtidos corroboram que pausas preenchidas obedecem a princípios de distribuição que determinam não apenas onde um elemento pode ocorrer, mas também reforçam sua função de manter o discurso fluente. As pausas preenchidas fazem parte da percepção da fala e de uma seleção de valores linguísticos-fonéticos e assim não necessariamente são sinônimos de disfluência, mas sim “um produto de dois tipos de informação: a voltada ao sinal e a independente do sinal; neste último caso, a informação relevante é de natureza sobretudo linguística” (SCARPA; FERNANDES-SVARTSMAN, 2012, p. 26). A motivação subjacente às pausas é consequência do processamento cognitivo, que se utiliza desses recursos com o objetivo de garantir ao falante um tempo necessário para o planejamento mais adequado do seu discurso. E, em uma escala de gradiência, “eh” e “ah” estão em pontos diferentes: os resultados distribucionais apontam que a pausa “ah” está associada à não disfluência e a pausa “eh”, à disfluência.

## 4 Considerações finais

Partimos da hipótese de que as pausas preenchidas não são necessariamente um fator de disfluência, mas sim indicativo de formulação e processamento, sendo

parte constitutiva da fala. As pausas preenchidas “ah” e “eh”, na amostra analisada, se comportam como hesitações e não ocorrem aleatoriamente do ponto de vista prosódico, “uma vez que sua ocorrência respeita o padrão geral de comportamento prosódico do português brasileiro” (SCARPA; FERNANDES-SVARTSMAN, 2012, p. 38). Os resultados apontaram para uma diferença de uso entre “ah” e “eh”, o que pode sugerir que a pausa “ah” está associada à não disfluência e a pausa “eh”, à disfluência.

Com este estudo, contribuímos para a caracterização fonética e acústica do português falado em Sergipe, ainda incipiente (SOUZA; SOUZA NETO, FREITAG, 2016; NUNES, 2015), e também para o tratamento de pistas prosódicas na caracterização de estilos de fala na abordagem da Sociolinguística Variacionista, ainda incipiente também.

## Referências

- ALMEIDA, V. B. Pausas preenchidas e domínios prosódicos: evidências para a validação do descritor fluência em um teste de proficiência oral em língua estrangeira. *Alfa: Revista de Lingüística*, v. 53, n. 1, p. 167-195, 2009.
- BARBOSA, A. P. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos Linguísticos*, vol. 20, n.1, p.11-27, 2012.
- BOERSMA, P. ; WEENINK, D. *Praat: doing phonetics by computer* (version 5.1.15). Retrieved August 30, 2009. Disponível em: [http://: www.praat.org/](http://www.praat.org/).
- FREITAG, R. M. K. Pistas prosódicas para a segmentação da entrevista sociolinguística. *Anais do Colóquio Brasileiro de Prosódia da Fala*, v. 2, p. 1-5, 2013a.
- FREITAG, R. M. K. (Re)discutindo sexo/gênero na sociolinguística. In: FREITAG, R. M. K., SEVERO, C. G. (org.). *Mulheres, Linguagem e Poder - Estudos de Gênero na Sociolinguística Brasileira*. São Paulo: Blücher, 2015b, p. 17-74
- FREITAG, R. M. K. Socio-stylistic aspects of linguistic variation: schooling and monitoring effects. *Acta Scientiarum. Language and Culture*, v. 37, n. 2, p. 127-136, 2015a.
- FREITAG, R. M. K. Banco de dados falares sergipanos. *Working Papers em Linguística*, v. 14, n. 2, p. 156-164, 2013.

- FREITAG, R. M. K.; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do português brasileiro e os estudos de terceira onda: potencialidades e limitações. *Alfa*, v. 56, p. 917-944, 2012.
- FRUEHWALD, J. Filled pause choice as a sociolinguistic variable. *University of Pennsylvania Working Papers in Linguistics*, v. 22, n. 2, p. 41-49, 2016.
- KENDALL, T. *Speech rate, pause and sociolinguistic variation: studies in corpus sociophonetics*. Nova Iorque: Palgrave/MacMillan, 2013.
- LEITE, C. M. B. Intersecção entre variação linguística dos róticos e a variável sexo. *Estudos Linguísticos*, v. 41, n. 2, p. 755-764, 2012.
- MERLO, S. *Hesitações na fala semi-espontânea: análise por séries temporais*. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade de Campinas: Campinas, 2006.
- MERLO, S.; BARBOSA, P. A. Séries temporais de pausas e de hesitações na fala espontânea. *Cadernos de Estudos Linguísticos*, v. 54, n. 1, p. 11-24, 2012.
- MONIZ, H. G. S. *Contributo para a caracterização dos mecanismos de (dis)fluência no português europeu*. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2006.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.
- Nunes, V. G. *A prosódia de sentenças interrogativas totais nos falares catarinenses e sergipanos*. Tese. Doutorado em Linguística, Universidade Federal de Santa Catarina, 2015.
- PODESVA, R. J. Three sources of stylistic meaning. *Texas Linguistic Forum*, p. 1-10, 2008.
- SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.
- SCARPA, E. M. Sobre o sujeito fluente. *Cadernos de Estudos Linguísticos*, vol. 29, p. 163-184, 1995.

SCARPA, E. M. FERNANDES-SVARTSMAN, F. A Estrutura prosódica das disfluência em português brasileiro. *Cadernos de Estudos Linguísticos*, vol. 54, n. 1. p. 25-40, 2012, (p. 25-40).

SOUZA, G. G. A.; SOUZA NETO, A. F.; FREITAG, R. M. K. . As vogais médias [e] e [o]: um estudo fonético-acústico e comparativo. In: FREITAG, R. M. K.; SEVERO, C. G., GÖRSKI, E. M. (Org.). *Sociolinguística e Política Linguística: Olhares Contemporâneos*. São Paulo: Editora Blucher, 2016, p. 21-34.

TOTTIE, G. Uh and um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics*, v. 16, n. 2, p. 173-197, 2011.

WITTENBURG, P. et al. Elan: a professional framework for multimodality research. In: *Proceedings of LREC*. 2006.



## Prosódia e fonoaudiologia: do fonostilo ao transtorno da linguagem

Aveliny Mantovan Lima  
Universidade de Brasília

Ana Carolina Constantini  
Unicamp

### 1 Introdução

A prosódia é um componente da fala, uma informação fônica, que está além do nível do segmento e que permite a estruturação e a organização dos enunciados em diversas línguas. De acordo com Barbosa e Madureira (2015), cabe à prosódia a coordenação dos gestos articulatórios ao longo do enunciado e, por isso, a prosódia molda a enunciação e, segundo os autores, imprime “ao que se fala” um “modo de falar”, que pode ser dirigido intencionalmente ou não ao ouvinte.

O estudo da prosódia pode ser dividido em estudos de produção e de percepção. No que diz respeito aos aspectos de produção da prosódia, três parâmetros fonético-acústicos são tradicionalmente estudados: duração, variação da frequência fundamental e variação da intensidade. Tais parâmetros fazem a intermediação entre a forma e o conteúdo, propriamente dito, do discurso (LOPES; LIMA, 2014). Os aspectos de percepção são identificados pelas noções de duração percebida, altura e volume (BARBOSA, 1999).

Além disso, a prosódia pode assumir diversas funções, segundo Barbosa e Madureira (2015), dentre elas distinguem-se as funções prosódicas, a saber:

- (a) atitudinais, que incluem atitude, postura interpessoal, estilo de elocução;
- (b) afetivas, que abrangem emoções como tristeza, alegria e raiva, e afetos como o humor;
- (c) indiciais, que compreendem gênero e origem dialetal.

Neste capítulo, abordamos a função prosódica indicial e a função prosódica atitudinal, especificamente ao que se relaciona ao estilo de elocução, ou, em seu sentido mais amplo, à fonostilística e comentamos as aplicações dos estudos prosódicos na comparação de locutores, tarefa comum realizada na fonética forense. O uso de um importante instrumento para avaliação da qualidade vocal a partir

de um modelo fonético é comentado. A função afetiva é abordada com relação ao transtorno de linguagem, mais especificamente, ao autismo.

Por fim, discutimos a necessidade de uma avaliação clínica que considere as habilidades expressão (como o sujeito produz) e recepção (como o sujeito compreende) da prosódia, a partir de um teste estruturado que está em fase final de adaptação para o português brasileiro.

## 2 Fonoestilística

A fonoestilística diz respeito à função social e organizadora da comunicação e trata das diferenças fonéticas perceptíveis ao indivíduo. Por possuir um caráter expressivo durante a comunicação, permite que diferentes tipos de informação sejam passados ao receptor da mensagem.

A atuação fonoaudiológica na fonoestilística é mais recente quando comparada com outros campos de atuação, e vem pautada em uma perspectiva relacionada ao aprimoramento do uso vocal e com a habilitação para o uso profissional da voz, ou seja, focada em uma perspectiva de promoção de saúde.

Dentro do campo da fonoestilística, a expressividade adquire papel fundamental e, segundo Madureira (2005), os recursos fônicos veiculam efeitos de sentido, estreitando assim, as relações entre som e sentido do que se quer dizer. Portanto, o estilo de fala pode ser ajustado pelo falante, de acordo com cada situação comunicativa.

Os diferentes estilos de fala são, geralmente, associados a determinadas profissões e, por isso, acabam por caracterizar social e profissionalmente uma categoria. Estudos demonstram que, mesmo com a retirada do conteúdo semântico e lexical de uma amostra de fala, os ouvintes ainda são capazes de diferenciar o estilo de fala de uma locução jornalística e de uma fala espontânea, como em uma situação de entrevista, por exemplo (CASTRO et al, 2010).

Dentre os diversos fonoestilos, podem-se destacar o do jornalismo, do radiolismo, dos políticos e religiosos (os dois últimos geralmente adquirem um caráter persuasivo no processo comunicativo).

Considerando os parâmetros prosódicos clássicos, os estudos da área focam na duração das pausas (preenchidas e silenciosas), taxa de elocução, média da frequência fundamental e tipos de contornos de frequência fundamental.

