

Análise entonativa formal: INTSINT aplicado ao português

CELESTE, L.C.^{1*}

REIS, C.^{1*}

¹Universidade Federal de Minas Gerais – UFMG

Abstract

There is a general consensus among scholars that one same sentence can be said in different ways. This consensus, however, falls apart when the issues are discussed related to the efficiency of prosody in allowing us to say the very same sequence of segments with different meanings. According to Hirst (1987) e t'Hartet. al. (1990) the difficulties of transcribing prosody overcome the difficulties of transcribing the segmental. They argue that in order to investigate how the melody contributes to the overall understanding of an enunciation a metalanguage was required through which the phenomenon can be discussed. The vocabulary of this metalanguage must consist of appropriate descriptive units through which it is possible to speak of entities and structures at various levels of observation. The semi-automatic analysis, by means of the programs MOMEL and INTSINT, was used in the description of some intonational languages such as: French (di Cristo, 2011), British English (Auran, Bouzon and Hirst, 2004; Brierley, 2011) and analysis of the Brazilian Portuguese rhythm by means of manual INTSINT (Gonçalves, 2000). The present study aims to determine if the program of semi-automatic intonational analysis - INTSINT - is capable of reproduce melodic variation trends of the Brazilian Portuguese. In our proposal, the effectiveness of the program to differentiate between declarative and interrogative forms during reading was evaluated. The following hypotheses were raised: (I) the analysis of intonation with the program INTSINT allows the choice of different phonological theories regarding the use of entonative units;(II) it is possible to distinguish between declarative and interrogative forms by means of coded target points to Brazilian Portuguese. An explanatory analysis of the codification performed by MOMEL-INTSINT was performed. The corpus of this study consists of 10 short texts, with approximately five sentences that were read by 10 people of the female sex, aged 20-30 years. The acoustic analysis of the data was performed using the program Praatversion 4.4.27, available at www.praat.org (Weenink and Boersma, 1997 ,i contains the necessary extensions for the implementation of the programs MOMEL / INTSINT. These last ones are freely available on the site <http://aune.lpl.univ-aix.fr/~auran/english/ressources.html>. The high beginning was common for the declarative and interrogative modalities. This fact was represented by the first target point, coded as T, or the second target point, coded as M; but followed by H, T or U. Nevertheless, they differ at the end of the unit: in interrogative sentences, a sharp increase represented by the symbols T and H can be observed just before the final fall. In declarative sentences it is possible to perceive a rise is followed by a downfall, but not as much and not as systematically as in the interrogative sentences. Far from closing the discussion on intonation units and in the chosen modalities for this study,our initial goal of evaluating the behavior of the program INTSINT against pitch variations was reached. We conclude that the program INTSINT is able to transmit trends and also the general melodic pattern of F0 curve.

Keywords: intonation, phonological representation, INTSINT, modalities.

1. Introdução

Há consenso entre os estudiosos quando se diz que uma mesma frase pode ser dita de diferentes formas. Porém, tal consenso se desfaz quando se incluem questões relacionadas à maneira mais eficiente de se entender como a prosódia nos possibilita dizer a mesma sequência de segmentos com significados diferentes.

Para Hirst (1987) e t'Hart et. al. (1990), as dificuldades de se transcrever a prosódia superam em muito as dificuldades de se transcrever o segmental. Defendem que, para se investigar como a melodia contribui para o entendimento global de um enunciado, é necessária uma metalinguagem por meio da qual podemos discutir o fenômeno. O vocabulário de tal metalinguagem deve consistir de unidades descritivas adequadas através das quais possamos falar de entidades e estruturas em vários níveis de observação.

Segundo Hirst e Di Cristo (1998), a prosódia contribui de alguma forma para o significado (função) em todas as línguas naturais. E Hirst (2005) afirma que as formas prosódicas são também certamente universais. No entanto, o mapeamento entre forma e função é diferente para as diversas línguas. Um exemplo claro pode ser exposto quando vimos em Ladd (1996) que o fenômeno da declinação é comum em frases declarativas, para o inglês. Por outro lado, Gussenhoven (2004) mostra que as frases declarativas em chickasaw apresentam configuração oposta, ou seja, a frequência sobe no final do enunciado. Assim, uma distinção clara entre forma e função é necessária para um estudo mais profundo da entonação (HIRST, 2005). Para tanto, é necessário que se tenha uma diferença clara entre os níveis de representação da entonação.

