

Transcrevendo a entonação

Carlos Alexandre V. Gonçalves*

Abstract

In this paper, I present and discuss the main intonational transcription systems, in order to show (i) the advantages of INTSINT e (ii) the similarities between the INTSINT and non-linear phonology about Stress and Rhythm (Metrical and Prosodic Phonology).

Introdução

Partindo do pressuposto de que *as variações de altura constituem o principal correlato físico para a percepção da estrutura entonacional das sentenças* (cf. Cutler & Ladd, 1983: 205), diferentes sistemas de transcrição têm sido desenvolvidos para representar movimentos de altura e respectivos contornos entonacionais considerados perceptivelmente relevantes nas línguas do mundo.

Na verdade, sistemas de representação necessariamente incorporam, como lembram Hirst & Di Cristo (1999: 18), certo número de hipóteses sobre que possíveis variações de traços prosódicos são linguisticamente pertinentes e, por isso mesmo, universais. Nesse sentido, um sistema notacional constitui, necessariamente, uma Teoria sobre a Entonação.

Neste texto, objetivo tecer comentários gerais sobre os sistemas notacionais mais usualmente empregados na transcrição da Entonação. De uma forma geral, procuro mostrar as vantagens do INTSINT (“INternational Transcription System for INtonation), elaborado por Hirst & Di Cristo (1999), sobre os modelos de representação de Pike (1943), baseado em níveis/escalas de altura, e o da Escola Britânica (cf. Firth, 1947), calcado na representação de curvas. Busco evidenciar que o INTSINT proporciona maior integração entre os níveis segmental e entonacional e, em decorrência, aproxima-se mais

* Doutor em Linguística pela UFRJ e Professor Adjunto II de Língua Portuguesa do Departamento de Letras Vernáculas da UFRJ.

das atuais abordagens fonológicas sobre (1) o Acento e (2) o Ritmo, dado que associa relações de proeminência não a vogais ou a sílabas, mas a estruturas fonológicas autônomas, como o Pé Acentual, a Palavra Prosódica e o Grupo Entonacional, categorias hierárquicas semelhantes às propostas pelos modelos prosódicos de Selkirk (1984 e 1995) e de Nespor & Vogel (1986 e 1989), por exemplo¹.

1. Representação em níveis

O sistema de transcrição desenvolvido por Pike (1943) utiliza ondas quadradas e linhas sigmóides para representar variações melódicas. Pike (*op. cit.*) reconhece quatro níveis de altura, com vistas a diferenciar o enunciado “expressivo” (ou “enfático”, como ressaltada à página 42) do considerado “normal” ou “básico”. Assim, para registrar as modulações de altura detectadas num dado enunciado, Pike define os seguintes níveis:

- (a) nível básico ou médio, que representa pelo número 2;
- (b) nível alto (representado por 3);
- (c) nível baixo (número 1); e, finalmente,
- (d) nível de ênfase (número 4).

Nesse sentido, a Ênfase, *para efeitos de contraste e expressão emocional do falante*, é realizada em um nível de altura considerado superior àquele empregado para o tom alto (cf. Pike, *op. cit.*: 43). Confirma-se o sistema de transcrição adotado por Pike na representação expressa pela Figura (01) a seguir:

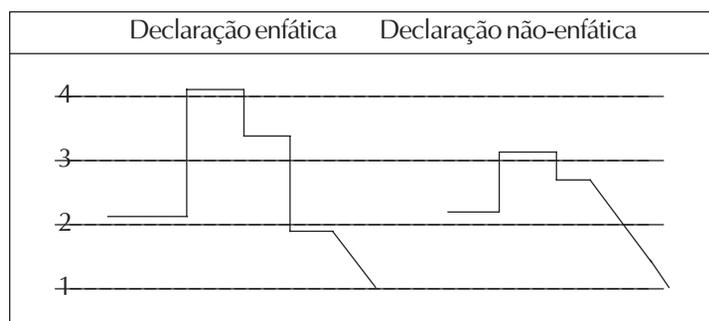


Figura (01): Representação do contorno entonacional, tal como concebido por Pike (1943).