O fonoestilo utilizado no telejornalismo combina traços de leitura oral e de fala espontânea, pois normalmente o telejornalista, em seu discurso, conta com apoio visual da leitura, mas busca uma produção com características de fala espontânea. Por isso, o trabalho com parâmetros prosódicos no fonoestilo do telejornalismo é de fundamental importância (CONSTANTINI, 2012). Aspectos de duração e frequência fundamental, por exemplo, podem ser trabalhados em concomitância para a produção de ênfases.

O uso de pausas também pode ser um recurso importante na estilo de narração telejornalístico. A inclusão de pausas silenciosas pode criar uma situação de suspense e trazer ênfase para o que será dito a seguir e pode ser um recurso expressivo de grande valor no telejornalismo. O uso de pausas preenchidas, por outro lado, não é um recurso presente nesse fonoestilo (CASTRO et al, 2010) e é frequentemente observado em outros estilos de elocução, como a fala espontânea.

O estilo de elocução presente no radialismo difere de acordo com o público-alvo. O locutor radialista assume grande versatilidade em suas funções, apresentando programas diversos, locução comercial, locução de vinhetas e, muitas vezes, realiza eventos sociais como mestres de cerimônia. As rádios FM direcionadas para o público jovem utilizam narrações com registro agudo, *loudness* aumentada e taxa de elocução elevada e maior variação melódica, enquanto que aquelas direcionadas para um público mais velho utilizam registro mais grave, *loudness* e taxa de elocução médias, além de menor variação melódica. Em relação às rádios AM, a locução prefere uso das marcas pessoais do radialista, com uso mais coloquial da voz (BORREGO, 2005).

Em função da necessidade de dar ênfase ao texto dito, pode ser observado um aumento significativo da variação de  $F_0$ , como marca da proeminência. A gama tonal durante a locução tende a ser maior do que na fala coloquial, para proporcionar uma maior ênfase ao texto com a variação de entoação que se julgar necessária. Por se tratar de um meio de comunicação que utiliza apenas a voz para transmitir as mensagens, acredita-se que, em comparação ao telejornalismo, a locução radialista faz uso de maiores recursos de variação da curva melódica, para atrair a atenção de seus ouvintes, de acordo com suas características socio-culturais (CAMPOS, 2012).

De acordo com Irvine (2001), as características de um estilo não podem ser explicadas independentemente de outros estilos, assim, a relação entre diferentes estilos, como seus contrastes, limites e semelhanças podem revelar informações importantes. Um estudo comparando o estilo de locução de telejornalistas, fala espontânea (situação de entrevista), políticos e religiosos mostrou que o estilo religioso e o político utilizam frequência fundamental com maior variação entre frequências altas, enquanto que o telejornalismo e a fala espontânea utilizam registros mais baixos.

Ainda nos estudos da fonoestilística, a qualidade vocal pode ser explorada através de seus aspectos de produção e percepção e, segundo Camargo, Madureira e Schmitz (2013), para além dos estudos de patologias que afetam a qualidade vocal, investigação deste parâmetro envolve também suas propriedades estéticas que estão relacionadas com a expressividade.

Em termos de instrumentos de avaliação destas características, citaremos o *Vocal Profile Analysis Scheme* (VPAS), proposto por Laver (1980), a partir da ne-

cessidade de elaboração de um modelo fonético de avaliação da qualidade vocal. O instrumento foi traduzido para o português brasileiro por Camargo e Madureira no ano de 2008 e, desde então, vem sendo amplamente utilizado em pesquisas no campo de estudos da prosódia como uma ferramenta que possibilita a análise dos ajustes de longo termo realizados pelo trato vocal modelo (chamado de *setting* neutro) e as suas características desviantes, também chamados de *settings*, que podem ser mais comuns em determinadas qualidades vocais. Cabe ressaltar que, os diferentes ajustes observados podem variar de acordo com o conteúdo do discurso e com o fonostilo estudado. O uso de recursos prosódicos nos fonostilos religioso e político, por exemplo, pode assumir ajustes desviantes e apresentar mudanças nas características da qualidade vocal a depender da atitude e das emoções a serem expressadas. Desta forma, o VPAS mostra-se um instrumento particularmente importante ao descrever ajustes laríngeos, supralaríngeos e de tensão presentes na qualidade vocal resultante e que traz informações valiosas nos estudos da fonostilística. Para maior aprofundamento deste instrumento, sugere-se a leitura de Camargo e Madureira (2008), Camargo, Madureira e Schmitz (2013) e Laver (1980).

A atuação fonoaudiológica voltado ao aprimoramento vocal dos diversos fonostilos pode ser iniciada desde a formação destes profissionais, com a participação dentro de cursos de graduação e cursos preparatórios e, já durante o exercício profissional, sob forma de consultoria. Especialmente, a atuação fonoaudiológica ainda no período de formação de futuros profissionais pode trazer diversos benefícios, já que a maioria dos profissionais que ingressa no mercado de trabalho busca inspiração em profissionais já consagrados, e acabam por usar de maneira inconsciente (nem sempre de forma adequada) diversos parâmetros prosódicos em suas falas (CASTRO, 2008, KYRILLOS, 2004).

### 3 Prosódia e fonética forense

Os parâmetros prosódicos, a partir da sua função indicial, também podem ser explorados no estudo da identificação de falantes. Especificamente, na tarefa de comparação de locutor, que é a mais frequente realizada por foneticistas forenses, a análise de parâmetros prosódicos clássicos tem apresentado resultados promissores, principalmente no que diz respeito ao parâmetro da duração (BARBOSA, 2006; WIGET et al, 2010; YOON, 2010; DELLWO, 2010; LEEMANN et al, 2014). Este parâmetro acústico tem relação direta com o ritmo da fala e pode ser definido como a variação a longo termo da duração percebida. O ritmo da fala não está somente restrito ao segmento e, sendo assim, assume um papel importante na organização global da fala em uma sentença; além disso, o ritmo é altamente individual, o que favorece as análises da fonética forense cujo obje-

tivo pode ser encontrar justamente as características mais particulares da fala de um sujeito. Diversos são os parâmetros que podem ser estudados, dentre eles o DeltaC, desvio-padrão dos intervalos consonantais, porcentagem de sonorização em um enunciado, diferença de duração entre intervalos vocálicos, entre outros.

Como abordado anteriormente no presente capítulo, o VPAS também pode contribuir para a descrição das características de qualidade vocal na área da fonética forense. Diversos autores consideram a qualidade vocal como um importante identificador de sujeitos e que pode ser útil em tarefas na tarefa de comparação de locutores. Atualmente, há uma proposta de versão modificada do protocolo original do VPAS (LAVER, 1980), proposta por San Segundo e colaboradores (2016) que sugere a eliminação de alguns *settings* que, segundo os autores, seriam de difícil avaliação perceptiva, como o abaixamento de laringe. Destaca-se que este protocolo modificado foi proposto recentemente e necessita de maiores investigações a respeito das modificações realizadas. Entretanto, dada a importância da qualidade vocal para a área da fonética forense, o uso do VPAS neste campo de estudo é promissor e pode contribuir para o avanço da área em questão.

## 4 Transtorno de linguagem

A prosódia é um dos aspectos essenciais para o desenvolvimento adequado da linguagem. Sabe-se que a criança tem a habilidade para entender as características prosódicas desde o nascimento (GERVAIN; MEHLER, 2010) e para usar a prosódia desde a fase mais inicial de seu desenvolvimento de linguagem (NAZZI; BERTONCINI; MEHLER, 1998), antes mesmo, portanto, de ter completado a aquisição dos níveis mais formais da linguagem. Entretanto, certos aspectos da prosódia não são dominados de maneira semelhante aos adultos antes da puberdade (WELLS; PEPPÉ; GOULANDRIS, 2004).

Embora o desenvolvimento típico da prosódia, especificamente, possa demorar até a puberdade, algumas alterações podem ser percebidas bem cedo. Os padrões prosódicos atípicos agem como barreiras significativas para a comunicação e podem também afetar o processo de aquisição de linguagem (FILIPE, 2014). Esse pode ser o caso de alguns indivíduos com transtorno do espectro autista, ou autismo. As características prosódicas do autismo serão descritas com base na função prosódica afetiva.

### 4.1 Prosódia e autismo

O autismo é uma desordem pervasiva do desenvolvimento que apresenta como características desenvolvimento atípico em relação à interação social, comunicação atípica e comportamento restrito, estereotipado e repetitivo (WING; GOULD, 1979; AITKEN et al., 1998). As principais marcas prosódicas que estão presentes na fala

dos indivíduos com autismo são as inflexões “pobres” (pouco *range*), acentuação excessiva ou inadequada (HARGROVE, 1997), entoação exagerada ou monótona, baixa ou alta taxa de elocução e a adoção de acento diferente em um par de palavras (BARON-COHEN; STAUNTON, 1994). Alguns dos termos que a literatura traz sobre a prosódia da fala do sujeito com autismo refletem mais as suas características perceptivas do que as suas características acústicas (PEPPÉ et al., 2007).

Em estudo de revisão crítica de literatura (MCCANN; PEPPÉ, 2003), as autoras relataram preocupação com a falta de um método consistente para a investigação das características prosódicas alteradas nos sujeitos com autismo, bem como levantaram muitos problemas nos estudos encontrados, tais como: (a) poucos estudos utilizaram a análise acústica para quantificar a prosódia expressiva; (b) incerteza de definição em alguns parâmetros prosódicos, como o acento, o que pareceu resultar em alguns achados conflitantes; (c) poucos estudos em áreas como a da prosódia afetiva, interacional e de compreensão da prosódia, sendo que nenhum abordou a compreensão no contexto de habilidades expressivas; (d) nenhum estudo abrangeu uma ampla gama de habilidades prosódicas expressivas e receptivas, tornando impossível investigar as relações entre a capacidade receptiva e expressiva dos sujeitos autistas. Para as autoras, a falta de sistematização na avaliação tornou difícil a condução terapêutica (e até facilitou a negligência) de aspectos da prosódia nessa população.