De acordo com t'Hart e Collier (1975), Pierrehumbert (1987) e Gussenhoven (2004), para representação da entonação é necessário distinguir pelo menos dois níveis de representação, um físico e outro abstrato.

Hirst, Di Cristo e Espesser (2000) argumentam que, no nível linguístico mais abstrato, o objetivo é representar de alguma forma o conhecimento que os falantes necessitam para aprender uma língua. No outro extremo, desejam relatar a representação da manifestação física desse conhecimento linguístico, que corresponde às características acústicas e fisiológicas dos enunciados. Nesse momento distinguem-se, claramente, os dois níveis de representação defendidos por t'Hart e Collier (1975), Pierrehumbert (1987) e Gussenhoven (2004): um físico e outro abstrato.

Porém, argumenta-se que uma teoria satisfatória e global da entonação necessita de algo mais entre o puramente físico e o extremo abstrato. O modelo para estudo da entonação desenvolvido pelo *Laboratoire Parole et Language* e seguido neste artigo apresenta quatro níveis de representação, de acordo com Hirst, Di Cristo e Espesser (2000): (1) Físico, (2) Fonético, (3) Fonológico de superfície e (4) Fonológico subjacente.

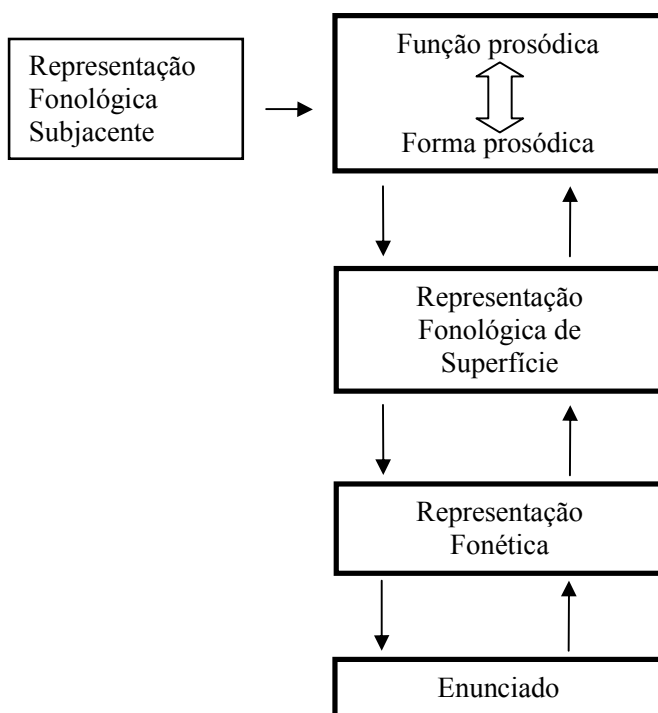
O nível físico exprime a acústica e a fisiologia da entonação. Os aspectos acústicos que estão tradicionalmente relacionados à prosódia são: frequência fundamental, duração e intensidade (Hirst, Di Cristo e Espesser, 2000). Esses são reflexos de alterações fisiológicas que ocorrem no trato vocal durante a fonação (por exemplo, o número de ciclos realizados por segundos de vibração das pregas vocais corresponde aos valores de frequência fundamental).

Para que não haja um salto dos dados acústicos para a representação abstrata, é usada a representação fonética.

O nível fonético consiste em transformar os dados acústicos em poucos valores quantitativos capazes de reproduzir os dados acústicos originais sem perda significativa de informação. Tal mudança é chamada de modelização e estilização da curva de frequência fundamental e pode ser realizada por diferentes programas. A finalidade do MOMEL (MODélisation de MELodie) é reduzir a curva de frequência fundamental a pontos-alvo, podendo ser utilizado como base para o INTSINT (International Transcription System for Intonation) (HIRST, 2005; 2011), e como base para conversão do contorno da variação melódica na aplicação em sistemas de conversão de voz (ODEBRECHT e SEARA, 2009). O MOMEL foi proposto originalmente por Hirst entre 1980 e 1983 e automatizado por Hirst e Espesser em 1993. O uso do algoritmo MOMEL se justifica pelo uso da função quadrática spline, que resulta numa curva contínua e suave. A estilização através dessa função produz uma curva bem próxima à original de frequência fundamental, sem perda de informações significativas (HIRST, 2005).

O terceiro nível citado, fonológico de superfície, é responsável por converter os valores quantitativos em qualitativos. Esse também é realizado sem perda significativa de informação. Assim, os dados quantitativos fornecidos pelo MOMEL podem ser passados através do INTSINT para representações qualitativas (HIRST, 2005).

