

# UMA ABORDAGEM FONÉTICA NA FONOLOGIA ENTOACIONAL

## UN ABORDAJE FONÉTICO EN LA FONOLOGÍA DE LA ENTONACIÓN

### PHONETIC APPROACH IN INTONATIONAL PHONOLOGY

**Luciana Lucente\***

Universidade Federal de Alagoas - UFAL, Arapiraca, BR

**RESUMO:** O presente artigo discute a dificuldade de implementar descrições de correlatos fonéticos a sistemas de notação entoacional. Discutindo as diferenças entre a proposta da fonologia métrica-autosegmental e uma perspectiva dinâmica de análise da entoação, fica exposto que sistemas exclusivamente fonológicos, como o sistema ToBI, carecem de medidas para uma notação mais exata de fenômenos com tons de fronteira. Da mesma forma, o sistema DaTo, de perspectiva dinâmica, procura apresentar correlatos fonéticos em seu sistema de rótulos de contornos entoacionais, no entanto, não pode evitar a associação de uma representação para esses fenômenos, o que nos coloca em um novo impasse, pois as representações não são suficientes para a descrição exata do fenômeno. A representação de altura da frequência fundamental não aparece nos sistemas abordados no artigo. Como exemplo da necessidade de representações fonéticas é discutido, por meio da apresentação de diferenças entoacionais funcionais, como esta função poderia ser aplicada para uma melhor descrição entoacional do português do Brasil. O artigo procura, dessa forma, trazer uma discussão sobre uma abordagem fonética na fonologia entoacional.

**PALAVRAS-CHAVE:** entoação; fonologia entoacional; fonética; contorno entoacional; DaTo.

**RESUMEN:** Este artículo discute la dificultad de implementarse descripciones de correlatos fonéticos a sistemas de notación entonativa. Se discuten las diferencias entre la propuesta del modelo métrico-autosegmental y la perspectiva dinámica del análisis de la entonación lo que demuestra que a los sistemas exclusivamente fonológicos, como el sistema ToBI, le faltan medidas para una anotación más exacta de fenómenos con tonos de frontera. De este mismo modo, el sistema DaTo, de perspectiva dinámica, busca presentar correlatos fonéticos en su sistema de etiquetas de los contornos entonativos, sin embargo, no puede evitar la asociación de una representación para esos fenómenos, lo que se trata de un nuevo impase, pues las representaciones no son suficientes para la descripción exacta del fenómeno. La representación de la altura de la frecuencia fundamental no aparece en los sistemas tratados en el artículo. Como ejemplo de la necesidad de representaciones fonéticas se discute, a través de la presentación de diferencias funcionales entonativas, cómo esta función podría aplicarse para una mejor descripción del portugués de Brasil. Así, en ese artículo se busca poner en discusión un abordaje fonético en la fonología de la entonación.

**PALABRAS CLAVE:** entonación; fonología de la entonación; fonética; contorno entonativo; DaTo.

**ABSTRACT:** The aim of this article is discuss the difficulties on implementation of phonetic features in intonational annotation systems. The differences between the autosegmental-metrical model and the dynamics perspective on intonation are discussed and evaluated. As example are discussed the ToBI system and the DaTo system, na autosegmental-metrical system and a dynamic system, respectively. The ToBI system presentes a lack of phonetic features, as an exact annotation of boundary tones. On the other hand, the DaTo system needs a set of label store present the phonetic features, however, the ways of representation are not sufficient etore present the phonetic features as a whole. As example are presented some differences in height between intonational contours which have the same representation, in other examples, the differences between utterances with the same representation but diferente shapes. This article wants to invite there searchers to discuss these aspects in a phonetic approach of intonational phonology.

**KEYWORDS:** intonation; intonationalphonology; phonetics; intonationalcontour; DaTo.

## 1 INTRODUÇÃO

Os estudos sobre a prosódia, mais especificamente sobre a entoação, do ponto de vista fonético-fonológico e não mais impressionístico, baseado nos estudos de retórica e enunciação, como os escritos de Platão e Baktin, teve seus primeiros resultados nos trabalhos de Lehiste (1970) e Halliday (1970) para o inglês. Tais trabalhos, por serem desenvolvidos em uma época em que a experimentação linguística estava a cargo de psicólogos experimentais e foneticistas, tiveram primeiramente um enfoque fonético, pois estes se encontravam centrados em aspectos perceptivos da fala, e em pesquisas a partir da observação de fenômenos físicos, como a pressão subglótica na produção de proeminências na fala e pistas acústicas para explicar fenômenos entoacionais específicos, como asserção e interrogação.

---

\* Professora Adjunta da Universidade Federal de Alagoas, Campus Arapiraca. Email: [lucente.luciana@gmail.com](mailto:lucente.luciana@gmail.com).

Contudo, os trabalhos de perspectiva fonético-perceptiva careciam de ferramentas de representação das relações intrínsecas entre os níveis linguístico e paralinguístico da fala. A questão era como associar a determinada forma e medida acústica características pragmáticas e expressivas da fala. O que faltava era justamente o componente fonológico na análise prosódica.

Nesse contexto, os trabalhos de Liberman (1975) e Pierrehumbert (1980) realizam análises fonológicas para o fenômeno entoacional, assim como a proposta de um sistema de notação fonológica para a entoação do inglês americano (PIERREHUMBERT, 1980), baseado em um modelo métrico-autossegmental. Na mesma perspectiva, surgiram posteriormente os trabalhos de Beckman (1994), Bolinger (1982), enquanto, na Europa, a Escola Holandesa propunha o modelo perceptivo IPO (Institute for Perception Research) para análise da entoação.

Ladd (1996) organiza e reúne em uma única publicação pressupostos e funcionamento do que se definiu como fonologia entoacional. A fonologia entoacional, principalmente por meio da teoria métrica-autossegmental, tem como objetivo fonológico ser capaz de caracterizar adequadamente os contornos entoacionais em termos de sequências de elementos categoricamente distintos; e como objetivo fonético promover o mapeamento de elementos fonológicos em parâmetros acústicos contínuos (LADD, 1996, p.42). Ou seja, fonologia e fonética tem nessa perspectiva papéis distintos, que não se misturam.

Atualmente, apesar da predominância da perspectiva fonológica, algumas pesquisas ainda apontam para uma dicotomia entre a necessidade ou não de se ter um sistema abstrato de representação entoacional. Esse desafio se caracterizava entre fazer medidas ou construir modelos. É sob essa perspectiva que o presente artigo se desenvolve, comparando modelos e analisando o tratamento de dados entoacionais.

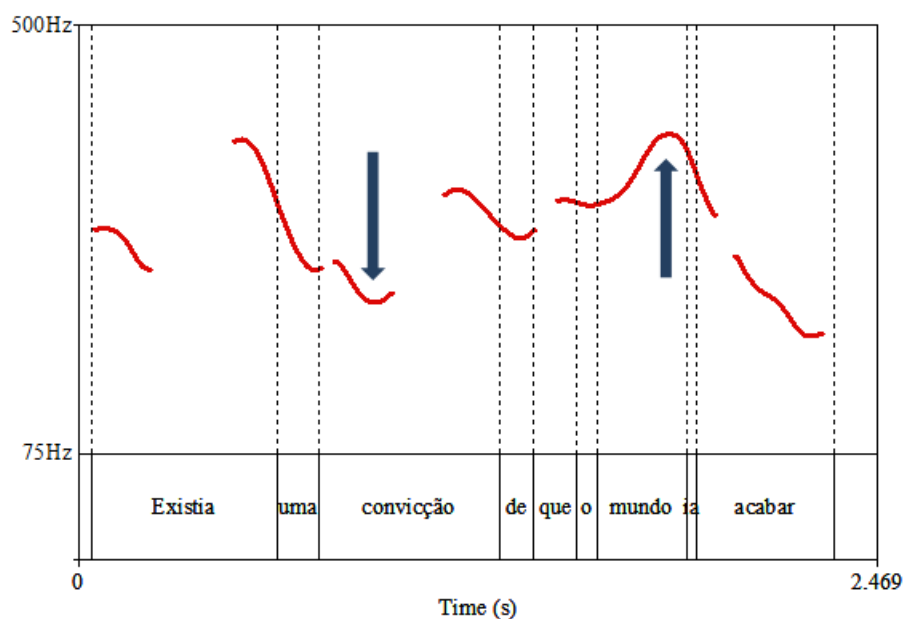
### 1.1 Entoação como fenômeno físico

A entoação aparece mais comumente na literatura relacionada aos aspectos físicos do contorno entoacional. Em um sentido estreito, a entoação é definida como “o uso de características fonéticas suprasegmentais para expressar significados pragmáticos no nível da sentença de forma linguisticamente estruturada” (LADD, 1996, p. 6). A entoação é um dos componentes da prosódia, que em termo mais amplo, inclui também a descrição de características dinâmicas e temporais associadas aos aspectos formais (BARBOSA, 2009).