Em busca de parâmetros de referência para a caracterização das habilidades prosódicas na população com autismo, Peppé et al. (2007) aplicaram um teste de avaliação da competência prosódica em crianças com autismo de alto funcionamento. O teste, chamado PEPS-C, será explicado na próxima seção deste capítulo. Os autores observaram que as crianças do grupo com autismo apresentaram desempenho significativamente menor do que as crianças do grupo controle em praticamente todas as tarefas do teste e, portanto, o desenvolvimento prosódico no grupo das primeiras parece ser atrasado em muitos aspectos da prosódia e desviante em alguns.

Por exemplo, no estudo (PEPPÉ et al., 2007), para as tarefas receptivas da prova de Afeto – em que era preciso julgar se o estímulo correspondia a uma produção de alguém que gostou (*like*) ou não gostou (*dislike*) de algo – o grupo de crianças com autismo cometeu mais erros do que o grupo de crianças controle. Na tarefa de expressão da mesma prova, as respostas preferenciais tendiam a ser “não gostou” e também julgavam “não gostou” como “gostou”. A expressão de afeto foi bastante inescrutável no grupo com autismo, visto que os julgamentos ambíguos (“não gostou” como “gostou” e vice-versa) foram significativamente mais frequentes neste grupo. Os autores observaram que as crianças estavam inclinadas a julgar os estímulos da tarefa receptiva de acordo com suas próprias preferências, apesar dos lembretes de que deveriam considerar as preferências do estímulo.

De um modo geral, o estudo (PEPPÉ et al., 2007) demonstra que as habilidades prosódicas receptivas e expressivas estão intimamente associadas no grupo autista.

Contudo, os autores sugerem que as habilidades prosódicas receptivas seriam um foco apropriado para a intervenção clínica. Ainda, afirmam que os indícios de que as habilidades prosódicas receptivas se desenvolvem tardiamente em crianças com autismo sugerem que a consciência prosódica pode ser mais explorada durante a intervenção. Uma das limitações do estudo, apontada pelos próprios autores, foi a não realização de análise acústica das tarefas expressivas do nível funcional do teste.

Filipe (2014) realizou análise acústica das tarefas ao aplicar o teste PEPS-C em crianças autistas falantes do português europeu. A autora mostrou que, embora as crianças com autismo pudessem produzir padrões prosódicos categoricamente precisos, seus contornos prosódicos eram percebidos como estranhos pelos ouvintes adultos, sendo que as medidas acústicas dos enunciados apresentavam alterações na duração e no *pitch*. Ainda, a autora pôde ver que as crianças autistas de alto funcionamento têm dificuldade para perceber e imitar padrões prosódicos, bem como dificuldades para compreender e produzir algumas das funções comunicativas transmitidas pela prosódia. Ao final a autora afirma que os resultados sugerem uma correlação positiva significativa entre a prosódia e outros domínios da linguagem e que, embora as deficiências prosódicas evidentes no autismo possam apresentar características semelhantes entre as línguas, existe alguma especificidade da linguagem que exige mais investigação.

## 5 Avaliação clínica da prosódia

Para que se amplie o conhecimento sobre o desenvolvimento típico da aquisição da prosódia, bem como para que se caracterize diferentes grupos clínicos, são necessários os instrumentos de avaliação dos aspectos prosódicos.

Apesar da importância de instrumentos que avaliem a prosódia, o teste mais completo atualmente disponível para avaliar as habilidades prosódicas de recepção e expressão é o *Profiling Elements of Prosodic Speech-Communication* – doravante PEPS-C – (PEPPÉ; MCCANN, 2003), que foi desenvolvido no Reino Unido para os falantes nativos de inglês com mais de 4 anos de idade. As tarefas do PEPS-C são realizadas nos níveis formal e funcional. O nível formal avalia a capacidade de recepção e processamento, bem como a capacidade de imitação de estímulos relacionados com o nível mais básico do processamento fonético. Nesses estímulos, não há envolvimento do significado, já que o sinal, filtrado, contém apenas a informação sonora emitida pela laringe, de forma que reste apenas a curva de  $F_0$ . Por sua vez, o nível funcional avalia a capacidade para compreender e expressar a prosódia utilizada para veicular as funções afetiva, pragmática, gramatical e interativa (FILIPE; FROTA; VICENTE, 2012).

O PEPS-C possui um total de nove provas. A primeira prova serve para averiguar se o indivíduo possui o vocabulário necessário para a realização do teste.

As oito provas restantes são constituídas por dois tipos de tarefas: as receptivas e as expressivas. Nas provas receptivas que avaliam o nível formal, são testados itens curtos (palavras simples) e itens longos (frases), mas utilizando apenas estímulos laríngeos, conforme explicado anteriormente. É solicitado ao sujeito que diga se os estímulos são iguais ou diferentes. Nas provas expressivas, é solicitado ao sujeito que imite o estímulo apresentado e a produção do sujeito é gravada.

Para a avaliação do nível funcional há, da mesma forma, tarefas receptivas e expressivas. Em versão revisada no ano de 2015, o teste passou a conter seis provas nesse nível. Cada prova tem o objetivo de avaliar uma capacidade específica, a saber: (a) Interação – capacidade de distinguir frases declarativas e interrogativas; (b) Afeto – capacidade de distinguir se gostou ou não gostou de algo apresentado; (c) Segmentação – capacidade de utilizar a prosódia para delimitar o discurso em unidades linguísticas (*chunks*); (d) Foco contrastivo – capacidade de distinguir o elemento importante na frase, por meio da realização de proeminência; (e) Acento lexical – capacidade de indicar a colocação de acento primário em palavras com duas sílabas; (f) Acento frasal – capacidade de diferenciar, de acordo com a realização do acento, entre dois substantivos simples ou um substantivo composto. Vale ressaltar que essa última prova foi modificada para a versão do português brasileiro, a partir da implementação de uma tarefa de identificação de foco informacional.

Uma bateria de testes de avaliação de prosódia dessa natureza auxilia sobremaneira clínicos e professores que lidam diariamente com a população que apresenta suspeita de perturbação na prosódia, como é o caso, por exemplo dos sujeitos com apraxia de fala, que é um transtorno relacionado à programação do gesto articulatório para a produção da fala. Estudos identificaram alterações prosódicas como uma característica importante desse transtorno, porém, não conseguiram descrever quais habilidades prosódicas estão alteradas. Por isso, os autores sugeriram estudos que utilizassem métodos de análises mais detalhados (ver SOUZA; PAYÃO, 2008; PAYÃO et al., 2012).

No Brasil, o teste PEPS-C está em fase final de adaptação por Lima, Barbosa e Celeste. Entende-se que a adaptação para o português brasileiro pode alavancar as pesquisas sobre prosódia no país e enriquecer as áreas de linguística, fonoaudiologia e psicologia, favorecendo a interdisciplinaridade, uma vez que o tema é de interesse comum.

## Considerações finais

O estudo e análise das características prosódicas da fala vai além dos ajustes laríngeos e supralaríngeos, de modo que o fonoaudiólogo precisa direcionar seu foco para as características que moldam o conteúdo que está sendo dito.

Com relação à atuação em fonoestilística, é sabido que a grande maioria dos ouvintes consegue diferenciar diversos estilos de elocução, mesmo quando não há conteúdo semântico e lexical presente. Por isso, estudos que investiguem a descrição de parâmetros prosódicos decorrentes da análise de seus componentes clássicos (duração, frequência fundamental e intensidade) tornam-se importante para o conhecimento de quais pistas acústicas levam os ouvintes a identificar corretamente um estilo de elocução.

Ao investigarmos os estudos sobre as características prosódicas no transtorno de linguagem, notamos o quão é importante a descrição dos mesmos componentes clássicos, porém, aplicados às habilidades de expressão e recepção da linguagem, bem como aos aspectos formais, que incluem as características perceptivo-auditivas, e os funcionais, que dizem respeito aos significados linguístico-pragmático. Nesse sentido, um teste estruturado que avalie a competência prosódica em diferentes grupos clínicos facilitaria a rotina de diagnóstico e reabilitação fonoaudiológica, pois permitiria ao profissional identificar as habilidades prosódicas mais prejudicadas.

## Referências

- AITKEN K, Papoudi D, Robarts J, Trevarthen C (1998). Children with autism, diagnosis and interventions to meet their needs. London: Jessica Kingsley.
- BARBOSA PA (1999). Revelar a Estrutura Rítmica de uma língua construindo máquinas falantes: pela integração entre ciência e tecnologia de fala. In: SCARPA, EM (org). Estudos de Prosódia. Campinas: Ed. da Unicamp.
- BARBOSA PA (2006). Incursões em torno do ritmo da fala. Campinas: Editora Pontes/FAPESP.
- BARBOSA PA, Madureira S. (2015). Manual de fonética acústica experimental. Aplicações a dados do português. Cortez Editora.
- BARON-Cohen S, Staunton R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language*, 14, 241–248.
- BORREGO MCM (2005). Expressividade no Rádio. In: Kyrillos, L. (org) Expressividade - da teoria à prática. Rio de Janeiro: Revinter.