Finalmente, o nível fonológico subjacente compreende tanto a representação da função prosódica como da sua forma (HIRST, DI CRISTO e ESPESSER, 2000). A figura 1 mostra a relação entre os quatro níveis de representação:



Adaptado de Hirst, Di Cristo e Espesser (2000)

Figura 1: Níveis de representação da entonação

A maioria dos estudos engloba o esquema acima representado simultaneamente, ou seja, há uma mistura entre o estudo da forma e da função. No entanto, o estudo exclusivamente da forma se faz necessário primeiramente porque os fenômenos prosódicos, apesar de quase universais, não apresentam a mesma forma em todas as línguas. E ainda, estudar a relação entre forma e função sem uma separação clara entre essas duas pode vir a formar um ciclo vicioso (HIRST, 2005).

1.1. O Programa INTSINT

Os pontos-alvo modelizados pelo programa MOMEL, descrito anteriormente, podem ser interpretados de diversas maneiras. Representar os pontos-alvo no nível fonológico de superfície utilizando o sistema de transcrição INTSINT é uma delas (HIRST, 2011). O INTSINT foi desenvolvido durante a preparação de um estudo entonacional de 20 línguas com a expectativa de se postular um sistema de códigos para análise prosódica que fosse capaz de representar qualquer distinção entonativa significativa (HIRST e DI CRISTO, 1998).

Tal proposta visa destrinchar as distinções entonativas usadas em diferentes línguas para construir os padrões que diferenciam os contornos melódicos de cada uma delas. Sob esse ponto de vista, os códigos do INTSINT seriam equivalentes ao sistema de transcrição de sons IPA. Assim, o modelo em si está

totalmente concentrado na representação prosódica, mais do que em sua função (HIRST e DI CRISTO, 1998; HIRST, 1999).

Para descrição da entonação, uma série limitada de símbolos tonais é utilizada nos pontos-alvo estimados pela técnica MOMEL (LOUW e BARNARD, 2004). Tais símbolos ortográficos abstratos definidos para representar esses pontos-alvo estão listados abaixo, seguidos de sua abreviatura e da proposta de tradução para o português:

- ✓ *Top*: T, topo;
- ✓ *Mid*: M, médio;
- ✓ *Bottom*: B, base;
- ✓ *Higher*: H, mais alto;
- ✓ *Same*: S, igual;
- ✓ *Lower*: L, mais baixo;
- ✓ *Upstepped*: U, subida suave;
- ✓ *Downstepped*: D, descida suave.

Dos símbolos listados acima, três apresentam valores absolutos, para cada falante: T, B e M. Os demais – H, S, L, U e D – são relativos aos tons precedentes. O topo (T) e a base (B) correspondem aos pontos-alvo mais alto e mais baixo respectivamente, para um mesmo falante, representando assim sua tessitura. Juntamente com o médio (M), o T e o B são os únicos que podem ocupar o primeiro ponto-alvo do segmento estudado. O M é utilizado quando o primeiro ponto não é nem o T nem o B (LOUW e BARNARD, 2004; HIRST, 2005).

O símbolo S (igual), como o nome indica, é utilizado quando o ponto analisado não possui diferença relevante ao se comparar com o ponto precedente; em outras palavras, usa-se o S quando um ponto for muito próximo do anterior (LOUW e BARNARD, 2004; HIRST, 1999).

Os pontos-alvo mais altos que os precedentes, mas que não são T, podem ser classificados como H (mais alto) ou U (subida suave), e os pontos mais baixos que os anteriores se dividem em L (mais baixo) e D (descida suave) (HIRST e DI CRISTO, 1998).

A figura 2 mostra o resultado final dado pelos programas MOMEL/INTSINT:

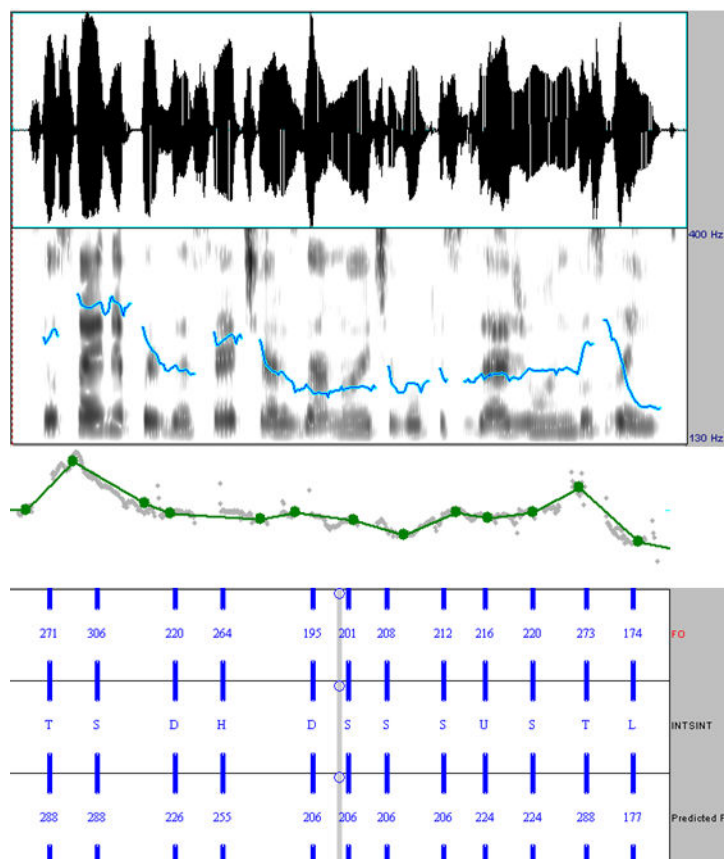


Figura 2: Oscilograma, curva de F0, pontos-alvo do MOMEL e codificação do INTSINT da frase “se eu sair depois das 7 horas da noite, eu vou precisar de um ônibus leito”.

Ressaltamos, neste momento, que os resultados apresentados pelo programa INTSINT representam uma ferramenta de análise fonológica apenas de superfície.

A utilização da análise semi-automática descrita acima foi utilizada na descrição entonacional de línguas como o francês (DI CRISTO, 2011) e o inglês britânico (AURAN, BOUZON e HIRST, 2004; BRIERLEY, 2011), e análise, via INTSINT manual, do ritmo no Português Brasileiro (GONÇALVES, 2000) e também automático (LUCENTE e BARBOSA, 2004; SILVA, MOUTINHO e TEIXEIRA, 2005; SILVA e SEARA, 2008; ARAUJO, 2012; ZENDRON Da CUNHA, 2012).

Para Lucente e Barbosa (2004), os resultados dados pelo programa INTSINT, apesar de parecerem ser uma boa alternativa para estudos entonativos, descrevem aspectos “não tão significativos do ponto de vista de uma notação entonacional” (LUCENTE e BARBOSA, 2004).

Em contrapartida, outros autores apresentaram uma visão mais positiva do programa em questão. Silva, Moutinho e Teixeira (2005), dentro do projeto Atlas Multimedia Prosodique de l’Espace Roman, apresentaram uma proposta de construção de um modelo entonativo do Português Brasileiro que seria aplicável a sintetizadores. Foi realizada demonstração de análise manual da curva melódica e também

automática, esta última por meio dos programas MOMEL/INTSINT. Para os autores, o resultado apresentado pelo INTSINT “foi capaz de extrair os movimentos essenciais da trajetória da frequência fundamental para os diversos tipos de frase” (SILVA, MOUTINHO e TEIXEIRA, 2005). Afirmaram, ainda, que a experiência de utilização do INTSINT foi eficaz e sugerem que trabalhos futuros considerem sua metodologia.

A fim de verificar o padrão entonacional de sentenças que não carregam foco sobre qualquer um de seus constituintes, Silva e Seara (2008) utilizaram o INTSINT e averiguaram que o evento tonal mais importante da frase recai sobre o verbo.

Também com propósito de correlacionar sintaxe e entonação, Zendron da Cunha (2012) utilizou o programa INTSINT a fim de buscar possíveis tendências em 624 frases produzidas por informantes do sexo feminino. Os resultados mostraram incidência de tons altos nos focos, sendo que os tons baixos são frequentemente relacionados ao sujeito e ao verbo.

Acrescentando aos estudos descritivos do português, pesquisas relacionadas à aplicabilidade em sistemas de conversão de voz e frequência podem ser encontradas (ODEBRECHT JÚNIOR e SEARA, 2009; ODEBRECHT JÚNIOR, 2012).