As características suprasegmentais a que se refere Ladd (1996) são a frequência fundamental (doravante  $f_0$ ), intensidade e duração, os correspondentes de ordem fonética dos fenômenos psicofísicos, de pitch, volume (loudness) e duração (length) (LEHISTE, 1970), respectivamente. Dentre essas características,  $f_0$  é o correlato direto do aspecto fonético que a entoação assume nos estudos prosódicos. A  $f_0$  é medida em Hz (Hertz) e corresponde ao número de vezes por segundo em que as pregas vocais completam um ciclo de vibração. Esses ciclos são controlados pelos músculos da laringe que determinam a tensão nas pregas vocais, como também por forças aerodinâmicas do sistema respiratório sublaríngeo (BOTINIS et al, 2001). Essa definição de entoação que envolve elementos físicos e psicofísicos, ou fonéticos e fonológicos, na sua produção faz com esta possa ser vista como um componente dinâmico da fala.

Como característica fonética, a  $f_0$  da fala, medida em Hertz, apresenta maior ou menor frequência de vibração de acordo com o contexto de produção. Por exemplo, para atribuir maior ênfase a uma palavra ou segmento do discurso, o falante aumenta ou diminui a frequência de  $f_0$  durante sua produção, fazendo com que a voz se torne mais aguda ou grave e garanta a atenção desejada para o segmento, como exemplificado na Figura 1, a seguir.

Figura 1 - Exemplo de curva de  $f_0$ ; as setas indicam diminuição e aumento da frequência de vibração das pregas vocais



Na concepção clássica de Saussure (1916), a fala se contrapõe ao conceito de língua. Enquanto a língua é um produto de natureza social não constituindo uma função do falante, a fala se constitui como um ato individual de vontade e inteligência. Como um ato individual, a fala possui caráter sincrônico, refletindo fenômenos e usos em um dado momento, enquanto a língua se volta para a diacronia com seu caráter evolutivo. Dessa forma, o conceito de entoação se liga à fala e à língua, expressando as funções comunicativas presentes na fala do indivíduo que estabelecem os padrões entoacionais que dizem respeito à língua.

Na produção da fala, a entoação se manifesta como a forma mais produtiva de expressão das intenções do falante. Através da produção e manipulação físicas de  $f_0$  é possível atribuir ênfase a um determinado item, determinar se o que está sendo dito é uma asserção, uma interrogação, uma interrogação parcial, determinar foco, pressuposição e status da informação, dentre outras funções.

Essa é a definição de entoação em termos físicos, ou seja, fonéticos. A questão que surge após o exposto até o momento é onde e como a representação fonológica dessas características fonéticas se apresentam.

## 2 FONOLOGIA ENTOACIONAL

Como apresentado na introdução do presente artigo, a fonologia entoacional procura caracterizar os contornos entoacionais em termos de sequências de elementos categoricamente distintos, para tanto são empregados sistemas ou formas de notação entoacional capazes de descrever tais elementos. O que veremos a seguir é como duas diferentes perspectivas de investigação científica interpretam os fenômenos entoacionais e a forma de representação que propõem. A primeira perspectiva segue os pressupostos da teoria métrica-autossegmental (doravante teoria MA), representada no sistema ToBI (BECKMAN et al., 1994; SILVERMAN et al., 1992) de notação entoacional, a segunda, uma perspectiva dinâmica de análise da entoação, representada no sistema DaTo de notação entoacional (LUCENTE, 2008; 2012).

### 2.1 Teoria Métrica-Autossegmental

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

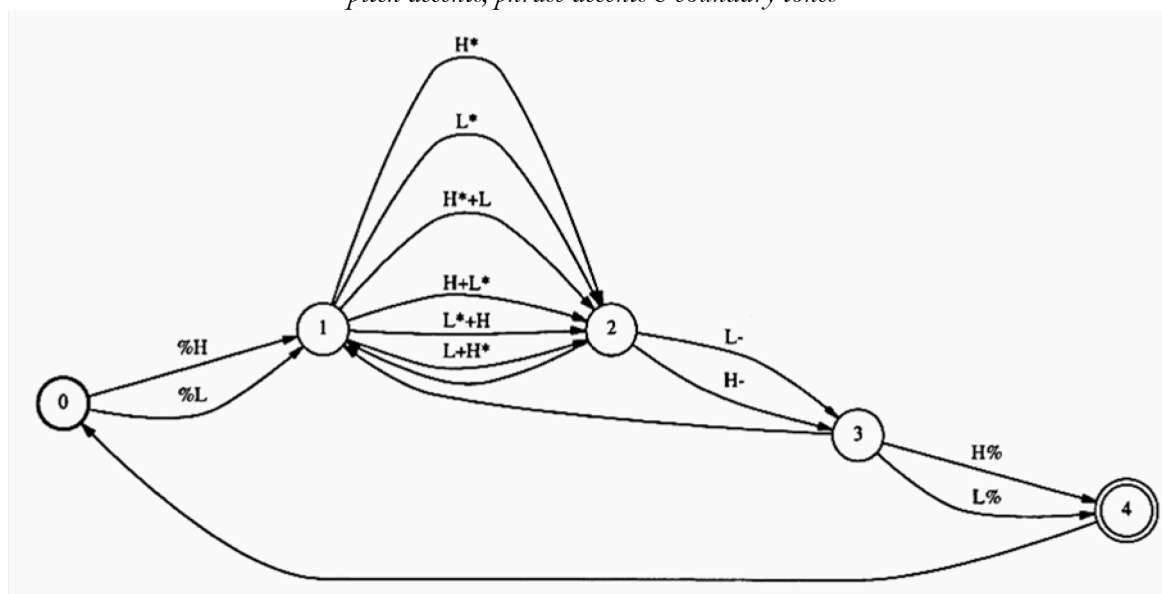
entre proeminência (*pitch accent*) e acento lexical (*stress*); iii) a análise dos *pitch accents*<sup>1</sup> em relação à altura dos tons; iv) e a relação entre características locais e tendências gerais.

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 de f0 entre eventos localizados. Os eventos considerados importantes do ponto de vista linguístico são os *pitch accents*. Os *pitch accents* são eventos locais que assinalam proeminência no nível do enunciado, e estão associados na teoria MA à altura da curva entoacional nos pontos em que esses eventos ocorrem.

Para essa teoria, as relações de proeminência são em primeira instância abstrações fonológicas, que se manifestam por meio de dois aspectos fonéticos distintos, o acento lexical (*stress*) e a posição das proeminências (*pitch accents*) (LADD, 1996). A literatura sobre o assunto, predominantemente em inglês, destaca a diferença entre *pitch accent* e *stress*, ou seja, entre proeminência e acento lexical. Diferentemente da proeminência, o acento é uma propriedade lexical de sílabas individuais, assinalada no nível da palavra. A teoria MA assume que proeminência e acento lexical são fenômenos distintos, e podem ocorrer alinhados ou não na estrutura métrica do enunciado.

Além da distinção dos *pitch accents* em relação aos acentos lexicais, estes ainda têm relação direta com a altura dos tons. Apesar de a 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 aparece na Figura 2, que mostra o inventário tonal proposto por Pierrehumbert (1980) adaptado em Botinis et al. (2001).

Figura 2 - Inventário tonal proposto por Pierrehumbert (1980), ilustrando respectivamente *boundary tones*, *pitch accents*, *phrase accents* e *boundary tones*



### 2.1.1 Sistema ToBI (*Tones and Break Indices*)

Seguindo esses pressupostos, o sistema ToBI (BECKMAN et al., 1994) apresenta uma sistematização mais econômica da proposta de Pierrehumbert (1980), ainda propondo uma notação que reflete a linearidade da estrutura tonal e a marcação dos *pitch accents*. Tal sistema foi desenvolvido por um grupo de pesquisadores

<sup>1</sup> Para uma melhor adequação com a literatura será usado o termo *pitch accent*, em inglês.

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).

Na transcrição do sistema ToBI, são utilizadas diferentes camadas, ou tiers, para marcações, que servem para a marcação de tons, transcrição ortográfica, agrupamento prosódico e para quaisquer observações sobre a transcrição consideradas importantes para a interpretação de elementos nas outras camadas.

O elenco de rótulos proposto em ToBI é composto por cinco possibilidades de marcação de *pitch accents*, e de quatro *boundary tones*, que são uma associação entre *phrase accents* e altura de fronteiras, como mostra a Tabela 1.