- CAMARGO Z, Madureira S (2008). Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In: Proceedings of the Fourth Conference on Speech Prosody, Campinas, SP, Brazil, 1: 57–60.
- CAMARGO Z, Madureira S; Schmitz, JR (2013). Qualidade vocal e produções de fala em três línguas: um estudo de caso. *Revista Intercâmbio*, v. XXVII: 110-140. São Paulo: LAEL/PUCSP. ISSN 2237-759x.
- CAMPOS LCP (2012). Radialista: análise acústica da variação entoacional na fala profissional e na fala coloquial. {Dissertação de Mestrado}: Unicamp, Campinas.
- CASTRO L, Serridge B, Moraes J, Freitas M (2010). The prosody of the TV news speaking style in Brazilian Portuguese Federal University of Rio de Janeiro, Brazil. Proceedings of the third ISCA Tutorial and Research Workshop on Experimental Linguistics ExLing.
- CASTRO L (2008). O comportamento dos parâmetros duração e frequência fundamental nos fonestilos político, sermonário e telejornalístico [PhD thesis]: UFRJ, Rio de Janeiro.
- CONSTANTINI AC (2012). Mudanças na estruturação prosódica de texto jornalístico antes e após intervenção fonoaudiológica. *Journal of Speech Sciences* 2(2):2342.
- DELLWO V (2010). The automatic extraction of time-domain based speaker idiosyncratic features. Abstracts IAFPA.
- FILIFE MG (2014). Prosodic abilities in typically developing children and those diagnosed with autism spectrum disorders: clinical implications for assessment and intervention. [PhD Thesis]: Universidade do Porto. Porto. 292p.
- FILIFE MG, Vicente S, Frota S (2012). Aquisição Prosódica entre 4 e 18 anos de idade. *Atas de Psicologia & Educação*. 15p.
- GERVAIN J, Mehler J (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, v. 61, p. 191-218.

- HARGROVE P, Sheran C (1989). The use of stress by language-impaired children. *Journal of Communication Disorders*, 22, 361–373.
- IRVINE, JT (2001). Style as distinctiveness: the culture and ideology of linguistic differentiation. In: Eckert, Penelope; Rickford, John. R. *Style and sociolinguistic variation*. Cambridge University. p. 2143.
- KYRILLOS L (2004). A voz do profissional de telejornalismo. In: Ferreira LP, Oliveira SMRP. *Voz Profissional: produção científica da Fonoaudiologia Brasileira*. São Paulo: Editora Roca; p. 75- 83.
- LAVER J (1980). *The phonetic evaluation of voice quality*. Cambridge: Cambridge University Press.
- LEEMANN A, Kolly MJ, Dellwo V (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic Science International*. 238: 59-67
- LOPES LW, Lima ILB (2014) Prosódia e transtornos da linguagem: levantamento das publicações em periódicos indexados entre 1979 e 2009. *Revista Cefac*. 16(2): 651-9.
- MADUREIRA S (2005). Expressividade na fala. In: Kyrillos L. (org) *Expressividade: da teoria à prática*. Rio de Janeiro: Revinter.
- MCCANN J, Peppé S (2003). Prosody in autism spectrum disorders: a critical review. *INT. J. LANG. COMM. DIS.*, 38(4):325–350.
- NAZZI T, Bertoncini J, Mehler J (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24(3): 756-66.
- PAYÃO LMC, Lavra-Pinto B, Wolff CL, Carvalho Q (2012). Características clínicas da apraxia de fala na infância: revisão de literatura. *Letras de Hoje*, 47 (1):24-29.
- PEPPÉ S, McCann J (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, 17(4/5):345-354.

- PEPPÉ S, McCann J, Gibbon F, O'Hare A, Rutherford M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech Language and Hearing Research*, 50, 1015-1028.
- SAN SEGUNDO E, Foulkes, P, French P, Harrison P, Hughes V (2016). Voice quality analysis in forensic voice comparison: developing the vocal profile analysis scheme. Conference: The 25th annual conference of the International Association for Forensic Phonetics and Acoustics, At York, UK.
- SOUZA TNU, Payão LMC (2008). Apraxia da fala adquirida e desenvolvimental: semelhanças e diferenças. *Rev Soc Bras Fonoaudiol*, 13(2):193-202.
- WELLS B, Peppé S, Goulandris A (2004). Intonation development from five to thirteen. *Journal of Child Language*. 31:749-78.
- WIGET L, White L, Schuppler B, Grenon I, Rauch O, Mattys SL (2010). How stable are acoustic metrics of contrastive speech rhythm. *J. Acoust. Soc. Am.* (127, 3): 1559-1569.
- WING L, Gould J (1979). Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *J Autism Dev Disord*. 9(1):11-29.
- YOON TJ (2010). Capturing inter-speaker invariance using statistical measures of speech rhythm. *Electronic proceedings of Speech Prosody*, Chicago/IL, USA.

# Contribuições sobre as características prosódicas de interrogativas totais neutras produzidas por sergipanos

Vanessa Gonzaga Nunes  
Universidade Federal de Sergipe

## 1 Introdução

Desde 2013, contemplamos em nossos estudos prosódicos as variedades do estado de Sergipe. Este estudo agrega dados que ainda não haviam sido manipulados em estudos anteriores, como em Nunes (2015). Para esta pesquisa específica, exploramos as interrogativas totais neutras, ou seja, aquelas que aceitam resposta sim/não, nas variedades dialetais de Aracaju, Estância, Itabaiana, Lagarto e Aquidabã.

Para tal pesquisa, utilizamos o corpus AMPER-POR (Atlas Multimídia Prosódico do Espaço Românico – Língua Portuguesa), grupo do qual fazemos parte. O projeto já reúne inúmeras equipes de trabalho em vários lugares do mundo, o que compreende pesquisadores em vários cantos do Brasil. Na região Nordeste, seguindo essa metodologia, além de Sergipe, já se há resultados sobre padrões prosódicos de declarativas e interrogativas totais nos estados do Maranhão, Ceará, Paraíba, Pernambuco e Bahia.

Apresentaremos aqui as curvas de  $F_0$  das interrogativas produzidas pelos locutores das cinco cidades sergipanas em questão, realizando uma comparação entre elas. Focaremos na região nuclear, ou seja, no movimento final das sentenças, onde ocorre a marca mais robusta de discriminação de interrogativas e declarativas. Olharemos, considerando uma parte dos dados, o detalhe fonético do movimento intrassilábico através da taxa de inclinação do movimento de  $F_0$  nas sílabas finais e a tessitura (intervalo entre o menor e o maior valor de  $F_0$  observados em um enunciado). Apresentamos, ainda, a taxa de elocução de algumas das variedades.

São perguntas e hipóteses desta pesquisa:

Q1. Os desenhos das curvas de  $F_0$  encontrados para interrogativas totais neutras irão estar mais próximos ao padrão circunflexo encontrado para sentenças correlatas estudadas por pesquisadores que investigam outras variedades regionais (MORAES, 1998; SILVA, 2011; REIS, 2011, etc.) ou mais de acordo com a descrição ascendente já encontrada para regiões do nordeste brasileiro descrita por pesquisas como a de Lira (2009)?

H1. Acredita-se que, de maneira geral, o desenho do contorno melódico da região nuclear das sentenças será, sobretudo, ascendente, como já demonstraram dados estudados anteriormente (NUNES, 2015).

H2. O movimento de subida acontecerá na região nuclear independentemente do tipo acentual da palavra final da frase. Estudos que investigam variedades outras demonstram o padrão circunflexo para frases terminadas por proparoxítonas e paroxítonas e, que apenas as frases que terminam por oxítonas podem apresentar movimento apenas ascendente, como é o caso de Milan (2015), que se dedicou ao falar de Curitiba ou de Silva (2011), em relação aos dados de Florianópolis. Nossa hipótese é de que os dados de todos os informantes de Sergipe apresentem muito frequentemente, para a região nuclear, movimento ascendente sendo a palavra final proparoxítona, paroxítona ou oxítona.

H3. A ascendência da curva, contudo, estará atrelada à região de acento, onde se inicia o movimento de subida.

Q2. De acordo com Callou e Leite (1990), a tessitura (intervalo entre o menor e o maior valor de  $F_0$  observados em um enunciado) é maior para mulheres do que para homens. Os dados das mulheres sergipanas apresentarão valores de tessitura maiores do que os valores encontrados para os homens ?

H1. É possível que os dados revelem valores de tessitura maiores para as mulheres sergipanas.

Q3. A taxa de inclinação do movimento de  $F_0$  das vogais da região nuclear pode confirmar que o movimento final das sentenças é ascendente e que o maior grau de inclinação se dá no movimento intrassilábico da tônica?

H1. Acredita-se que esse parâmetro, embora pouco testado em pesquisas do gênero, pode servir para demonstrar se o movimento é preponderantemente ascendente ou mais à direita (positivo) ou preponderantemente descendente ou mais à esquerda (negativo) e que pode identificar qual a região de maior variação de  $F_0$  e, conseqüentemente, pesquisas futuras podem revelar que se trata de um parâmetro importante para discriminação de variedades dialetais.

Q4. Os valores encontrados para a taxa de elocução dos dados analisados estarão em conformidade com média encontrada para outros estudos da área ou poderão classificar a velocidade de fala do sergipano como lenta, que é uma descrição estereotipada sobre os falares do nordeste?

H1. Muito provavelmente alguns dados não confirmarão a ideia pré-estabelecida sobre a velocidade de fala dos nordestinos.

## **2 Elementos prosódicos de interrogativas totais neutras em variedades brasileiras**

A modalidade interrogativa total é aquela que aceita resposta sim ou não. No entanto, na prática, nem sempre é fácil discriminar ou selecionar uma interrogativa neutra porque as questões que aceitam resposta sim ou não também podem desempenhar papéis modais (relacionados com a expressão de emoções e atitudes) com base o contexto pragmático em que operam.