O presente estudo tem como objetivo verificar se o programa de análise semi-automática da entonação INTSINT é capaz de reproduzir tendências de variação melódica do português brasileiro. Dentro da nossa proposta, avaliamos a eficácia do programa durante a leitura na diferenciação das modalidades declarativa e interrogativa. A partir do exposto, foram levantadas as seguintes hipóteses: (I) A análise da entonação através do programa INTSINT permite escolhas de diferentes teorias fonológicas no que se refere ao uso de unidades entonativas, uma vez que seus resultados se restringem à análise fonológica de superfície; (II) E, uma vez que o programa codifica os principais pontos de variação da curva melódica, acreditamos que é possível distinguir as formas declarativa e interrogativa por meio dos pontos-alvo codificados para o português brasileiro. Foi realizada, ainda, uma análise descritiva exploratória da codificação realizada pelo MOEML-INTSINT. A justificativa da primeira hipótese tem como base a subdivisão em fonologia de superfície e fonologia subjacente, seguindo a proposta de Hirst, Di Cristo e Espesser (2000): as teorias fonológicas subjacentes relacionam forma e função, o que não é dado pelo INTSINT. Dessa forma, o INTSINT é visto como uma ferramenta, não como uma teoria.

2. Métodos

2.1 Seleção e descrição dos participantes

Participaram desta pesquisa 10 informantes, todos falantes nativos do português brasileiro, nascidos e criados na região metropolitana de Belo Horizonte, sexo feminino, com idades entre 20 e 30 anos, com mediana de 26. Todos estudavam na graduação ou pós-graduação da Faculdade de Letras da Universidade

Federal de Minas Gerais. Foram escolhidos apenas participantes do sexo feminino a fim de eliminar a variável sexo. Os participantes foram convidados a participar da pesquisa e assinaram o termo de consentimento livre e esclarecido antes de realizar as gravações. Todos os informantes se declararam bons leitores, sem qualquer queixa de problemas de fala, linguagem ou audição.

2.2. Informação técnica

O corpus deste estudo se constituiu de 10 pequenos textos retirados do corpus *EUROMI* adaptado ao Português Brasileiro (CELESTE, 2007). Cada texto apresenta aproximadamente cinco frases, as quais estão relacionadas entre si e estabelecem coerência textual. Os informantes receberam os textos e leram silenciosamente quantas vezes consideraram necessárias para que a leitura fluísse da forma mais natural possível.

As gravações foram realizadas dentro de uma cabine tratada acusticamente, por meio de um computador da marca Macintosh, modelo MacBook, 2GHz, 512 MB, com microfone interno com entrada de 16 bits de quantização e faixa de frequência de 20 até 20.000Hz. Os sujeitos leram cada passagem o número de vezes que foram necessárias para se sentissem confortáveis com a gravação. Porém, só foi salva no computador uma leitura de cada passagem por informante.

Para análise acústica dos dados foi utilizado o programa Praat, versão 4.4.27, disponível no site www.praat.org (BOERSMA e WEENINK, 1997), com as extensões necessárias para a aplicação dos programas MOMEL/INTSINT. Essas últimas encontram-se livremente disponíveis no site <http://aune.lpl.univ-aix.fr/~auran/english/ressources.html>.

Foi aplicado, então, o programa MOMEL e, feitas as correções manuais necessárias, aplicou-se o INTSINT. O resultado da aplicação do INTSINT é dado em um arquivo *textgrid*, contendo os valores previstos de F0 (frequência fundamental), a codificação do INTSINT e os valores finais de cada ponto (Fig. 3).

Após realizar a transcrição entonativa por meio do INTSINT, realizamos dois tipos de análises divididas em duas modalidades: afirmativa e interrogativa. Inicialmente fizemos uma análise qualitativa de todos os informantes e, em seguida, contrapusemos as duas unidades.

Essa divisão foi realizada a fim de verificar se o INTSINT nos daria algum padrão ou tendência para análise entonativa. Outro questionamento foi sobre a relação dos alvos codificados como pontos iterativos (U, D, H, L, S) e não-iterativos (T, B, M), e como os mesmos se relacionavam entre si.

2.3. Estatística

Para a análise estatística do INTSINT, questionamos: “há alguma homogeneidade na produção dos informantes?”. Os cálculos estatísticos de média, desvio padrão, nível de confiança (95%) e intervalo de confiança (95%) foram realizados através do programa Microsoft Excel® versão 2003.

Para os cálculos do intervalo de confiança, foi utilizada a técnica do *Bootstrap* não-paramétrico. O *Bootstrap* é um método de reamostragem de dados introduzido por Efron (1979) usado na inferência estatística quando há interesse, por exemplo, na construção de intervalos de confiança com uma amostra de tamanho n finito, em uma tentativa de aproximar a distribuição de uma determinada função das observações pela distribuição empírica de dados (RIBEIRO, 2006).