¹ Por exemplo, a palavra prosódica, não necessariamente isomórfica em relação ao *output* do componente morfológico, é definida de modo bastante parecido na abordagem de Hirst & Di Cristo (1999) e na de Nespor & Vogel (1986): categoria/domínio de atribuição do Acento Lexical.

2. Representação em curvas

Na década de 40, paralelamente aos avanços do Estruturalismo Clássico, a Escola Inglesa (ou *The Prosodic Analysis*), cujo principal expoente vem a ser John R. Firth, desenvolveu trabalho próprio para os fenômenos entonacionais, independentemente das orientações teóricas em vigor. Essa abordagem teve como um de seus principais objetivos a confecção de sistemas notacionais para a Entonação. A proposta de Firth se põe explicitamente como alternativa à Fonologia Segmental, praticada pelos estruturalistas europeus continentais (Praga) e pelos norte-americanos, especialmente K. Lee Pike).

A *Prosodic Analysis* caracteriza-se por ser bidimensional, dado que relaciona entidades de níveis distintos – segmental e suprasegmental – e polissistêmica, uma vez que formula uma série de sistemas para diferentes estruturas fonológicas ou para diferentes contextos. Conforme ressalta Soares (1986: 12), Firth analisa as características da sílaba considerando a estrutura prosódica como um sistema de relações sintagmáticas. Nas palavras da autora, a *Prosodic Analysis* assume uma visão polissistêmica, ao abstrair *traços de acordo com domínios dos quais eles são extraídos e, uma vez extraídos, tais traços serão vistos como prosódias* (grifo da autora).

Os “Estudos Funcionais da Entonação” foram marcados por um sistema de notação do tipo interlinear, através do qual seqüências de pontos (curvas) representam variações de tom (altura). Esse sistema de notação foi utilizado em diversos trabalhos sobre a língua portuguesa, tanto a falada em Portugal (cf., entre outros, os artigos de Freitas, 1993, e de Mata 1992), quanto a falada no Brasil (cf., por exemplo, as pesquisas de Aubert, 1978; e de Aubert & Hochgreb, 1981). Nesse modelo de representação, as linhas de topo e de base funcionam como limites máximo e mínimo das variações de frequência fundamental ao alcance da voz humana. Conforme se vê na Figura (02) a seguir, a combinação de pontos representa contornos de altura e os sinais (-) e (•) simbolizam sílabas não-acentuadas e acentuadas, nesta ordem.

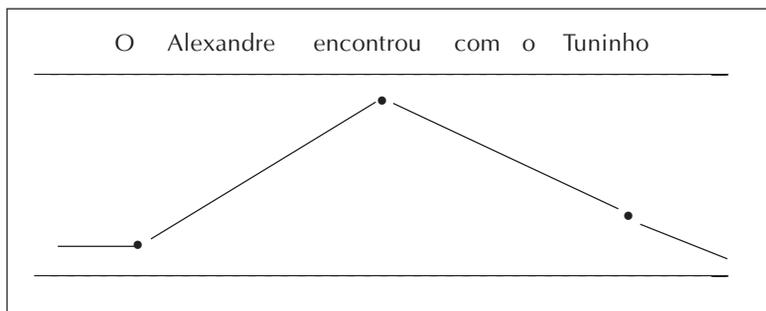


Figura (02): Sistema de notação desenvolvido pela Escola Britânica (cf. Crystal, 1969).

Os traços prosódicos fisicamente manifestados – frequência fundamental (F₀), intensidade e duração – podem, na teoria funcionalista de Firth (op. cit.: 25), expandir-se por vários domínios: desde a sílaba até a sentença propriamente dita. Além disso, as relações de proeminência são derivadas da interação direta entre os três correlatos mencionados e *dão origem a padrões rítmicos característicos*.