Em relação à curva  $F_0$  de interrogativas totais neutras, em geral, apresentam, no início da frase, um movimento de subida na primeira tônica seguido por um movimento descendente até a pretônica final, quando a curva volta a subir na tônica para novamente cair nas sílabas postônicas. A descrição clássica para interrogativas totais neutras é de Moraes (2008). O pesquisador, que examinou a variedade do Rio de Janeiro, encontrou para esta modalidade, uma curva que pode ser descrita por um movimento ascendente na primeira sílaba acentuada da sentença; subida esta um pouco maior do que a observada em sentenças declarativas correlatas. No entanto, a principal diferença entre uma declarativa neutra e uma interrogativa neutra está na região nuclear ou parte final da frase. As interrogativas têm, então, um movimento ascendente-descendente característico capaz de discriminá-las das declarativas. Muitos pesquisadores, a exemplo de Truckenbrodt et al. (2008), encontraram esse mesmo padrão para outras variedades e, a partir de tais estudos, pode-se dizer que este desenho ascendente-descendente é o padrão das curvas melódicas de interrogativas totais neutras do Português Brasileiro (PB). No entanto, outros pesquisadores que confirmaram os contornos circunflexos, também encontraram movimentos apenas ascendentes para a região final de interrogativas totais neutras, como Lira (2009) que investigou cinco variedades do Nordeste do Brasil; Silva (2011) que explorou várias capitais brasileiras; Reis et al. (2011) que estudaram as variedades de Mariana e Belo Horizonte, no estado de Minas Gerais; Lemos e Cruz (2013) e Costa e Cruz (2014), que olharam as variedades de Baião e Mocajuba no Estado do Pará; Paixão (2014), que investigou a variedade de Rio de Janeiro e Wildner (2013) e Milan (2015) que estudaram, respectivamente, as variedades de Ponta Grossa e Curitiba, ambas cidades do Estado do Paraná. No entanto, boa parte destes estudos encontrou curvas ascendentes apenas em frases que terminam por palavras oxítonas, ou seja, os movimentos de subida das interrogativas estariam, nessas variedades dialetais, atreladas à sílaba acentuada. Em outros casos, o desenho da curva que sobe e que

não desce ocorre quase que sistematicamente por causa desvozeamento das sílabas finais. Neste caso, a perda ou a ausência de material fonético impede a curva de descer. Uma curva afetada pelo desvozeamento de uma ou mais vogais não tem a mesma estrutura de uma curva eminentemente ascendente.

No campo fonológico, Moraes (2008) propõe para curvas melódicas de interrogativas totais neutras a notação  $/L+\langle H *L\% /$ . Outros estudos encontraram o mesmo comportamento e decidiram por notações semelhantes, como Milan (2015). No entanto, a pesquisadora atribuiu a notação  $L+L*H\%$  para frases que terminam por palavras oxítonas. Wildner (2013) escolheu descrever as frases que terminam por palavras oxítonas como  $L+H*(L\%)$ , entendendo que, no nível subjacente, existem vogais, mas a máquina não foi capaz de registrá-las.

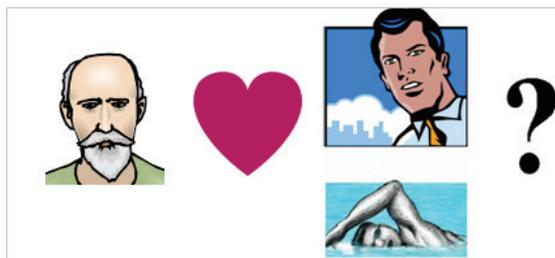
A duração também revelou sua importância para os estudos prosódicos. Ela manifesta-se, por exemplo, na velocidade de fala. Este parâmetro fortemente relacionado à caracterização das atitudes, estilo, gênero etc., pode também ser considerado nos estudos dialetais (LAVER, 1994). No PB, não temos muitos estudos que calcularam a velocidade de fala de variedades regionais e, infelizmente, as pesquisas que se dedicaram ao parâmetro utilizam metodologias distintas, o que nos impede de comparar os resultados.

Meireles e Barbosa (2009), Meireles et al. (2010), Meireles e Gambarini (2011) analisaram as variedades regionais brasileiras consideradas como lentas e rápidas, usando o mesmo procedimento e demonstraram que a duração parece ser um parâmetro importante para os estudos prosódicos. Eles encontraram para falantes de Minas Gerais, Bahia e Espírito Santo, respectivamente, 6,3, 4,2 e 4 sílabas por segundo, como velocidade de fala.

## 3 Metodologia

### 3.1 O corpus

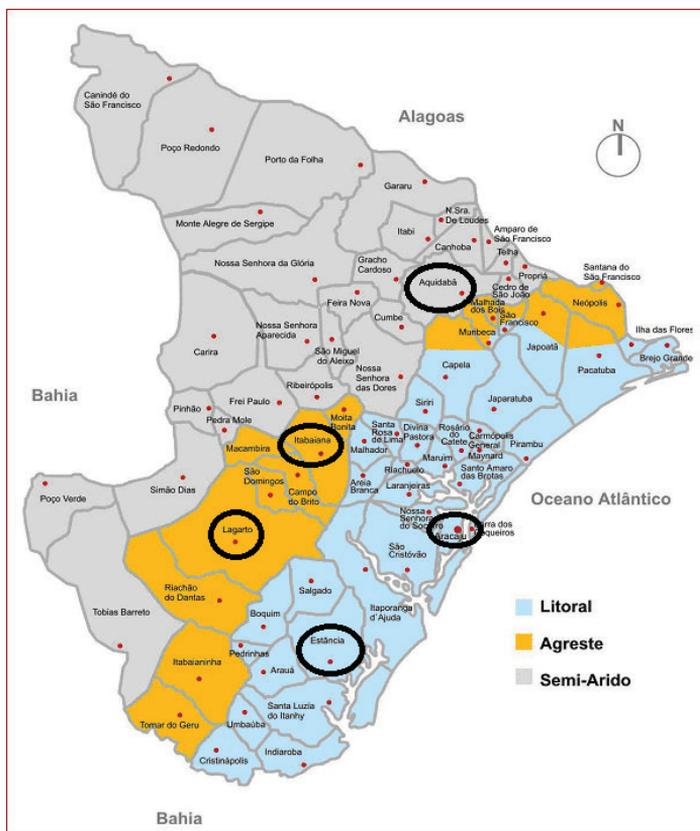
Este estudo faz parte do projeto AMPER-POR que tem um *corpus* próprio e que proporciona a comparação de variedades, sobretudo no que concerne às curvas de  $F_0$ . O corpus completo contempla 33 frases declarativas neutras e 33 frases interrogativas neutras que são produzidas três vezes - em diferentes tempos de coleta de dados - a partir de imagens. As sentenças têm estruturas: sujeito + verbo + complemento, podendo haver extensões. O sintagma nominal (SN), ou seja, a região pré-nuclear (antes do verbo) e a frase verbal (SV), ou seja, a região nuclear (depois do verbo), têm todos os tipos de acentos lexicais do PB. As palavras que compõem as sentenças podem ser, então, proparoxítonas (ex.: pássaro), paroxítonas (ex.: Renato) ou oxítonas (ex.: bisavô).



**Figura 1** Modelo de estímulo visual para a produção de frases. Modelo da frase interrogativa total composta por SN sujeito + verbo + SV (objeto paroxítono + complemento adjetivo oxítono): *o bisavô gosta do Renato nadador?*

### 3.2 Locais de pesquisa e locutores

As cidades aqui contempladas são Aracaju, Estância, Lagarto, Itabaiana e Aquidabã, todas elas pertencentes ao estado de Sergipe, conforme mostra a ilustração a seguir.



**Figura 2** Cidades contempladas nesta pesquisa. Em sentido horário: (1) Aquidabã, (2) Aracaju, (3) Estância, (4) Lagarto e (5) Itabaiana.

O corpus contempla frases de 10, 13 e 14 sílabas. Depois de uma seleção das três melhores repetições, tivemos 198 sentenças (99 declarativas e 99 interrogativas) por locutor, 396 sentenças para cada cidade, totalizando 1980 (990 sentenças declarativas e 990 frases interrogativas) produzidas pelos nove locutores que participaram desta pesquisa.

### 3.3 Os parâmetros de análise

Realizamos as rotinas do protocolo AMPER-POR. Portanto, dispomos de gráficos e histogramas criados a partir das medidas de  $F_0$ , duração e intensidade de cada vogal. Os scripts também proporcionaram dados importantes sobre os valores de  $F_0$ , duração e intensidade, medidos em três pontos de cada vogal (F01, F02, F03).

#### 3.3.1 Normalização de dados

Todos os dados correspondentes aos valores de  $F_0$ , duração e intensidade obtidos através da interface AMPER foram exportados para planilhas, onde foram organizados de acordo com o número de vogais de cada frase e do tipo de acento lexical das palavras. Os valores de  $F_0$  em Hertz permitiram verificar a frequência máxima, mínima e média de cada locutor. Para eliminar valores discrepantes e para que pudéssemos comparar produções de diferentes falantes, os valores de frequência foram normalizados em semitons (st). Para isso, foi utilizada a seguinte fórmula:  $n=12x\log_2 (fn/m)$ , onde  $n$  é o valor normalizado,  $fn$  é o valor de  $F_0$  em Hertz e  $m$  é a média da frequência laríngea do informante.

#### 3.3.2 A tessitura

Como a terminologia prosódica pode gerar hesitações sobre o parâmetro que está sendo observado. Vamos definir os aspectos observados nos dados para esta pesquisa. Olhamos para: (i) a gama tonal ou a variação de amplitude de  $F_0$  no movimento de subida final e (ii) tessitura ou variação de  $F_0$  entre o menor e o maior valor de um enunciado e nesta pesquisa apresentaremos tais valores. Para essas análises, foram utilizados os valores fornecidos pelo script AMPER. Utilizamos, mais especificamente, os valores F01, F02, F03 da vogal acentuada na região final da frase, onde ocorre o movimento de subida. Para as frases de comportamento final ascendente, é possível que o movimento de subida corresponda ao parâmetro de tessitura, pois podemos ter, nesta região, o ponto mais baixo e o mais alto da curva melódica de toda a sentença. Entretanto, para frases de configuração final circunflexa, é possível que o terceiro ponto da vogal (F03) já esteja em mo-

vimento de descida e, nesse caso, ele não poderia ser levado em consideração no cálculo do movimento de subida. Por isso, para algumas variedades ou dados que apresentavam o padrão ascendente-descendente, escolhemos calcular apenas os dois primeiros pontos, evitando, assim, mascarar a variação de  $F_0$ .