Tabela 1 - conjunto de rótulos de notação do sistema ToBI

<i>Pitch accents:</i>	L*	Tom baixo
	L+H*	Tom baixo seguido de tom alto alinhado a sílaba tônica
	L*+H	Tom baixo alinhado à sílaba tônica seguido de tom alto
	H*	Tom alto
	H+!H*	Downstepping
<i>Boundary Tones:</i>	L-L%	Fronteira baixa precedida por phraseaccent baixo
	H-L%	Fronteira baixa precedida por phraseaccent alto
	L-H%	Fronteira alta precedida por phraseaccent baixo
	H-H%	Fronteira alta precedida por phraseaccent alto

Os *phrase accents* se associam fonologicamente às fronteiras, porém não se alinham foneticamente a nenhum elemento segmental, como a sílaba. Esses elementos possuem uma aceitação controversa nos estudos da entoação, pois marcam proeminência no nível da sentença e sem correlato fonético, o que gera complicações teóricas, Grice et al. (2000), ao defenderem a existência de *phrase accents* para o grego e para o sueco, consideram em sua discussão que os *phrase accents* têm sido objeto de muito debate entre as línguas. Primeiramente por ser considerado por alguns estudiosos uma complicação desnecessária quando aplicado na notação do inglês e de outras línguas germânicas. Mesmo em análises que incluem os *phrase accents*, estes não aparecem com um status teórico claro (GRICE et al., 2000, p. 143)

A notação do agrupamento prosódico das palavras da sentença obedece a uma escala de números que vai de 0 a 4, de forma crescente - da sequência mais coesa para a mais disjunta - assim ficam: 0, entre palavras com agrupamento prosódico muito forte, como elisões; 1, entre diferentes palavras prosódicas; 2, para marcar forte disjuntura marcada por pausa ou pausa virtual sem marcas tonais, ou para marcar disjuntura mais fraca do que a indicada por 4, marcando uma fronteira entoacional; 3, para marcar uma fronteira intermediária na sentença; e 4, marcando uma fronteira completa, ou seja, o final de um sintagma entoacional (*intonational phrase*) (BECKMAM et al., 1994).

A implementação desta notação em enunciados específicos é feita por meio de oficinas de notação entoacional. Essas oficinas consistem na reunião de um número de participantes, cujas experiências em relação à notação entoacional sejam de diferentes níveis, para transcrição de arquivos previamente escolhidos. A coincidência de transcrições entre os participantes é que define o melhor rótulo a ser empregado na representação de determinado tom ou fenômeno entoacional. O sistema ToBI propõe uma coincidência entre transcritores de ao menos 80% (SILVERMAN et al. 1992) para o estabelecimento de

uma notação. Em Wightman (2002), que coloca em discussão o método de notação ToBI, essa coincidência é de apenas 50%.

É interessante observar que o primeiro passo para a notação no sistema ToBI e nos demais sistemas que seguem o mesmo modelo (Sp\_ToBI, GToBI, DToBI) é a percepção de outiva de correlatos como *pitch*, *loudness* e *length* (LEHISTE, 1970). A percepção dessas características irá determinar a existência de proeminências, fronteiras nos enunciados e movimento de f<sub>0</sub>; é a coincidência entre a percepção desses elementos que irá determinar a coincidência entre transcritores. Em um segundo momento se confirma em uma análise dos correlatos físicos presentes na curva entoacional.

## 2. 2 Perspectiva Dinâmica

A perspectiva dinâmica nas ciências teve seu início a partir dos estudos da motricidade humana e animal. As pesquisas de Kugler e Turvey (1987), Thelen e Smith (1994) e Kelso (1994, 1995) sobre motricidade em geral, revelaram a existência de uma coordenação funcional e auto-organizada de movimentos que pressupõe o abandono das dicotomias mente-corpo e planejamento-execução em detrimento da existência de sistemas dinâmicos atuando no controle desses movimentos.

Essa execução perfeita de movimentos no espaço é possível pela ação de um sistema dinâmico, que compreende um sistema físico composto por um número de diferentes componentes que interagem entre si e com o ambiente de uma forma auto-organizada, e que muda seu estado inicial com a passagem do tempo (van GELDER; PORT, 1995; BARBOSA, 2006).

O conceito de auto-organização se refere a uma tendência espontânea não-linear de ordenamento observada em sistemas complexos, que podem ser naturais ou artificiais. Esse ordenamento, ou interação, deve-se a características intrínsecas ao próprio processo desses sistemas complexos. Tais sistemas são sensíveis às suas condições iniciais de partida, podendo estar imersos em um regime caótico<sup>2</sup>, e à sua relação com o meio ambiente. Exemplos de auto-organização podem ser encontrados na natureza na forma de sistemas químicos, físicos, biológicos e sociais, em certos tipos de magnetismo, na estrutura de um furacão, na organização de colmeias, de movimentos sociais, na coloração de espécies animais e nos mecanismos de produção da fala.

A relação de auto-organização pressupõe o abandono das dicotomias entre mente e cérebro e planejamento e execução em detrimento de ações coordenadas. No entanto, para o controle dessas ações, o número de variáveis a serem controladas devem ser reduzidas e se correlacionarem por meio de uma sinergia<sup>3</sup> (KELSO, 1995), que permite que o sistema exiba propriedade de adaptação. Dessa relação resulta um padrão de funcionamento que age de acordo com uma frequência interna e que é capaz de alterar seu estado inicial ao longo do tempo. Esse comportamento possibilita ao observador a capacidade de previsão de eventos futuros a partir da observação do estado inicial desse sistema e de sua frequência em um dado momento.

A coordenação entre as diferentes partes desse sistema dinâmico auto-organizado possibilita a emergência de parâmetros de controle que podem influenciar o sistema como um todo. Quando ocorre algum tipo de modificação, mesmo que suave, em algum dos parâmetros de controle desse sistema, pode ocorrer em seguida o que chamamos de bifurcação no sistema. Em termos de sistemas dinâmicos, uma bifurcação é uma transição brusca em relação ao estado anterior em que se apresentava o sistema. Transições desse tipo podem ser de origem intrínseca ou ocorrer a partir de uma perturbação imposta externamente ao sistema. Quando há uma perturbação no sistema é que se torna possível determinar a relação funcional entre os parâmetros de controle, os quais podem ser identificados nesses momentos de transição (KELSO, 1995).

Por apresentar uma relação estreita com estados temporais, um sistema dinâmico, além de ter bifurcações como propriedades que ocorrem ao longo do tempo, pode dispor também de atratores cíclicos (caso seja um

<sup>2</sup> É importante notar que nem todo sistema está imerso no caos. O caos é um dos regimes possíveis em alguns sistemas a partir de certos valores e parâmetros

<sup>3</sup> Sinergia (do grego *synergía*: *syn*, cooperação + *érgon*, trabalho) é o efeito ativo e retroativo do trabalho ou esforço coordenado de vários subsistemas na realização de uma tarefa complexa. O efeito resultante da sinergia entre de vários agentes que atuam de forma coordenada para um objetivo comum tem valor superior ao da soma do trabalho desses agentes efetuados separadamente.

sistema periódico) que são padrões de movimentos oscilatórios para os quais o sistema tende com o passar do tempo ou com a realização de uma tarefa específica, ou ainda, quando sujeito a alguma perturbação externa (BARBOSA, 2006).

Segundo essa perspectiva dinâmica, além da coordenação de movimentos que possibilitam a um sujeito a realização de uma tarefa específica, existe ainda uma tendência biológica à coordenação de movimentos entre sujeitos ou entre sujeito e ambiente, fato que pode ser exemplificado por dois experimentos: o de Kelso (1984), em que sujeitos ao executarem movimentos oscilatórios com os dedos indicadores das mãos em anti-fase, entram em fase quando alcançam uma velocidade alta, passando por um limiar de transição de fase; e pelo experimento reportado em Barbosa et al. (2005) sobre uma sequência de repetições de sílabas CV (consoante-vogal) do PB que entram em fase com o ritmo de um metrônomo, coincidindo os *onsets* vocálicos das sílabas com as batidas do metrônomo.

Se esta coordenação possibilita um mecanismo de controle do movimento como abordado até o momento, na produção da fala, mais especificamente da entoação, esta hipótese também pode ser aplicada, pois esta sincronia entre elementos auto-organizados pode forçar os movimentos articulatórios que possibilitam o alcance de um alvo específico na curva entoacional a coincidirem com certos ciclos articulatórios recorrentes. Existem evidências, por exemplo, de que a sílaba sirva como estrutura coordenativa com a qual vários movimentos articulatórios estão alinhados (FUJIMURA, 2000). Segundo Xu (2005), isso significa que pode ser difícil para um falante manter uma relação de fase entre movimentos de f0 e sílaba senão por uma sincronia que revela um mecanismo atrator.