3. Separando trigo do joio

A principal diferença entre o modelo de Pike (1943) e o da Escola Inglesa está no enfoque dado à notação em níveis, no primeiro caso, e à notação em curvas, no segundo. Nos dois modos de representação, pode-se dizer que há pouca integração entre os níveis segmental e entonacional, o que, de certa forma, prejudica uma abordagem que procure observar o grau de interação entre a Fonologia e a Sintaxe, como a que empreendi no tratamento da Focalização (ou Foco stricto sensu), fenômeno discursivo-pragmático materializado pela Sintaxe e pela Prosódia (cf. Gonçalves, 1999). Por essa razão, para representar a Entonação dos dados enfáticos que analisei, fiz uso do INTSINT, da autoria de Hirst & Di Cristo (1999). Os motivos pelos quais optei por esse sistema são os seguintes, além do já apresentado:

- (a) em primeiro lugar, esse sistema de transcrição sem dúvida alguma é mais apropriado que os demais para lidar com relações de proeminência – jogo de valores em que um batimento é considerado forte em oposição a outro contíguo, mais fraco –, uma vez que faz uso de diferentes símbolos para representar cada variação de altura perceptivelmente relevante (levando-se em conta sempre o ponto precedente na cadeia). Por descrever um fenômeno de manifestação também acentual, a Focalização, julguei mais oportuno trabalhar com o INTSINT;
- (b) em decorrência disso, optei pelo uso do INTSINT devido ao fato de ele se aproximar mais das atuais abordagens sobre o Acento e sobre o Ritmo (cf, por exemplo, Nespor & Vogel, 1986 e 1989, Selkirk, 1984 e Hayes, 1991), dado que associa relações de proeminência não a vogais ou a sílabas, mas a estruturas fonológicas autônomas, como o Pé Acentual, a Palavra Prosódica e o Grupo Entonacional (cf. Hirst e Di Cristo, op. cit. 30-1), o que de forma alguma causa estranheza, haja vista que as subteorias não-lineares vêm se caracterizando, cada vez mais, pela utilização/incorporação de aspectos prosódicos há muito detectados por foneticistas (cf. Selkirk & Shen, 1990).

No modelo de Hirst & Di Cristo (cf. p. 29-35), está implícita a idéia de função de agrupamento, uma vez que cada sílaba acentuada combina, de uma forma ou de outra, com sílabas não-acentuadas adjacentes, formando unidades prosódicas mais altas, como os Pés Acentuais (*Stress-foot*) e os Grupos Acentuais (*Stress-group*). Nesse sentido, a parentetização etiquetada, reivindicada pelo modelo de Nespor & Vogel (1986), parece estar contemplada no INTSINT.

Como se sabe, elencar categorias fonologicamente relevantes, que envolvam domínios mais extensos que as unidades rítmicas propriamente ditas (pés e sílabas), é o principal objetivo da Fonologia Prosódica, uma das subteorias que, desde a década de 70, vem sendo englobadas pelo rótulo **não-linear**. Dessa maneira, os pés, centro de interesse na Fonologia Métrica (cf. Liberman & Prince, 1977, e Hayes, 1991), são apenas um dos constituintes abordados pelo modelo prosódico.

Nespor & Vogel (1986) enumeram as unidades relevantes para uma descrição de natureza prosódico-entonacional e evidenciam sua significância, recorrendo a processos segmentais que carecem desses constituintes como domínio. Os níveis de estruturação prosódica, nesse tipo de abordagem, são em número de sete: (a) a sílaba, (b) o pé, (c) a palavra prosódica, (d) o grupo clítico, (e) o enunciado fonológico, (f) o sintagma entonacional, e, por fim, (g) o enunciado fonológico. Em Nespor & Vogel (op. cit.), as unidades de agrupamentos superiores são formadas por membros inferiores, isto é, cada categoria consiste de uma seqüência discreta de categorias do próximo nível, mais baixo, constituindo a aludida parentetização etiquetada, como se vê na Figura (03)¹.