### 3.3.3 Taxa de inclinação

Tendo como base as análises de Colamarco (2009), nós decidimos calcular a inclinação da vogais tônicas que compõem a região nuclear das frases e, para isso, nós utilizamos a fórmula *Tx de inclinação* =  $F01 - F03/dur \times 1000$ , onde F01 e F03 são o primeiro e o último pontos da vogal, respectivamente; *dur* se refere à duração da vogal. Os valores de  $F_0$  estão em semitons (st) e a duração em milissegundos (ms). As curvas ascendentes são consideradas positivas e as descendentes negativas. Esse parâmetro demonstra se o movimento é preponderantemente ascendente ou mais à direita (positivo) ou preponderantemente descendente ou mais à esquerda (negativo).

### 3.3.4 Taxa de elocução

A partir dos scripts Amper, podemos obter o valor das durações das vogais, entretanto, observamos que a literatura da área trabalha com unidades que compreendem a articulação da vogal com o segmento vizinho. Fundamentados por Barbosa (2000), para este trabalho, a taxa de elocução das modalidades declarativas e interrogativas e dos informantes de cada ponto de inquérito foi calculada através da relação V2V (do início de uma vogal até o começo da vogal seguinte) por segundo. Utilizamos o *Descriptive statistics* (SPSS) para obter os valores mínimos, máximos e médias das unidades V2V, por locutor e por modalidade. Para esta pesquisa apresentaremos a taxa de elocução apenas de uma parcela dos informantes.

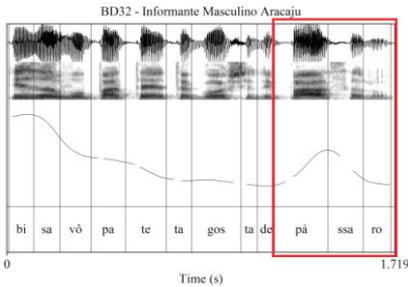
## 4 Os resultados

### 4.1 O padrão e a subida da curva de $F_0$ na região final da frase

As regiões nucleares das interrogativas totais neutras produzidas por falantes de Sergipe foram analisadas separadamente por ponto de enquete e pelo tipo acentual da última palavra da frase (proparoxítona, paroxítona ou oxítona).

No que diz respeito às interrogativas totais que terminam por palavras proparoxítonas, nós encontramos dois comportamentos : (1) de maneira menos frequen-

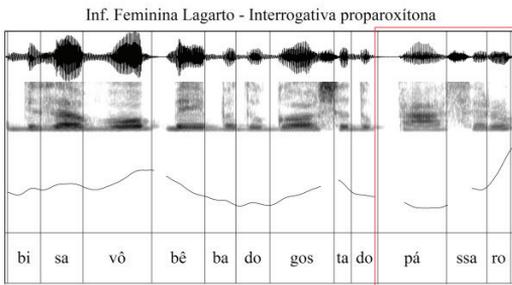
te, um movimento ascendente-descendente (figuras 3a e 3b) e, mais recorrente, um movimento ascendente (figuras 3c e 3d). As curvas de  $F_0$  de desenho circunflexo podem apresentar movimento ascendente até o final da tônica (Figura 3a) e subsequente queda ou uma subida que pode alcançar a parte medial postônica, para posterior queda (Figura 3b). As curvas de  $F_0$  de desenho ascendentes apresentam movimento de subida na parte final da tônica ou na postônica 1, subida esta que continua até o final da postônicas 2. Alguns movimentos de subida são mais ou menos íngremes nas postônicas, como podemos observar nas figuras 3c e 3d.



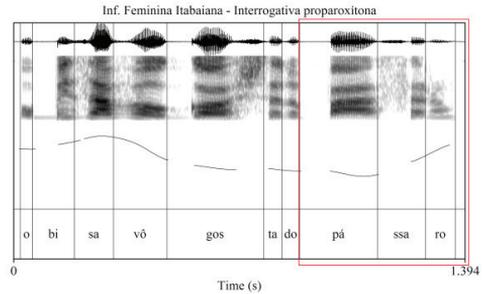
(a)

FALTOU IMAGEM

(b)



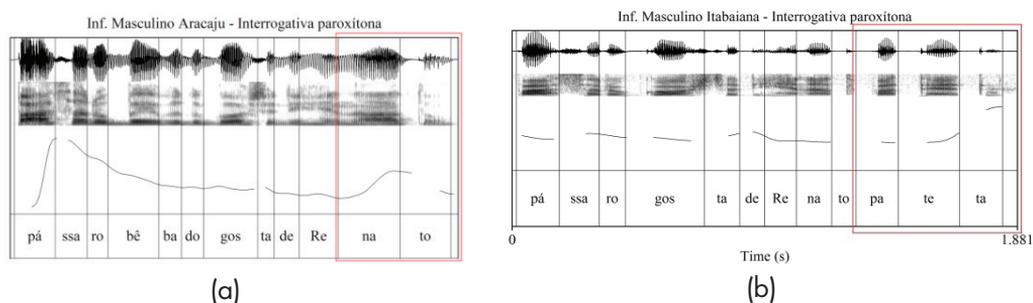
(c)



(d)

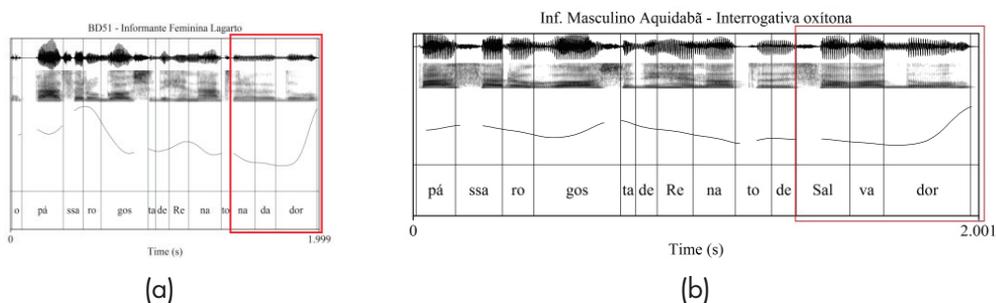
**Figura 3** (a) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o bisavô gosta do pássaro?”, pelo informante masculino de Aracaju. (b) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o pássaro gosta do Renato de Mônaco?”, pela informante feminina de Estância. (c) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total o “bisavô gosta do pássaro?”, pela informante feminina de Lagarto. (d) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total o “bisavô gosta do pássaro?”, pela informante feminina de Itabaiana.

Para frases que terminam por paroxítonas também encontramos os dois padrões, sendo o comportamento ascendente mais frequente do que o comportamento ascendente-descendente. No padrão circunflexo, o movimento de subida ocorre na tônica (Figura 4a), com subsequente queda na postônica e, o desenho mais frequente para o padrão ascendente também realiza seu movimento de subida na tônica com pico de  $F_0$  na postônica.



**Figura 4** (a) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o pássaro bêbado gosta do Renato?”, pelo informante masculino de Aracaju. (b) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o pássaro gosta do Renato pateta?”, pelo informante masculino de Itabaiana.

As curvas de  $F_0$  de sentenças interrogativas totais que terminam por oxítonas apresentam apenas comportamento ascendente. O movimento de subida ocorre na tônica e os picos de  $F_0$  podem ser, com maior frequência, íngremes até o final do enunciado (Figura 5a) ou mais curvilíneos e menos íngremes (Figura 5b).



**Figura 5** (a) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o pássaro gosta do Renato nadador?”, pela informante feminina de Lagarto. (b) Forma de onda, espectrograma e curva de  $F_0$  da sentença interrogativa total “o pássaro gosta do Renato de Salvador?”, pelo informante masculino de Aquidabã.

## 4.2. A tessitura

Nós apresentamos os valores médios para a tessitura (intervalo entre o menor e o maior valor de  $F_0$  observados em um enunciado) de interrogativas totais neutras. Os locutores de Sergipe obtiveram média de 11,68 St. Os valores médios encontrados a partir de cada sentença podem ser verificados na Tabela 1.

**Tabela 1** Valores médios de tessitura das sentenças interrogativas totais produzidas pelos sergipanos.

<b>F<sub>0</sub></b>	<b>Região da sentença</b>	<b>Locutores</b>	<b>Mín</b>	<b>Máx.</b>	<b>Tessitura</b>
Valores em Semitons (st)	SN + SV	Aracaju Fem.	6,20	22,11	14,32
		Aracaju Masc.	6,53	15,57	10,67
		Estância Fem.	6,05	16,56	11,87
		Estância Masc.	9,68	16,52	13,36
		Lagarto Fem.	6,67	17,41	10,82
		Lagarto Masc.	7,30	19,47	13,02
		Itabaiana Fem.	9,95	19,49	14,70
		Itabaiana Masc.	6,16	11,58	8,9
		Aquidabã Masc.	1,93	11,06	7,5
<b>Média</b>					11,68
<b>Número de dados</b>					<b>990</b>

### 4.3 A taxa de inclinação

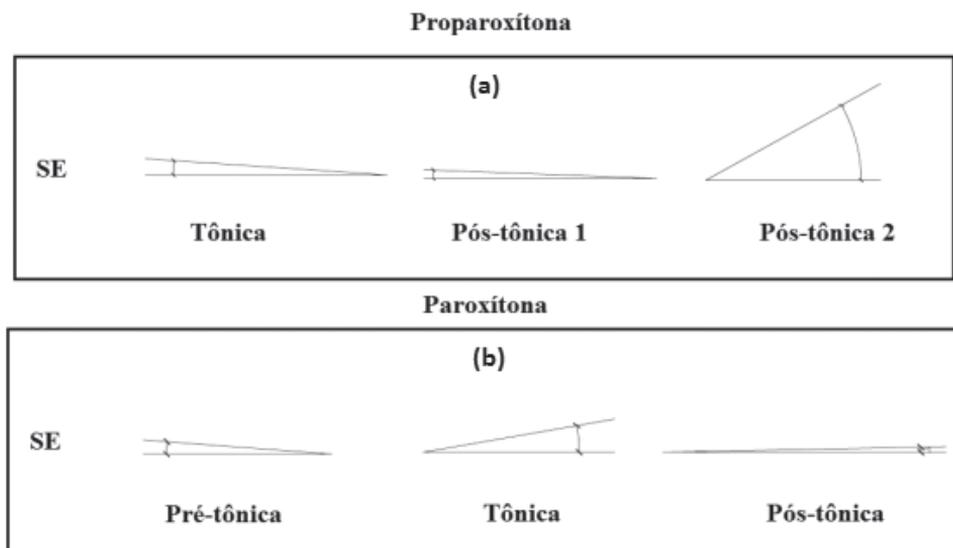
Apresentamos aqui os valores encontrados para a taxa de inclinação dos dados de Aracaju, Estância e Lagarto. Em virtude da quantidade de dados, apresentamos um recorte das análises que contemplaram 366 dados referentes às sentenças de interrogativas totais de 13 sílabas.