### 2.2.1 Perspectiva dinâmica no estudo da entoação

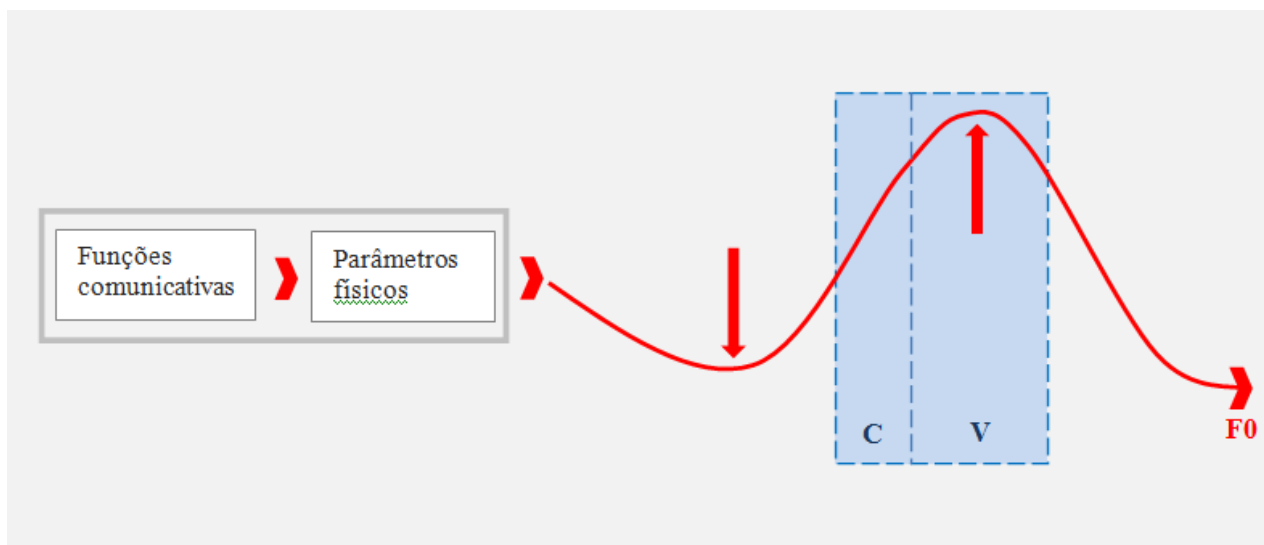
Enquanto a teoria MA se baseia em aspectos lineares da estrutura tonal, na identificação dos *pitch accents* e no alinhamento abstrato dessa estrutura com o material linguístico, o programa dinâmico-entoacional se concentra na convergência de aspectos fonéticos – velocidade, intensidade, altura, duração – 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.

Os contornos entoacionais, enquanto um dos primitivos melódicos desse programa é definido por uma trajetória ideal da curva entoacional especificada por um alvo a ser atingido (*pitch target*) associado a uma unidade segmental, provavelmente uma sílaba (XU, 2005). Na especificação dessa trajetória da curva até alcançar seu alvo estão envolvidas, além dos parâmetros físicos, funções comunicativas.

O modelo PENTA (*Parallel Encoding and Target Approximation*) proposto por Xu (2005, 2006) ilustra como diferentes funções e parâmetros estão codificados nos contornos entoacionais. Nessa proposta, funções comunicativas de ordem lexical, frasal, focal, de tópico, etc. são codificadas em parâmetros físicos, como duração, gama de variação tonal e intensidade, que possibilitam o alcance do alvo e seu alinhamento com a sílaba na curva entoacional a uma altura e taxa de aproximação específicas (XU, 2005; 2006).

Considerando a definição de contorno dinâmico e a proposta do modelo de Xu (2005; 2006), a Figura 3 ilustra o conceito de contorno dinâmico, que é definido aqui 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.

Figura 3 - Em um contorno entoacional, as (i) funções comunicativas são codificadas em parâmetros físicos; expressas em uma (ii) trajetória ideal de  $f_0$  com um alvo a ser atingido (setas); e (iii) alinhado a uma unidade segmental (sílabas tônicas)



A gama de variação tonal delimita o intervalo de *pitch* no qual um alvo é implementado, e pode ser definido pelos parâmetros do valor do alvo e intervalo de variação. A gama de variação tonal é de grande importância para a especificação de foco, pois é de acordo com as variações do valor do alvo de *pitch* que antecedem e seguem o foco é que é possível percebê-lo. Essas variações de *pitch* podem compor um esquema de codificação do foco, ou mesmo constituírem propriedades intrínsecas a este (XU, 2005). Da mesma forma, por meio das variações no valor do alvo de *pitch*, podemos determinar fenômenos como mudanças de tópicos na fala e distinções entre tipos de enunciados, como declarativas e interrogativas. As Figuras 4 e 5 ilustram, respectivamente, um enunciado declarativo e um enunciado interrogativo. É possível observar nessas figuras que os padrões de  $f_0$  são distintos e que a gama global de variação no enunciado declarativo é de aproximadamente 90Hz (201Hz-111Hz), enquanto no interrogativo é de aproximadamente 136Hz (233Hz-98Hz).

Figura 4 - Enunciado declarativo com a marcação dos valores dos picos de  $f_0$  em Hertz

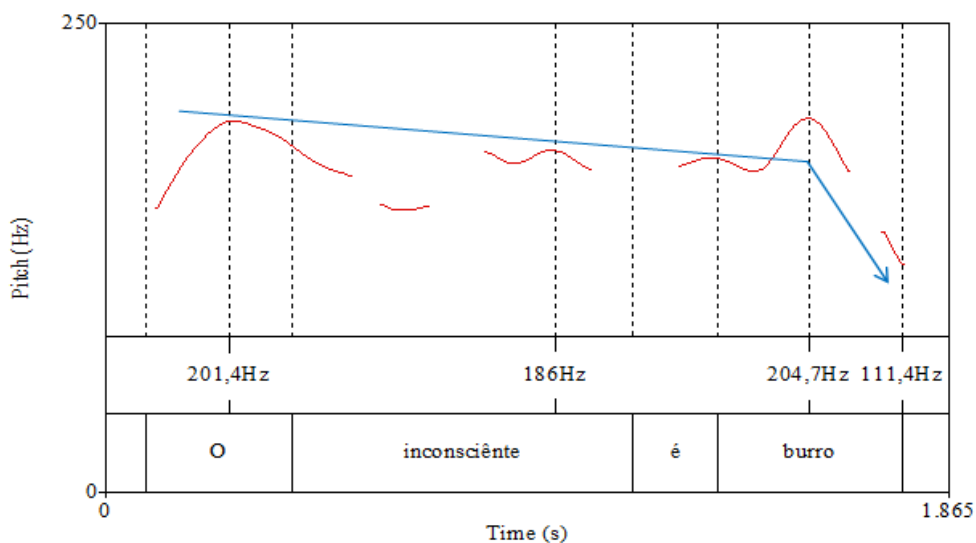
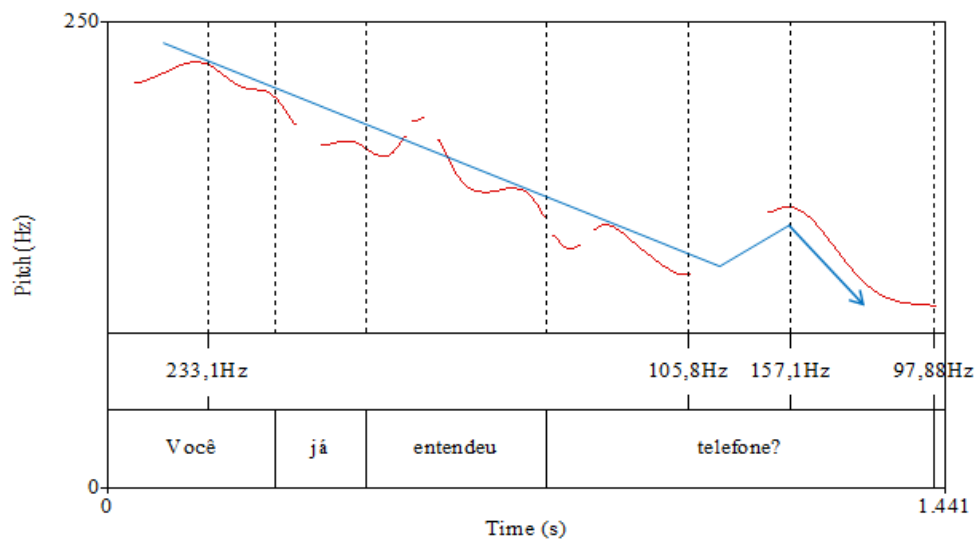