Uma vez que o *bootstrap* não-paramétrico não depende da distribuição que os dados seguem, como o presente estudo, Ribeiro (2006) afirma que o mesmo pode ser usado para qualquer conjunto de dados.

A fim de contribuir para a descrição dos valores dados pelo programa INTSINT, foi aplicado o teste t de Student, com nível de significância de 5%, na comparação entre os valores de F0 originais e os preditos pelo programa.

3. Resultados

Os resultados serão apresentados e discutidos inicialmente com as modalidades separadas, primeiro a frase declarativa e logo após a frase interrogativa e, em seguida, far-se-á uma comparação entre as duas.

A tendência de queda em final de frase declarativa (FERNANDES, 1976; CAGLIARI, 1981; MACIEL e ROTHE-NEVES, 2007) foi evidente, com uma configuração clara na curva de F0 que foi captada pelo MOMEL e codificada pelo INTSINT. Para caracterizar tal fenômeno, os símbolos finais deveriam ser D, L ou B, seguidos ou não de S. No quadro 1, é possível visualizar a codificação do INTSINT para a frase “Aqui é do corpo de bombeiros”, revelando essa queda final na fala de todas as informantes.

A característica ascendente para interrogativas é questionada em pesquisas que envolvem essa modalidade no Português Brasileiro (MORAES, 1984; MATA, 1992; REIS, 1995), sendo que no presente vemos que não há subida no último ponto-alvo codificado em 90% das frases interrogativas, como mostra o quadro 1:

Informante	INTSINT - declarativa	INTSINT – interrogativa
1	MTDTB	TSLLSUHB
2	MTDTDTB	MHUTMSDSHLS
3	MUDUSUBS	MHLDSHL
4	MTSLBHL	MTLSDTL
5	TDTLD	TDLDTB
6	BHTDB	MUDLDHB
7	MHSSL	TDTDDDDTB
8	MTDUSUDHDB	MUHLLULHD
9	MHDUB	MUDDSSHSD
10	MHDUDUB	MTSLLTL

Quadro 1: Codificação do INTSINT para as duas modalidades, declarativa e interrogativa, de todas as informantes.

Se dividirmos os resultados encontrados em subida, descritos por T, H e U, e em descida, sob notação dos símbolos B, L e D, vemos que apenas 10% dos casos pertencem ao primeiro grupo, ascendente, e 90% ao segundo grupo, descendente. Ambas as modalidades parecem apresentar um início alto, de forma geral, se não no primeiro ponto-alvo, codificado por T, no segundo, que tem por início um M, mas é seguido de H, T ou U.

Nas frases interrogativas, vemos que há, antes da queda final, uma subida acentuada representada pelos símbolos H e T. Se considerarmos o penúltimo ponto-alvo codificado como coincidente ou próximo à última tônica proeminente da interrogativa, pode-se observar sim uma subida final relevante. Segundo Moraes e Abraçado (2005), as interrogativas totais apresentam tons altos logo no início do enunciado. Em seguida, o movimento descende e ascende, com pico na última tônica da sentença. Esse movimento melódico visualizado acusticamente nas interrogativas também foi descrito por Cemin, Andrade e Nunes (2011).

Nas declarativas é possível visualizar essa subida seguida de descida, mas não tanto e nem tão sistematicamente quanto nas frases interrogativas. Para Fernandes (1976), as frases declarativas são marcadas por um movimento descendente final, mais especificamente na última sílaba tônica. Nos resultados dados pelo INTSINT nesta pesquisa, parte dessa descrição pode ser observada pelo final claramente mais baixo, apontados pelos símbolos B, L e D (seguidos ou não de S).

Padrões similares foram encontrados por Cemin, Andrade e Nunes (2011) em enunciados declarativos. As autoras verificaram que, na região medial dos enunciados, há mudança de foco, marcados

ora por tons mais baixos, como o B, ora por tons mais altos, como o H, e tal alteração representa desambiguação de sentenças semanticamente ambíguas.

Ao tentar distinguir as modalidades aqui estudadas via codificação realizada pelo INTSINT, vemos que o início alto foi comum tanto na forma declarativa quanto na forma interrogativa, fato representado pelo primeiro ponto alvo, codificado como T, ou segundo ponto alvo, codificado por M, mas seguido de H, T ou U. Porém, as modalidades se diferem no final da unidade: nas frases interrogativas, uma subida acentuada representada pelos símbolos T e H pode ser observada logo antes da queda final. Já nas frases declarativas vemos, sim, uma subida seguida de descida, mas não tanto e nem tão sistematicamente quanto nas frases interrogativas. Dessa forma, a segunda hipótese proposta no presente estudo foi confirmada.