**Tabela 2** Valores da taxa de inclinação de F<sub>0</sub> (st/s) das vogais tônicas de proparoxítonas, paroxítonas e oxítonas em região nuclear das sentenças interrogativas totais com 13 sílabas produzidas por sergipanos.

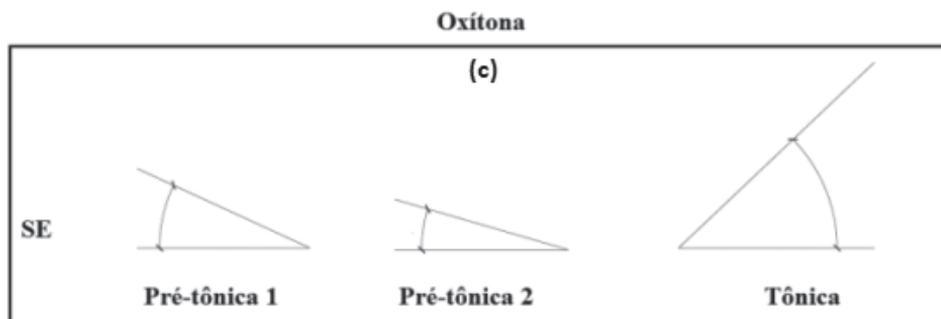
<b>F<sub>0</sub></b>	<b>Tipo Acentual no núcleo</b>	<b>Locutores</b>	<b>Tônica Proparoxítonas</b>	<b>Tônica Paroxítonas</b>	<b>Tônica Oxítonas</b>
Valores em Semitons (st)	Proparoxítona	Fem. Aracaju	-8,07	-5,52	64,42
		Masc. Aracaju	24,51	20,33	22,95
		Fem. Estância	15,68	16,79	16,70
		Masc. Estância	-1,83	0,63	36,54
		Fem. Lagarto	-5,57	3,54	37,82
		Masc. Lagarto	-3,39	1,21	61,14
<b>Número de dados</b>		<b>366</b>	<b>160</b>	<b>144</b>	<b>62</b>

As configurações das inclinações das vogais dos sergipanos são distintas entre si. As tônicas das palavras proparoxítonas finais apresentam três comportamentos: (1) movimento mais ascendente na tônica (valores positivos) com queda nas postônicas (valores negativos), que é o caso da informante de Estância e do informante de Aracaju; (2) movimento mais descendente na tônica (valores negativos) e movimento de subida que se inicia na postônica 1 e que continua subindo na postônica 2, que é o caso da informante de Aracaju e do informante de Lagarto e (3) movimento de subida que inicia apenas na última sílaba do enunciado, que é o caso do informante de Estância e da informante de Lagarto (Ver inclinação de  $F_0$  da tônica em ás postônicas na Figura 6a). As vogais pretônicas apresentam, sobretudo, valores negativos, o que identifica o movimento descendente. Já as tônicas apresentam valor positivo, indicando que o movimento nesta posição é de subida e que ele continua na postônica (valores positivos). Outras configurações são possíveis, como, por exemplo, subida que começa na pretônica. (Ver inclinação de  $F_0$  da tônica em relação à pretônica e a postônica na Figura 6b) Para oxítonas, os valores para taxa de inclinação de  $F_0$  são sempre positivos nas vogais tônicas e, no geral, negativos nas pretônicas, indicando movimento descendente (Ver inclinação de  $F_0$  da tônica em relação às pretônicas na Figura 6c).

Na Figura 6, podemos ver a representação estilizada das taxas de inclinação das vogais pertencentes à palavras proparoxítonas, paroxítonas e oxítonas produzidas por sergipanos.



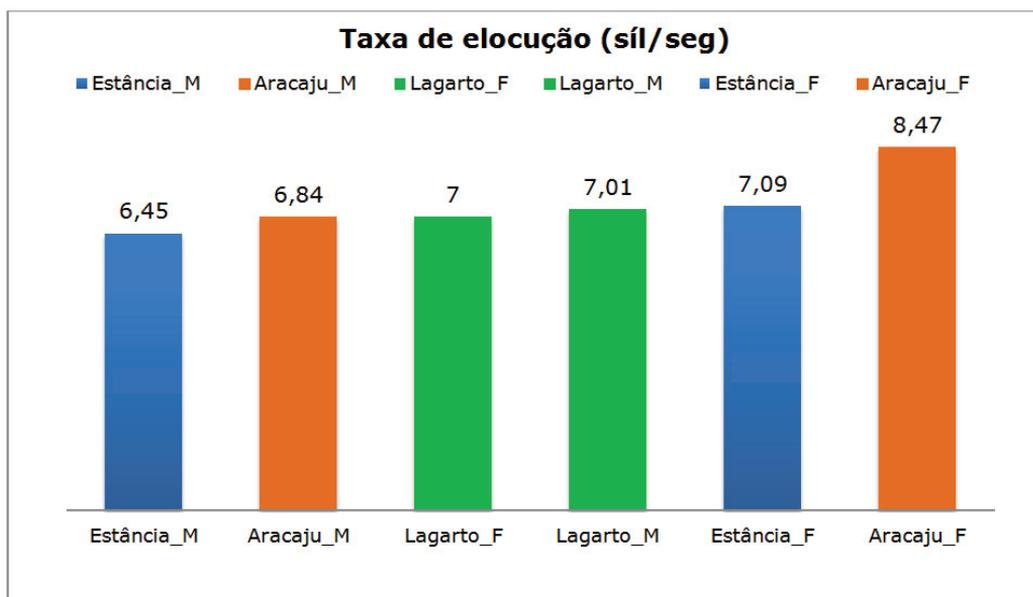
**Figura 6** Representação estilizada da inclinação média das vogais finais em região nuclear das sentenças interrogativas totais produzidas por sergipanos. (*continua*)



**Figura 6** Representação estilizada da inclinação média das vogais finais em região nuclear das sentenças interrogativas totais produzidas por sergipanos. (continuação)

#### 4.4 A taxa de elocução

A taxa de elocução (número de sílabas por segundo, calculada a partir de unidades V2V), não nos permitiu agrupar sergipanos por pontos de enquete. Apresentamos aqui os valores encontrados para os dados de Aracaju, Estância e Lagarto. Os valores encontrados para todos os locutores estão na Figura 7.



**Figura 7** Gráfico da taxa de elocução de locutores de Sergipe. Em azul a taxa de elocução de locutores de Estância, em laranja, de Aracaju e, em verde, de Lagarto.

A informante aracajuana apresentou a maior taxa de elocução (8,47 sil/s). Na sequência, das velocidades mais altas para as mais baixas, tivemos Estância

fem. (7,09 sil/s), Lagarto masc. (7,01 sil/s), Lagarto fem. e Aracaju masc. (ambas com 6,84 sil/s) e Estância masc. (6,45 sil/s). A taxa de elocução da locutora feminina de Aracaju foi 19,24% mais rápida do que a do locutor masculino de Aracaju e 23,84% mais rápida do que a do locutor masculino de Estância, que teve a menor taxa do grupo de sergipanos. Meireles e Barbosa (2009) encontraram como taxa de elocução para mineiros e paulistas 7,5 e 6,4 sílabas por segundo, respectivamente. Meireles e Gambarini (2011) também realizaram uma pesquisa com locutores mineiros de Belo Horizonte, baianos de Jacuípe e capixabas de Vitória e obtiveram como taxa de elocução 6,3, 4,2 e 4,0 sílabas por segundo, respectivamente. A partir desses dados, podemos inferir que a maioria dos locutores sergipanos não difere muito das taxas registradas pela literatura, que alguns locutores apresentaram taxas elevadas, como é o caso da mulher aracajuana (8,47 sil/s) e que se tomarmos como parâmetro às taxas de elocução consideradas lentas (4,2 síl/seg.), os dados dos nossos locutores vão de encontro com a ideia impressionista de que todos os nordestinos falam devagar.

## 5 Considerações finais

Agora vamos tentar responder às questões de pesquisa apresentadas no começo do trabalho e vamos verificar se as hipóteses se confirmam ou não.

Q1. Os desenhos das curvas de  $F_0$  encontrados para interrogativas totais neutras irão estar mais próximos ao padrão circunflexo encontrado para sentenças correlatas estudadas por pesquisadores que investigam outras variedades regionais (MORAES, 1998; SILVA, 2011; REIS, 2011 etc.) ou mais de acordo com a descrição ascendente já encontrada para regiões do nordeste brasileiro descrita por pesquisas como a de Lira (2009)?

H1. Encontramos dois padrões para o movimento da curva melódica no seu trecho final. Temos, então, o movimento circunflexo já descrito por muitos pesquisadores como sendo o mais frequente para interrogativas do gênero em PB, mas temos a maior recorrência de curvas ascendentes, o que corrobora com os dados de Lira (2009) para regiões do nordeste. Os novos dados, sobretudo referentes aos falantes de Itabaiana e Aquidabã, cidades até então não exploradas pela pesquisadora, apresentam maior recorrência de movimento ascendente para a região nuclear, corroborando com os dados apresentados em Nunes (2015) sobre os falares sergipanos.