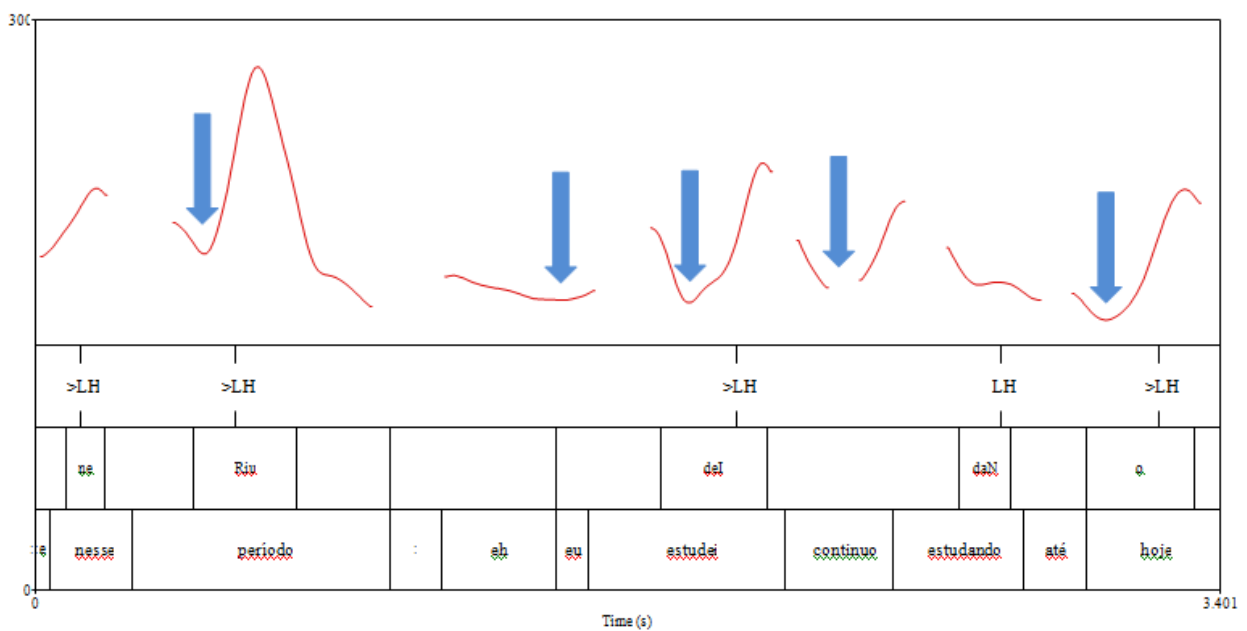


Figura 5 - Enunciado interrogativo com a marcação dos valores dos picos de f0 em Hertz



Cada contorno dinâmico, marcador de proeminência, tem movimento e altura relacionados à altura e ao movimento de f0 que o antecede e 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 f0, mas a um movimento da curva entoacional ao qual se atribui marcação de proeminência. A Figura 6 é ilustrativa de como cada movimento na curva se relaciona com o anterior, pois, para cada movimento de subida ou descida na curva, o movimento anterior tem que proporcionar condições para que isso ocorra, ou seja, uma descida para que ocorra subida, e vice-versa.

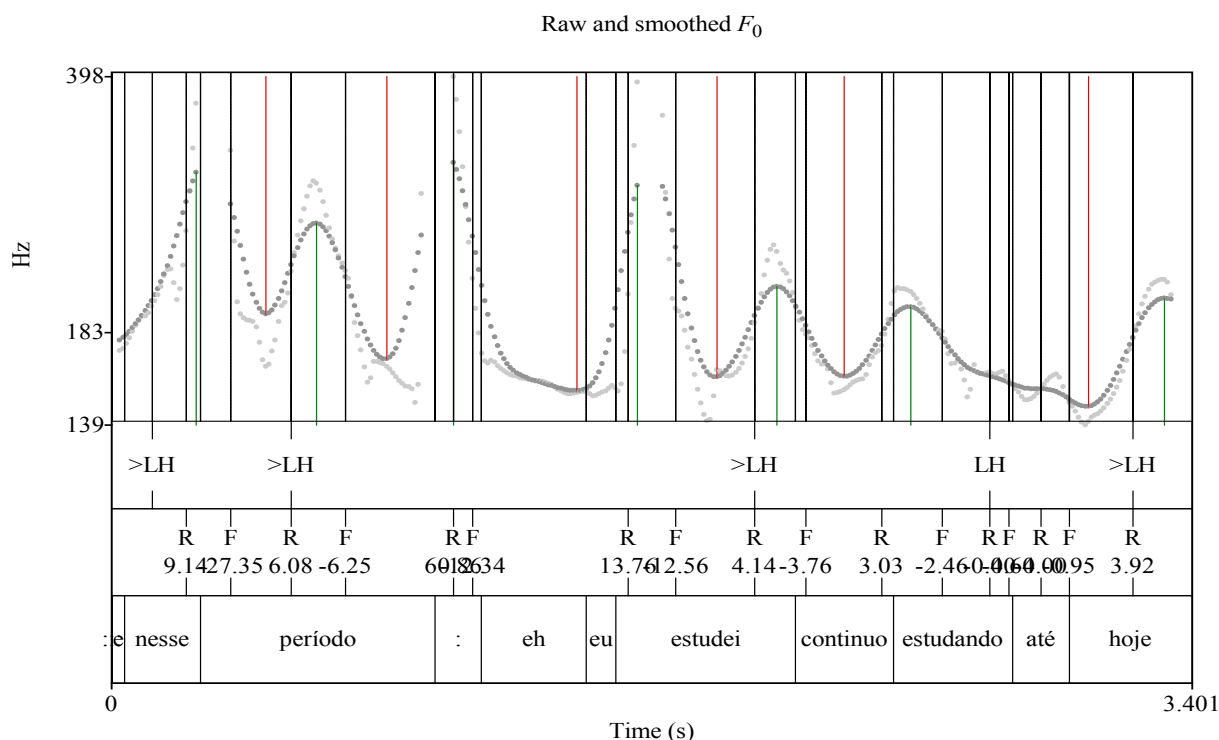
Figura 6 - Movimentos dinâmicos dos contornos entoacionais com as setas indicando os vales antes das subidas; na primeira camada a notação entoacional; na segunda a marcação das sílabas tônicas; na terceira a transcrição ortográfica



Essa inter-relação entre os movimentos da curva entoacional defendida pela interpretação dinâmica dos fenômenos que compõem a estrutura entoacional fica evidente ao se extrair a função derivada da curva de f0. Ao realizar a extração da função derivada de f0 por meio do *script* FO\_extrema.praat (ARANTES, 2009), pode-se observar a existência de um controle na produção de f0 que faz com que esta coordene os

movimentos da curva de acordo com a velocidade de subida e descida de  $f_0$  e com o alvo a ser atingido. A Figura 7 mostra essa relação: em cinza a curva suavizada de  $f_0$ ; em vermelho, estão indicados os pontos de mínima altura de  $f_0$ ; em verde, indicados os pontos de máxima altura. Os pontos que indicam os picos de máxima velocidade de subida (R) e descida (F) de  $f_0$  estão alinhados com a notação na segunda camada da figura.

Figura 7 - Curva entoacional de enunciado; na primeira camada indicando os contornos entoacionais; na segunda os pontos de máxima velocidade de subida e descida de  $f_0$  (R e F; respectivamente *rising* e *falling*); e na terceira a transcrição ortográfica



O alinhamento específico entre contornos dinâmicos e o material linguístico compreende uma questão central na defesa desta abordagem, pois enquanto a teoria MA estabelece um alinhamento abstrato da curva entoacional cujo ponto específico é decidido por regras de implementação fonética (PIERREHUMBERT, 1980; 2000), a perspectiva dinâmico-entoacional estabelece um alinhamento específico (ATTERER e LADD, 1994) entre esses elementos, a ponto de este evidenciar a existência de uma sincronia entre os movimentos que produzem  $f_0$  e movimentos articulatorios que geram os padrões espectrais, apesar de esses movimentos serem controlados separadamente. O fato de existir uma sincronia entre movimentos independentes não significa que haja uma total interdependência entre eles, pois estes podem estar limitados por determinados graus de liberdade na execução de movimentos concomitantes (XU, 2005). Os pressupostos e funcionamento da notação DaTo, dentro de uma perspectiva dinâmica da entoação, com contornos dinâmicos e níveis de fronteira, assim como em outros componentes da notação serão apresentados na próxima seção.

### 2.2.2 Sistema DaTo (*Dynamic Tones*)

O sistema DaTo, além de apresentar uma notação dos fenômenos entoacionais do PB a partir de uma perspectiva dinamicista, procura também apresentar em suas camadas de notação informações pragmático-comunicativas e aspectos fonéticos do ritmo da língua. Para tanto, a notação completa se divide em seis camadas, das quais três são fixas e servem para a notação de: i) contornos dinâmicos, ii) segmentação semi-automática da fala em unidades V-V, e iii) transcrição ortográfica; e outras três camadas, que podem ser omitidas ou apresentadas de acordo com o objetivo de cada notação, que são: iv) detecção automática de pausas/grupos acentuais, v) informação pragmática e vi) segmentação em unidades discursivas (DS).