(c) em terceiro lugar, conforme já mencionei mais acima, o sistema de notação desenvolvido por Hirst & Di Cristo (1999) permite estabelecer melhor conexão entre os níveis segmental e entonacional dos enunciados, o que possibilita integrar à transcrição outras características prosódicas também relevantes na veiculação do Acento e do Ritmo, como a duração e a intensidade².

1 Em Gonçalves (1998), afirmei que Topicalização/Deslocamento à Esquerda, de um lado, e Focalização, de outro, diferenciam-se, entre outras coisas, pelo domínio prosódico de materialização de cada fenômeno. TOP/DE manifestam-se na chamada Frase Entonacional (I), ao passo que a Focalização Contrastiva tem como domínio o Sintagma Fonológico (j). Os exemplos abaixo permitem detectar tal diferença. Observe-se que em (01), um caso de Deslocamento à Esquerda, a pausa é maior entre o constituinte alçado para a posição de SN inicial e externo à sentença, havendo, ainda, alongamento consideravelmente grande na tônica desse mesmo constituinte, o que nos leva a postular o domínio de I. Em (02), um caso de Focalização, há uma pausa breve entre os constituintes e elevação da freqüência fundamental não na tônica do termo enfatizado, mas na segunda átona pretônica, o que nos leva a pressupor o domínio de (j). Essas diferenças em termos de categorias prosódicas tem seqüências nas realizações alofônicas. (01), com fronteira de I, bloqueou a regra de degeminação (sândi vocálico externo), ao passo que (02), com fronteira de j, não impediu a fusão dos [a] contíguos.

(01) [A Zilál]I, a Aline não briga. [+ DE, não-contrastivo: a sílaba **an** é a proeminente]

(02) [A Zilál]j a Aline não briga; não a Rose. [+ Foco: o artigo é a proeminente do grupo]

2 O Acento Enfático, por exemplo, manifesta-se através da atuação conjunta dos três correlatos prosódicos (função pluriparamétrica). Há necessidade, portanto, de integrar à transcrição (comportamento da freqüência fundamental) também referências acerca da duração e da intensidade.

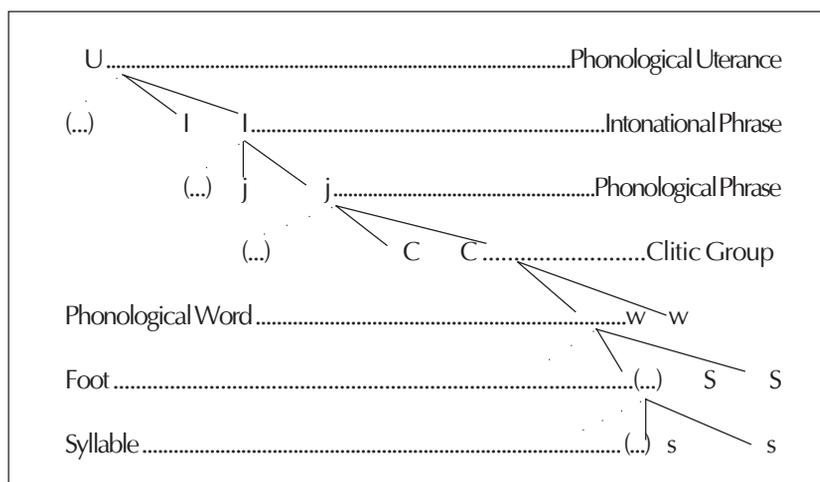


Figura (03): Categorias prosódicas, na concepção de Nespor & Vogel (1986: 230-8).