Realizamos, ainda, uma análise descritiva exploratória da codificação realizada pelos programas MOMEL e INTSINT, utilizando como base a F0 original e a prevista pelo INTSINT, bem como a análise da codificação dos pontos não-iterativos.

Quanto aos valores absolutos das duas modalidades, foi possível observar que as variações maiores do código T não se mantiveram como na análise de uma frase com duas unidades entonativas.

Ainda comparando as modalidades, realizamos uma análise global dos pontos não-iterativos. Para tanto, separamos três frases: uma afirmativa, contendo uma unidade entonativa (para esta divisão de unidade entonativa consideramos as pausas, as grandes variações melódicas e alongamento da vogal tônica), uma frase afirmativa contendo duas unidades entonativas e uma frase interrogativa.

Ao analisarmos os pontos-alvo codificados como não-iterativos, vemos que, para a frase declarativa “Aqui é do corpo de bombeiros”, as maiores diferenças entre a F0 do sinal original e a frequência prevista pelo INTSINT foram concentradas no símbolo M (tabela 1), com diferenças de até 18 Hz, diferentemente do encontrado até então.

Tabela 1: Pontos não-iterativos, seus valores correspondentes ao sinal original e ao cálculo do INTSINT e a diferença entre eles da unidade entonativa “Aqui é do corpo de bombeiros”, em Hertz (Hz).

SI	F0 or (Hz)	F0 pr (Hz)	Dif F0 or e pr (Hz)
T	278	273	5
	268	273	-5
M	220	222	-2
B	174	180	-6
T	254	251	3
	252	251	1
	247	251	-4
M	214	204	10
B	165	166	-1
M	214	218	-4
B	158	154	4

Tabela 1 (cont.): Pontos não-iterativos, seus valores correspondentes ao sinal original e ao cálculo do INTSINT e a diferença entre eles da unidade entonativa “Aqui é do corpo de bombeiros”, em Hertz (Hz).

SI	F0 or (Hz)	F0 pr (Hz)	Dif F0 or e pr (Hz)
T	270	281	-11
M	212	228	-16
B	194	185	9
T	305	307	-2
	318	307	11
T	330	333	-3
B	210	205	5
	206	205	1
T	300	301	-1
M	198	192	6
T	306	293	13
M	205	207	-2
B	153	146	7
M	218	224	-6
B	160	158	2
M	222	240	-18
B	180	176	4

Legenda

SI: símbolo do INTSINT

F0 or: valor de F0 referente ao sinal original

F0 pr: valor de F0 predito pelo INTSINT

Dif F0 or e pr: diferença entre os valores de F0 referente ao sinal original e o valor predito pelo INTSINT

Ao comparar os resultados da F0 original e o valor predito pelo INTSINT, não foi encontrada diferença estatisticamente significativa, com $p=0,5$.

Para a frase “Se eu sair depois das sete horas da noite, eu vou precisar de um ônibus leito”, a diferença da F0 do sinal original e da frequência predita pelo INTSINT tinha valores maiores para o símbolo T. Mas, para a frase “Aqui é do corpo de bombeiros”, os valores de diferença para o T são os menores, como mostra o gráfico 1:

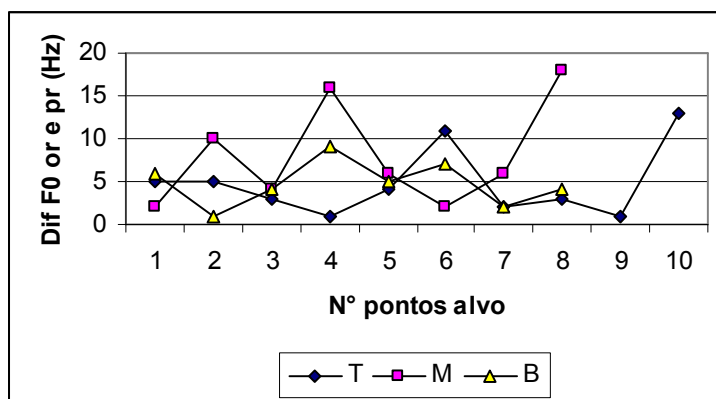


Gráfico 1: Diferenças em Hertz (Hz) entre o valor de F0 do sinal original e o valor predito pelo INTSINT para os símbolos não-iterativos na unidade entonativa “Aqui é do corpo de bombeiros”.