Os contornos dinâmicos, como definidos anteriormente, são movimentos contínuos de  $f_0$ , que percorrem a transição de um ponto a outro na curva entoacional até atingir o seu alvo. Esse 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.

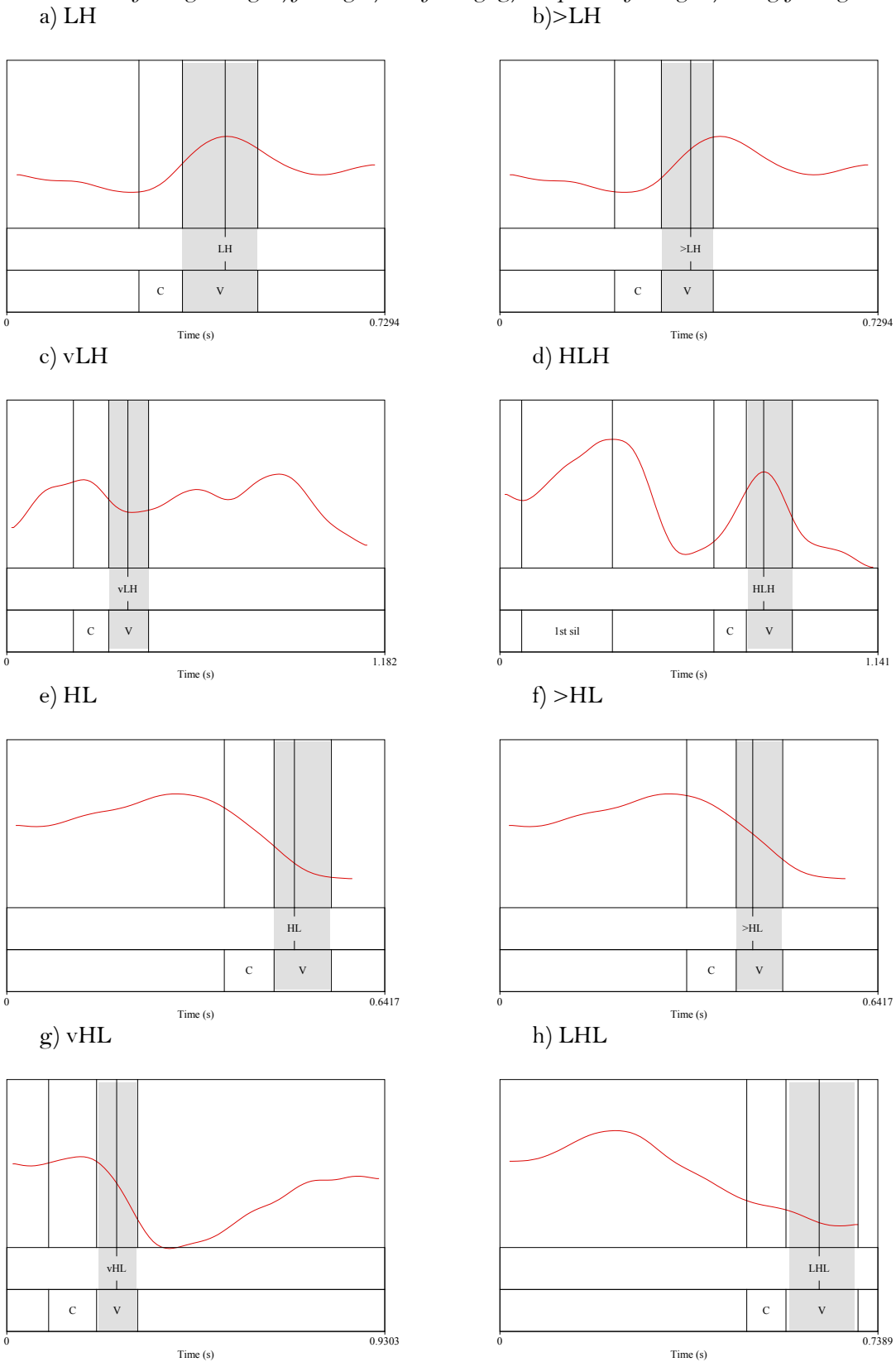
No sistema DaTo, os contornos dinâmicos se dividem entre ascendentes, descendentes e níveis de fronteira, e são representados por rótulos que ilustram o movimento de  $f_0$  pelas letras H (*high*) e L (*low*), assim com na teoria MA, esses rótulos foram mantidos por julgar tais rótulos produtivos e ilustrativos, e, portanto, de mais fácil compreensão.

O conjunto de contornos entoacionais do sistema DaTo, descritos na Tabela 2 e exemplificados na Figura 8, 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. Diferentemente, a teoria MA pressupõe inicialmente a percepção de outiva dos *pitch accents* e posteriormente sua verificação na curva. Um dos pressupostos do treinamento para notação nos sistemas que seguem a teoria MA é a concordância entre sujeitos sobre o tipo de movimento tonal que conseguem detectar de outiva, sem a inspeção na curva entoacional.

Tabela 2 - Descrição dos contornos entoacionais e níveis de fronteira do sistema DaTo

Contornos Dinâmicos		
LH	<i>rising</i>	contorno que parte de uma posição baixa na sílaba pré-tônica, alcança seu pico alinhado à vogal tônica e tem um <i>resetting</i> final.
>LH	<i>late rising</i>	contorno que parte de uma posição baixa na consoante da sílaba tônica e tem todo o movimento de subida alinhado à vogal tônica, alcançando seu pico após a vogal tônica, acompanhado de um <i>resetting</i> final
vLH	<i>compressed rising</i>	contorno formado entre dois picos em palavras distintas; não há espaço temporal para que ocorra um pico na vogal tônica nessas condições, fazendo com que se realize comprimido entre dois picos.
HLH	<i>falling-rising</i>	apresenta a mesma estrutura que LH, porém antecedido por um pico na mesma palavra; geralmente ocorre em palavras com mais de três sílabas, tendo o primeiro pico alinhado à primeira sílaba da palavra e o seguinte à sílaba tônica.
HL	<i>falling</i>	contorno que parte de uma posição alta na sílaba pré-tônica, alcança seu nível mais baixo alinhado à vogal tônica; ocorre em fronteiras de enunciados assertivos.
>HL	<i>late falling</i>	contorno que parte de uma posição alta na consoante da sílaba tônica e tem todo o movimento de descida alinhado à vogal tônica, alcançando seu nível mais baixo após a vogal tônica; também ocorre em fronteiras de enunciados assertivos.
vHL	<i>compressed falling</i>	contorno que parte de uma posição alta na sílaba pré-tônica, porém não alcança seu nível mais baixo alinhado à vogal tônica, pois não há espaço temporal para alcançar o nível mais baixo, tendo que realizar um movimento de subida logo em seguida.
LHL	<i>rising-falling</i>	contorno que se inicia em posição alta e realiza um <i>downstepping</i> durante a duração das palavras seguintes; ocorre em fronteiras de enunciados assertivos.
Níveis de Fronteira		
L	<i>low</i>	nível de fronteira baixo.
H	<i>high</i>	nível de fronteira alto.

Figura 8 - Contornos dinâmicos do sistema DaTo para o PB: a) *rising*; b) *late rising*; c) *compressed rising*; d) *falling-rising*; e) *falling*; f) *late falling*; g) *compressed falling*; h) *rising-falling*



### 2.2.2.1 Unidades V-V

A camada da transcrição que apresenta a segmentação dos enunciados em unidades V-V compreende a divisão de todo enunciado em segmentos compreendidos entre os *onsets* de duas vogais. Essa segmentação é feita semi-automaticamente através do programa Beat Extractor (ver BARBOSA, 2006), que gera no programa Praat uma camada contendo a segmentação dessas unidades. Esse programa identifica os *beats*, pontos específicos associados ao aumento da amplitude na transição consoante – vogal (C-V). A detecção desses *beats* indica fronteiras próximas aos *onsets* das vogais e, portanto, auxilia na segmentação das unidades V-V.