- (d) no INTSINT, devido ao fato de os símbolos serem escritos numa linha sob o texto ortográfico ou fonético, é possível transcrever padrões entonacionais independentemente da divisão em grupos sintáticos, o que condiz, de uma forma ou de outra, com a proposição de um módulo prosódico autônomo, independente do sintático. Essa orientação vem sendo defendida no âmbito das Fonologias Não-Lineares, que propõem um componente fonológico autônomo, relacionado com o sintático pelas categorias prosódicas (cf Nespor & Vogel, 1986 e 1989 e Gonçalves, 1999); e, por fim,
- (e) os símbolos utilizados no INTSINT são de mais fácil acesso, pois são disponíveis em termos de programas de editoração de textos. Além disso, por ter sido desenvolvido com o intuito de ser internacional esse sistema incorpora símbolos já empregados pelo IPA para representar movimentos de tons (elevação e abaixamento).

4. O INTSINT

De acordo com Hirst & Di Cristo (1999: 21), a maneira mais interessante de analisar as curvas de freqüência fundamental vem a ser representá-las através de uma seqüência de alvos de altura (*pitch targets*), partindo do pressuposto de que, na ausência de outra indicação, a altura é intercalada por uma curva nivelada (*smooth*) entre os alvos. Os autores evidenciam que no INTSINT está implícita a idéia de que o fator mais importante na transcrição é a altura de cada ponto sucessivo, relativamente ao ponto anterior. Dessa

maneira, não é necessário especificar a escala de altura em cada ponto da unidade entonacional, visto que o principal elemento da representação vem a ser a altura de cada ponto sucessivo em relação ao ponto precedente.

No INTSINT, fronteiras entre unidades entonacionais são delimitadas por meio de colchetes, que representam, mais objetivamente, os Grupos Entonacionais (GEs) sob análise (ou *Phonological Utterance*, na proposta de Nespor & Vogel, 1986). Nesse sistema de transcrição, cada símbolo de altura é definido como um alvo, que pode ser determinado de duas maneiras:

- (1) referindo-se a uma altura relativa à extensão total da unidade entonacional, caso em que toma o valor de topo ('*top*') ou de base ('*bottom*') e
- (2) referindo-se a pontos de altura definidos como relativamente (a) mais altos ('*highers*'), (b) mais baixos ('*lowers*') ou (c) os mesmos ('*the same*') que um ponto precedente, dentro da mesma unidade prosódica.

Como se vê, o uso dos colchetes pode servir à delimitação de qualquer unidade prosódica e não apenas à separação dos Enunciados Fonológicos (ou GEs). Os colchetes sinalizam fronteiras entonacionais e, por isso, podem ser utilizados de forma encaixada/recursiva, evidenciando diferentes níveis de hierarquização prosódica. Mais uma vez, o INTSINT se aproxima das abordagens métricas/prosódicas, que têm como premissa básica a estruturação das categorias fonologicamente relevantes. A principal vantagem desse modelo está na flexibilidade com que constituintes podem ser construídos. Esse fato possibilita que uma mesma sentença seja fatiada de diversas maneiras, gerando nuances de significado em cada parseamento³.

O INTSINT opera, ainda, com dois outros símbolos, desta feita para representar um leve abaixamento ('*downstep*') ou uma ligeira elevação ('*upstep*') em relação ao ponto precedente (cf. Quadro (01)). *Downstep* (⤵) e *Upstep* (⤴) refletem, pois, movimentos de altura sensivelmente menores que os estabelecidos por *Lower* (⤴) e *Higher* (⤵), respectivamente.

O ponto inicial de uma unidade entonacional pode ser marcado, dentro do respectivo colchete, como topo (⤴) ou como base (⤵). Um colchete inicial não-marcado representa uma altura inicial média ('*mid*'). O ponto final, diferentemente, pode ser representado como 'topo', 'base', 'mais alto', 'mais baixo' ou 'o mesmo'. Esses são os pontos que ocorrem mais freqüentemente em final de unidades entonacionais. No entanto, nada impede que outros

3 Em "Encontrei a Joana assustada", o termo 'assustada' pode ou não constituir unidade prosódica distinta de [Joana]. No caso de valer como predicativo do sujeito, 'assustada' forma um sintagma fonológico distinto de [Joana]. No caso de predicativo do objeto, não ha fronteira fonológica entre [Joana] e [assustada]: ambos os termos figuram dentro da mesma unidade: o sintagma fonológico.

símbolos sejam utilizados nessa posição, a depender dos padrões melódicos da língua ou do dialeto em estudo. Confirmam-se as representações para o início e para o fim de um GE na Figura (04) a seguir.