H2. A hipótese se confirma uma vez que os movimentos de subida, seja para desenhos melódicos circunflexos, seja para desenhos melódicos ascendentes, ocorrem na região nuclear da frase. Além disso, temos movimento apenas ascendente independente do tipo acentual da palavra final, ou seja, temos padrão acendente para frases que terminam por palavras oxítonas, paroxítonas ou proparoxítonas.

H3. Para as frases que apresentam movimento ascendente-descendente, temos a subida na posição da tônica. Entretanto, nas frases que apresentam padrão ascendente, o movimento de subida não se dá necessariamente na região do acento. Em palavras proparoxítonas ou paroxítonas, por exemplo, é muito comum observarmos movimento ainda descendente ou muito plano na tônica, para só depois, na postônica, haver o movimento de subida. No geral, o alçamento se dá da penúltima sílaba para última, independente do tipo de acento da palavra final.

Q2. De acordo com Callou e Leite (1990), a tessitura (intervalo entre o menor e o maior valor de  $F_0$  observados em um enunciado) é maior para mulheres do que para homens. Os dados das mulheres sergipanas apresentarão valores de tessitura maiores do que os valores encontrados para os homens ?

H1. A hipótese não se confirmou para todos os dados. Os dados não foram agrupados por sexo. Do grupo dos nove informantes, o valor mais alto é o da mulher de Aracaju (14,32 st), seguido pelo homem de Estância (13,36 st) e pelo homem de Lagarto (13,02 st). Os valores mais baixos foram registrados nos dados do homem de Itabaiana (5,13 st), seguido da mulher da mesma cidade (9,53 st). A tessitura também não se mostrou um parâmetro eficaz para a discriminação de variedades, uma vez que não agrupou os valores por cidades.

Q3. A taxa de inclinação do movimento de  $F_0$  das vogais da região nuclear pode confirmar que o movimento final das sentenças é ascendente e que o maior grau de inclinação se dá no movimento intrassilábico da tônica?

H1. A hipótese se confirmou parcialmente. A taxa de inclinação de  $F_0$  demonstrou ser um parâmetro importante para a caracterização da prosódia dos sergipanos uma vez que ratifica as análises que indicam que temos, para a região nuclear, movimento apenas ascendente. Os dados de Itabaiana e de Aquidabã vão ao encontro dos resultados encontrados por Nunes (2015) para alguns dos informantes sergipanos. Entretanto, o maior grau de inclinação não necessariamente ocorre no movimento intrassilábico das tônicas. Considerando que as curvas melódicas mais frequentes apresentam movimento ascendente, o maior grau de inclinação pode ocorrer no trecho final da região nuclear, que é o caso da inclinação da postônica 2 de proparoxítonas.

Q4. Os valores encontrados para a taxa de elocução dos dados analisados estará em conformidade com média encontrada para outros estudos da área ou poderá classificar a velocidade de fala do sergipano como lenta, que é uma descrição estereotipada sobre os falares do nordeste?

H1. Os valores encontrados para a taxa de elocução dos sergipanos são superiores ou estão em paridade aos valores encontrados por pesquisadores que estudam outras variedades, o que significa dizer que os resultados vão de encontro com a visão impressionista de que os nordestinos falam mais lentamente do que falantes de outras regiões. Não é um parâmetro que tenha se mostrado eficiente

para a discriminação de variedades dialetais de Sergipe. Os dados de Itabaiana e de Aquidabã não foram analisados neste quesito, mas de oitiva, estima-se que os dados da mulher de Itabaiana também apresentará valores altos se considerarmos a classificação da literatura da área.

No geral, curvas de  $F_0$  de interrogativas totais neutras produzidas por sergipanos apresentam uma ataque alto (início da curva), poucas proeminências de  $F_0$  ao longo da curva e íngremes alçamentos de  $F_0$  na região nuclear, tanto para o padrão circunflexo, tanto para o padrão ascendente. Obviamente, estamos analisando locutores de um mesmo estado (e de um estado muito pequeno) e isso tem implicações sobre o agrupamento de informantes, seja por ponto de enquete, seja por sexo. O mais relevante é que as curvas melódicas de interrogativas totais neutras produzidas por sergipanos são muito semelhantes entre si, principalmente o grupo que apresenta padrão final ascendente, e muito distantes de outras curvas de  $F_0$  apresentadas por outros pesquisadores que investigam outras variedades, o que demonstra um comportamento bastante identitário para essa variedade dialetal.

## Referências

- BARBOSA, P. (2000). "Syllable-timing in Brazilian Portuguese". *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo, 16(2), p. 369-402.
- CALLOU, Dinah & LEITE, Yonne. Iniciação à fonética e à fonologia. Rio de Janeiro: Zahar, 1990.
- COLAMARCO, M. C. P. (2009) "A expressão das emoções em atos de fala no Português do Brasil: produção e percepção". 2009. 189 f. *Dissertação*. Programa de Pós-Graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro.
- COSTA, M. S; Cruz, R. C. F. (2014). "Contribuições para o Atlas do Projeto AMPER-Norte: Variedade Linguística de Mocajuba (PA)". In: Raimunda Benedita Cristina Caldas; José Guilherme dos Santos Fernandes; Fernando Alves da Silva Júnior; Larissa Fontinele de Alencar. (Org.). *Mídias e Mediações Culturais*. Recife. Pipa Comunicação, p. 73-82.
- LADD, D. R (2008). *Intonational phonology*. Cambridge University Press.
- LAVÉ, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

- LEMOS, R.; Cruz, R. C. F (2013). “Contribuições para o Atlas do Projeto AMPER-Norte: Variedade Linguística de Baião (PA)”. In: IV Colóquio Brasileiro de Prosódia da Fala, Maceió (AL). *Anais do Colóquio Brasileiro de Prosódia da Fala*. Belo Horizonte, p. 230-234.
- LIRA, Z. de. (2009). “A entoação modal em cinco falares do Nordeste brasileiro”. *Tese* (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal da Paraíba, João Pessoa.
- MEIRELES, A. R. (2007). “Reestruturações rítmicas da fala no português brasileiro”. *Tese* (Doutorado em Linguística). IEL, Universidade Estadual de Campinas, Campinas.
- MEIRELES, A. R.; Barbosa, P. A. (2009). “O papel da taxa de elocução nos processos dinâmicos de mudança linguística”. *Revista (con) textos linguísticos*, (3)3, 91-116.
- MEIRELES, A.; Gambarini, V. (2011). “Tipologia rítmica de dialetos do português brasileiro. 3º Colóquio Brasileiro de Prosódia de Fala”, Belo Horizonte. *Anais*. Belo Horizonte, 4-7.
- MILAN, P. (2015). “Subsídios para uma análise prosódica do dialeto de Curitiba: uma contribuição ao projeto AMPER-POR”. Dissertação. Programa de Pós-Graduação em Linguística. Universidade Federal do Paraná.
- MORAES, J. A. (1998). “Intonation in brazilian portuguese”. In: Hirst, D.; Di Cristo A. (Ed.). *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, p. 179-194.
- MORAES, J. A. (2006) . “Melodic contours of yes/no questions in Brazilian Portuguese”. In: International Speech Communication Association (ISCA) *Tutorial and Research Workshop on Experimental Linguistics*, Atenas, p. 117-120.
- MORAES, J. A. (2008). “The Pitch Accents in brazilian portuguese: analysis by synthesis”. In: Fourth Conference on Speech Prosody. Proceedings of the Speech Prosody. Campinas : Unicamp, p. 389-397.
- MORAES, J. A.; Abraçado, M. (2005). “A descrição prosódica do português do Brasil no AMPER”. *Geolinguistique – Hors série – no. 3*, 337- 345.

- MORAES, J.A.; Colamarco, M. (2007). “Você está pedindo ou perguntando? Uma análise entonacional de pedidos e perguntas no português do Brasil”. *Revista de Estudos da Linguagem*, 15(2), p. 113-126.
- NUNES, V. G. (2011). “Análises entonacionais de sentenças declarativas e interrogativas totais nos falares lageano e florianopolitano”. *Dissertação* (Mestrado em Linguística). Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina.
- NUNES, V. G. (2015). “A prosódia de sentenças interrogativas totais nos falares catarinenses e sergipanos”. *Tese de Doutorado*. Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina.
- PAIXÃO, V. B. (2014). “A prosódia das interrogativas totais na fala carioca: fala espontânea versus leitura”. *Dissertação*. Programa de Pós-Graduação em Letras Vernáculas, Universidade Federal do Rio de Janeiro. Universidade Federal do Rio de Janeiro.
- PRIETO, P (2014). “The Intonational Phonology of Catalan”. In: *Sun-Ah Jun* (ed.). *Prosodic Typology 2. The Phonology of Intonation and Phrasing*. Oxford University Press: Oxford, p. 43-80.
- REIS, C.; Antunes L. B.; Pinha, V. C. de J. (2011). “Prosódia de declarativas e interrogativas totais nos falares marianense e belorizontino no âmbito do projeto AMPER”. In: III Colóquio Brasileiro de Prosódia da Fala, Belo Horizonte (MG). *Anais do III Colóquio Brasileiro de Prosódia da Fala*, Belo Horizonte, p. 104-109.
- SEARA, I. C.; Figueiredo, M. C. (2008). “Mais sobre a entoação de sentenças com ordem SV”. *Revista Letras*, Curitiba, v. 75-76, p. 1-10.
- SILVA, J. C. B. A (2011). “Prosódia regional em enunciados interrogativos espontâneos do português do Brasil”. *Revista Gatilho*, 8(13), p.1-13.
- TRUCKENBRODT, H.; Sandalo, F.; Abaurre, M. B. (2009). “Elements of Brazilian Portuguese intonation”. *Journal of Portuguese Linguistics*, (8)1, 75-114.
- WILDNER, A. K. (2013). “Padrões entoacionais de interrogativas totais e parciais no falar paranaense.” *Leitura*, 2(52), 185-206.