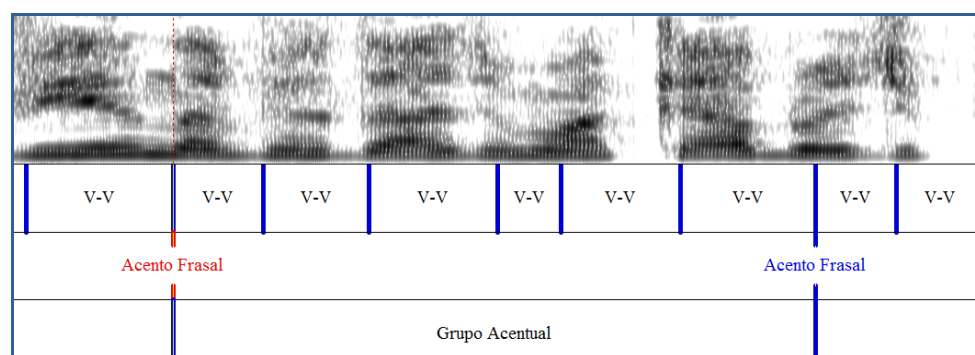
Este tipo de segmentação no sistema DaTo, assim como proposto em Barbosa (2006) e Barbosa et al. (2005), é usada para delimitar uma unidade cujas fronteiras são distintas das da sílaba fonológica e possui a vantagem de dinamizar a transcrição do sistema na medida em que a detecção de transições C-V por meio desse tipo de segmentação auxilia na observação de fenômenos de ordem física que podem estar relacionados à produção e percepção (MARCUS, 1981; DOGIL and BRAUN, 1988; WONG and SCHREINER, 2003), e na determinação do ritmo (BARBOSA, 2006) e da entoação.

De acordo com Barbosa (2006), as unidades V-V apresentam uma estabilidade duracional que oferece um efeito de compensação e uma extensão homogênea sobre a duração dos segmentos. Devido a essas propriedades, unidades V-V são usadas como blocos de construção para delimitação de grupos acentuais.

### 2.2.2.2 Grupos Acentuais

Os grupos acentuais (doravante SG do inglês *stress group*)<sup>4</sup> são unidades delimitadas por dois acentos frasais consecutivos. Os acentos frasais não se referem a aspectos sintáticos, mas sim a aspectos fonéticos, pois são definidos como proeminências no domínio da produção da fala, assinaladas por picos locais de duração ao longo dos enunciados (BARBOSA, 2006, p.10). Por sua vez, os picos locais de duração que assinalam os acentos frasais são determinados pelos padrões de duração das unidades V-V. A relação entre unidades V-V, acentos frasais e grupos acentuais aparece ilustrada na Figura 9.

Figura 9 - Relação entre a segmentação em unidades V-V na primeira camada; acentos frasais na segunda camada; e grupos acentuais (SGs) na terceira camada



No sistema DaTo, os grupos acentuais são determinados automaticamente através do programa SG Detector (ver BARBOSA, 2006) que, assim como no caso das unidades V-V, gera uma camada com a segmentação dos grupos acentuais no programa Praat. O programa SG Detector detecta os acentos frasais pelo cálculo do desvio padrão das médias de duração das unidades V-V normalizadas pelo cálculo do *z-score* estendido (ver fórmula em BARBOSA, 2006). Dessa forma o programa pode indicar onde estão os pontos de maior duração das vogais e, portanto, os acentos frasais.

<sup>4</sup> Aqui também será adotada a terminologia em inglês.

A delimitação dos SGs através da detecção dos acentos frasais possibilita a localização e marcação automática, baseada em aspectos fonéticos, das fronteiras entoacionais, as mesmas fronteiras que em ToBI são marcadas com os números 1, 2 e 3 baseando-se apenas na percepção (BECKMAN et al., 1994). A marcação dos SGs se mostrará bastante funcional no sistema DaTo, pois, como se verá a seguir, ela será responsável pela coordenação entre aspectos prosódicos e discursivos.

### 3 A POSSIBILIDADE DE UM SISTEMA DE NOTAÇÃO ENTOCIONAL FONÉTICO-FONOLÓGICO

A intensão do sistema DaTo é apresentar um sistema de representações fonológicas dos contornos entoacionais do PB, e de outras línguas, baseando-se em medidas fonéticas e na percepção do ouvinte. Pela aplicação de programas computacionais para extração dos *onset* de vogais e de acentos frasais usando medidas fonéticas é possível implementar análise fonética ao sistema. Contudo, tais medidas apresentadas automaticamente não possuem representação no sistema, a escolha pela segmentação automática de fronteiras foi incorporada justamente pela ausência de representação fonológica, o que oferece economia e maior exatidão na marcação das fronteiras entoacionais.

O emprego do conceito de contornos entoacionais também é uma tentativa de aproximação da fonética. A definição dos contornos é por si uma definição fonética, pois leva em conta padrões de movimento dos contornos de  $f_0$ . No entanto, tais contornos não podem ser empregados na ausência de representação fonológica. O sistema ToBI oferece uma notação exclusivamente fonológica, que parte da percepção e é representada por rótulos de *pitch accents*, *phrase accents* e tons de fronteira. No sistema ToBI, não é realizada qualquer medida para a determinação de seus rótulos.

A questão, portanto, é como prescindir de representação? O modelo PENTA, proposto por Xu (2005; 2006), se apresenta como um modelo exclusivamente fonético e, portanto, não prevê qualquer tipo de notação. Nesse sentido, este pode ser empregado por meio de implementações ou mesmo com fins computacionais, isso não ocorre, por exemplo, com o sistema DaTo.

A análise da entoação do PB, usando o sistema DaTo, mostra que existe uma necessidade de diferenciar contornos entoacionais não somente pela forma da curva, mas também pela diferença de gama tonal. A Figura 10 mostra a ocorrência de dois contornos ascendentes sucessivos, em um mesmo item lexical, que se diferenciam pelo alinhamento em relação à sílaba tônica e pela altura do pico de  $f_0$ , que apresenta uma diferença de 90 Hz. Em Lucente (2010; 2012), é descrito um experimento de percepção cujo resultado mostra que os sujeitos são mais sensíveis à percepção de proeminência quando está associada a mudanças paralelas de alinhamento e altura de  $f_0$ .

Essa evidência mostra a necessidade de uma forma de representação de diferenças de altura de  $f_0$  para a descrição exata da proeminência encontrada em dados de fala. Outro fator que corrobora para a procura de uma forma de representação da gama de  $f_0$  é o fato de enunciados diferentes apresentarem a mesma sequência de notação, porém com efeito comunicativo/pragmático diferente, e vice-versa, como se pode inferir a partir da observação das Figuras 11 e 12.

Figura 10 - Diferenças de alinhamento e altura dos contornos entoacionais nas repetições da palavra “mesmo”

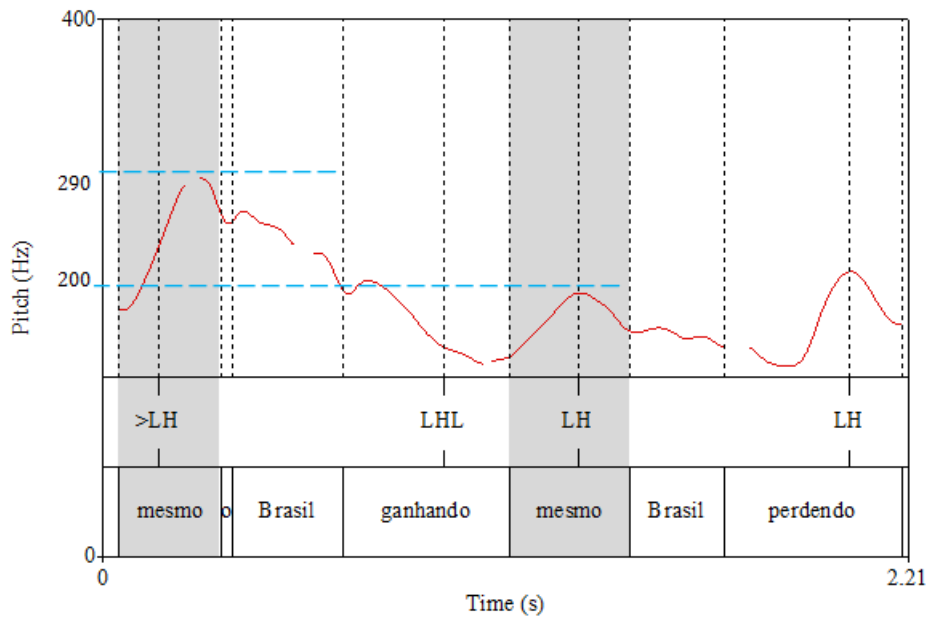


Figura 11 - Exemplo de fala com padrão entoacional semelhante com notação diferente: o abaixamento de  $f_0$  no segmento inicial de “vitória” não foi suficiente para a realização do pico marcado em “vítima”.