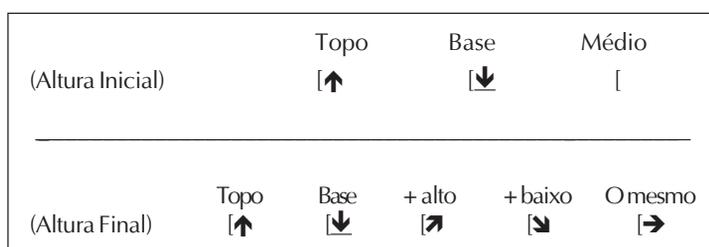


Figura (04): Início e fim de Grupo Entonacional no INTSINT.

A fim de representar o espraçamento de F₀ de uma sílaba a outra, comum em determinados tipos de Focalização que rastreei pela ACVL (cf. ‘Ênfase Contrastiva Corretiva – Gonçalves, 1997), optei por ampliar o leque de símbolos existentes no INTSINT. Assim, incorporei à transcrição os sinais (↗) e (↘), respectivamente para o espraçamento em elevação e para o espraçamento em declínio. Prefiro trabalhar com esses símbolos, em virtude da ambigüidade que pode surgir entre espraçamento e elevação/declínio progressivo, característicos de certos contornos entonacionais, como, por exemplo, a própria linha de declinação. Confira-se, na Figura (05) a seguir, de que maneira o INTSINT representa o Padrão Não-Enfático Básico (PNEB), típico de sentenças declarativas simples, sem muitas complicações estruturais, como encaixamentos sintáticos e deslocamentos.

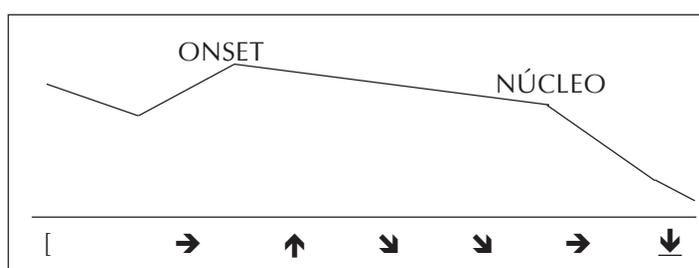


Figura (05): Representação do PNEB (Padrão Não-Enfático Básico) pelo INTSINT.

Dado que o INTSINT é um sistema restrito à transcrição da frequência fundamental, sugiro formalização adicional para representar outras

símbolos	significados
[]	Unidade (grupo) entonacional
↑	Topo ('top')
↓	Base ('bottom')
↗	Mais alto ('higher')
↘	Mais baixo ('lower')
→	O mesmo ('the same')
↗	Leve elevação ('upstep')
↘	Ligeiro declínio ('downstep')
↗	Espraiamento em elevação
↘	Espraiamento em declínio
'	Sílaba intensa
"	Sílaba superintensa
—	Sílaba longa
=	Sílaba ultralonga
È	Sílaba breve
^	Pausa
#	Fronteira vocabular
S	Sílaba tônica
s	Sílaba átona

Quadro (01): Lista de símbolos utilizados (sistema INTSINT e outros)

Conclusão

Por todos os motivos enumerados no decorrer da exposição, pode-se considerar o INTSINT o sistema notacional mais apropriado para representar contornos entonacionais relevantes na descrição de fenômenos de manifestação acentual e/ou rítmica. Além da facilidade na utilização dos símbolos, o INTSINT tem a vantagem de operar de modo independente do nível segmental, o que torna mais produtiva uma abordagem que busque descrever a interação entre os módulos sintático e fonológico da Gramática das línguas. Acredito, no entanto, que a maior vantagem do INTSINT é o que se pode chamar de semelhança com modelos formais, o que proporciona, no tratamento de fenômenos métricos, melhor integração entre Fonética e Fonologia.