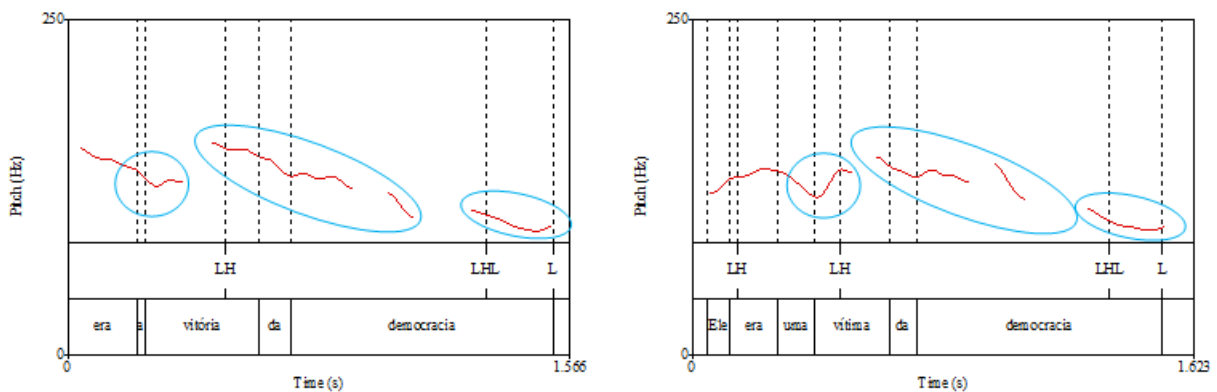
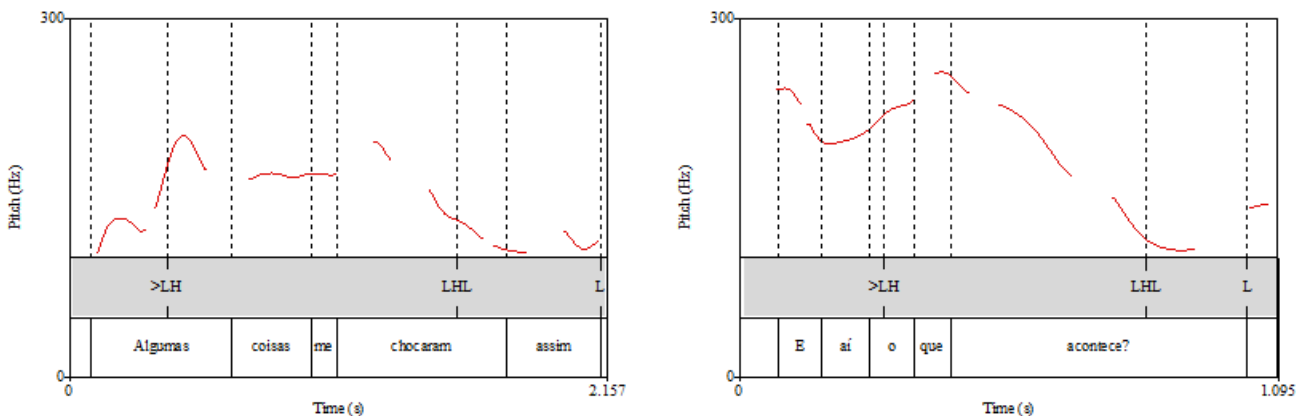


Figura 12 - Enunciados assertivo e interrogativo com mesma notação entoacional



Essa discussão pretende avaliar o uso de sistemas fonéticos de notação entoacional, pois, como evidenciado pelos exemplos, estes podem deixar de fora da notação aspectos importantes da descrição entoacional, como a variação de altura de  $f_0$ . Por outro lado, a tentativa de implementar aspectos fonéticos aos sistemas de notação não é tarefa fácil, pois a informação fonética abre mão de representações, representações estas que

se tornam essenciais para a descrição entoacional. Retornamos aqui ao impasse apresentado na introdução, entre fazer medidas e seguir um modelo.

Fica aqui um convite à reflexão sobre a construção e o uso de modelos fonético-entoacionais.

## REFERÊNCIAS

- ATTERER, M. and LADD, D. R. On the phonetics and phonology of 'segmental anchoring' of Fo: evidence from German. *Journal of Phonetics*, v. 32, p.177–197, 2004.
- ARANTES, P.; BARBOSA, P.A. Duration and fundamental frequency pattern in Brazilian Portuguese in polysyllabic words. *Journal of Acoustical Society of America*, v. 124, n. 4, 2009.
- BARBOSA, P. A. *Incursões em torno do ritmo da fala*. Campinas: Pontes, 2006.
- \_\_\_\_\_. Prominence- and boundary-related acoustic correlations in Brazilian Portuguese read and spontaneous speech. In: SPEECH PROSODY 2008 CONFERENCE. *Proceedings...* Campinas, 2008.
- BARBOSA, P. A.; ARANTES, P.; MEIRELES, A. R.; VIEIRA, J. M. Abstractness in speech-metronome synchronisation: P-centres as cyclic attractors. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOG, 9., Interspeech 2005. *Proceedings...* Lisbon, Portugal, 2005. p. 1441-1444.
- BECKMAN, M. E.; HIRSCHBERG, J.; PITRELLI, J. F. *Evaluation of prosodic transcription labeling reliability in the ToBI framework*, 1994. Disponível em [http://www.ling.ohio-state.edu/~tobi/ame\\_tobi](http://www.ling.ohio-state.edu/~tobi/ame_tobi).
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer* (Version 5.1.05) [Computer program]. Retrieved May 1, 2009. From: <http://www.praat.org/>.
- BOLLINGER, D. Intonation and its parts. *Language*, v. 58, p 505–533, 1982.
- BOTINIS, A.; GRANSTRÖM, B.; MÖBIUS, G. Developments and paradigms in intonation research. *Speech Communication*, v. 33, n. 4, p. 263–296, 2001.
- DOGIL, G.; BRAUN, G. *The PIVOT model of speech parsing*. Vienna, Áustria: Verlag, 1988.
- FUJIMURA, O. The C/D model and prosodic control of articulatory behavior. *Phonetica*, v. 57, p.128–138, 2000.
- GRICE, M.; LADD, D. R.; ARVANITI, A. On the place of phrase accents in intonational phonology. *Phonology*, v. 17, n. 2, p.143–185, 2000.
- HALLIDAY, M.A.K. *A course in spoken English: intonation*. Oxford: Oxford University Press, 1970.
- KELSO, S. Phase transitions and critical behavior in human bimanual coordination. *American journal of Physiology: regulatory, integrative and comparative*, v. 15, 1984.
- \_\_\_\_\_. *Dynamic Patterns*. Cambridge: MIT Press, 1995.
- KUGLER P.N.; TURVEY, M.T. *Information, natural law and the self assembly of rhythmic movement*. Hillsdale: Erlbaum, 1987
- LADD, D. R. *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- LEHISTE, I. *Suprasegmentals*. Cambridge: MIT Press, 1970.
- LIBERMAN, M. *The intonational system of English*. Ph.D. Thesis, MIT, 1975.
- LUCENTE, L. *Aspectos dinâmicos da entoação e da fala no português brasileiro*. Tese (Doutorado em Linguística) - Unicamp, Campinas, SP, 2012.
- \_\_\_\_\_. *DaTo: um sistema de notação entoacional do português brasileiro baseado em princípios dinâmicos. Ênfase no foco e na fala espontânea*. Dissertação (Mestrado em Linguística) - Unicamp, Campinas, SP, 2008.
- \_\_\_\_\_; BARBOSA, P. A. The role of alignment and height in the perception of LH contours. In: CONFERENCE ON SPEECH PROSODY, 5. *Proceedings...* Chicago, 2010.
- MARCUS, M. A *Theory of syntactic recognition for natural language*. New York: MIT Press, 1980.



- PIERREHUMBERT, J. *The Phonology and Phonetics of English Intonation*. Ph.D thesis, MIT, 1980.
- SAUSSURE, F. et al. *Curso de linguística geral*. 24. ed. São Paulo, SP: Cultrix, [1916] 2002.
- SILVERMAN, K.; BECKMAN, M., PITRELLI, J. et al. TOBI: A standard for labeling English prosody. *International Conference on Speech and Language Processing (ICSLP)*, v. 2, 1992. p. 867-870.
- van GELDER, T.; PORT, R. It's about time: an overview of the dynamical approach to cognition. In: PORT, R.; van GELDER, T. (Eds.). *Mind as motion: dynamics, behavior, and cognition*. Cambridge, MA: MIT Press, 1995.
- WIGHTMAN, C.W. ToBI or not ToBI. In: SPEECH PROSODY, Aix-en-Provence. *Proceedings...* France, 2002. p. 25-29.
- WONG, S.W.; SCHREINER, C.E. Representation of CV-sounds in cat primary auditory cortex: intensity dependence. *Speech Communication*, v. 41, p. 93-106, 2003.
- XU, Y. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, v. 27, p.55-105, 1999.
- \_\_\_\_\_. Speech melody as articulatorily implemented communicative functions. *Speech communication*, v. 46, p. 220-251, 2005.
- \_\_\_\_\_. Speech prosody as articulated communicative functions. In: SPEECH PROSODY 2006, *Proceedings...* Dresden: Germany, 2006.

***Recebido em 20/10/2013. Aprovado em 25/05/2014.***