Referências Bibliográficas

AUBERT, F. Foco, sintaxe e entoação. *Foco e pressuposição*. Uberaba: Faculdades Integradas de Santo Tomás de Aquino, n. 4, p. 42-52, 1978.

- _____. & HOCHGREB, N. Descrição perceptiva da entonação da frase interrogativa em português: variante mineira – Uberaba. *Série Estudos: português, estudos lingüísticos*. Uberaba: Faculdades Integradas de Uberaba, n. 7, p. 9-22, 1981.
- CUTLER, A. & LADD, D. R. *Prosody: models and measurements*. Berlin: Springer-Verlag, 1983.
- FIRTH, F. Sounds and prosodies. In: JONES, P. & LAVER, W. (orgs.). *Phonetics in Linguistics*. London: Longmann, 1947. p. 47-65.
- FREITAS, M. J. Contributo para o estudo de padrões de estruturação temporal da fala no Português. Europeu. In: DELGADO MARTINS, M. R. (ed.). *Estudos em prosódia*. Lisboa: Colibri, 1993. p. 75-103.
- GONÇALVES, C. A. V. *Focalização no português do Brasil*. Rio de Janeiro, 1997. 401 p. Tese (Doutorado em Lingüística) – UFRJ, Faculdade de Letras.
- _____. Foco e topicalização: delimitação e confronto de estruturas. *Revista de Estudos da Linguagem*, v. 7, n.1, p. 31-50, 1998.
- _____. Ênfase no português do Brasil: fonologia e sintaxe. In: GÄRTNER, E.; HUNDT, C. & SCHÖNBERGER, A. (eds.). *Estudos de lingüística do texto*. Frankfurt am Main: TFM, 1999. p. 32-65.
- HAYES, B. *Mletrical stress theory: principle and case studies*. UCLA, 1991.
- HIRST, D. & Di CRISTO, A. *A survey of intonation Systems*. Cambridge: Cambridge University Press, 1999.
- LIBERMAN, M. & PRINCE, A. 5. On stress and linguistic rhythm. *Linguistic Inquiry*, v. 8, p. 249-336, 1977.
- MATA, A. I. A questão da entoação na interrogação em português: isso é pergunta? In: DELGADO MARTINS, M. R. (org.). *Estudos em Prosódia*. Lisboa: Colibri, 1992. p. 33-73.
- NESPOR, M. & VOGEL, I. *Prosodic Phonology*. Dordrecht: Foris Publications, 1986.
- _____. On clashes and lapses. *Phonology*, v. 6, p. 69-115, 1989.
- PIKE, K. L. *Phonetics*. 13a. ed. Ann Arbor: The University of Michingam Press, 1943.
- SELKIRK, E. *Phonology and syntax: the relations between sound and structure*. Cambridge: The MIT Press, 1984.
- _____. Sentence prosody: intonation, stress and phrasing. In: GOLDSMITH, J. A. (org.). *The handbook of phonological theory*. Oxford: Basil Blackwell, 1995. p. 550-69.
- _____. & SHEN, T. Prosodic domains in Shanga Chinese. In: INKELAS, S. & ZEC, D (orgs.). *The Phonology-Syntax connection*. Chicago: University of Chicago Press, 1990. p. 313-37.
- SOARES, M. F. Suprassegmentos e fonologias pós-gerativas: teorias de domínios e processos. *Cadernos de Estudos Lingüísticos*, v. 16, n. 1, p. 65-73, 1986.