Apesar da irregularidade dos pontos na linha do gráfico 1, a questão foi aqui analisada com apoio de um teste estatístico. Assim, sempre que apresentamos a tabela com as variações entre o sinal original e o predito pelo programa INTSINT, complementamos com a análise estatística.

Quando analisamos a frase “Você faz reserva por telefone?” também encontramos os maiores valores de diferença entre F0 do sinal original e da frequência predita pelo INTSINT relativos aos pontos-alvo codificados por M, como mostram a tabela 2 e o gráfico 2. O gráfico 2 também nos mostra que o símbolo não-iterativo com o maior número de codificações foi o T, com 11 pontos-alvo, seguido pelo M e pelo B, com oito e quatro, respectivamente.

Tabela 2: Pontos não-iterativos, seus valores correspondentes ao sinal original e ao cálculo do INTSINT e a diferença entre eles da unidade entonativa “Aqui é do corpo de bombeiros”, em Hertz (Hz).

Informante	SI	F0 or	F0 pr	Dif F0 or e pr
1	T	320	324	-4
	B	167	151	16
2	T	329	325	4
	M	201	193	8
3	M	175	193	-18
	M	257	230	27
4	T	267	271	-4
		275	271	4
	M	227	213	14
5	T	268	264	4
		253	264	-11
	B	162	162	0

Tabela 2 (cont.): Pontos não-iterativos, seus valores correspondentes ao sinal original e ao cálculo do INTSINT e a diferença entre eles da unidade entonativa “Aqui é do corpo de bombeiros”, em Hertz (Hz).

Informante	SI	F0 or	F0 pr	Dif F0 or e pr
6	M	206	213	-7
	B	127	122	5
7	T	232	233	-1
		244	233	11
		231	233	-2
	B	155	154	1
8	M	227	225	2
9	M	234	242	-8
10	T	255	260	-5
		273	260	13
	M	203	204	-1

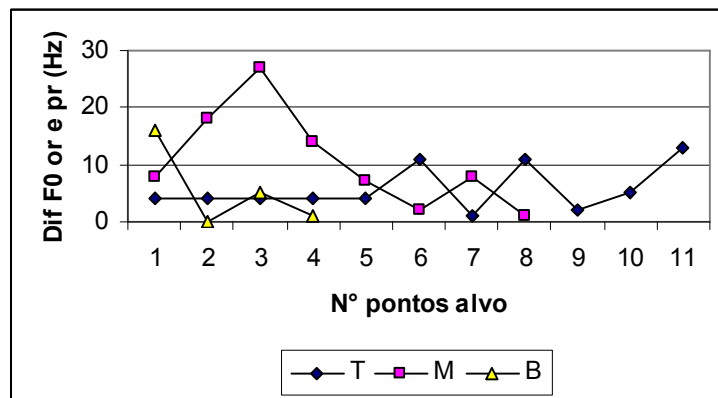


Gráfico 2: Diferenças em Hertz (Hz) entre o valor de F0 do sinal original e o valor previsto pelo INTSINT para os símbolos não-iterativos na unidade entonativa “Aqui é do corpo de bombeiros”.

Ao comparar os resultados da F0 original e o valor previsto pelo INTSINT, não foi encontrada diferença estatisticamente significativa, com $p=0,2$.

3.1. Análise dos pontos não-iterativos

Após analisarmos as informantes de forma individual, reunimos as informações de todos os valores de T, M e B encontrados para todas as informantes nos três grupos analisados referentes à F0 do sinal original, à frequência prevista pelo INTSINT e à diferença entre eles.

Os valores de T apresentaram variação significativa, com mínimo de 212 Hz e máximo de 363 Hz, como pode ser visualizado na tabela 3:

Tabela 3: Média de F0 referente ao sinal original, ao predito pelo INTSINT, à diferença entre eles e ao valor de p ($p < 0,05$).

	F0		Dif F0	
	or	pr	or e pr	p
T	269	271	5	0,4
M	214	218	7	0,2
B	168	166	4	0,3

Legenda

F0 or: valor de F0 referente ao sinal original

F0 pr: valor de frequência predito pelo INTSINT

Dif F0 or e pr: diferença entre o valor de F0 referente ao sinal original e o predito pelo INTSINT

Em todas as comparações entre valores originais e preditos pelo INTSINT realizadas neste estudo, não há diferença estatisticamente significativa. Dessa forma, um possível desvio dos valores de F0 durante a aplicação do programa não deve ser preocupação dos pesquisadores.

Para os pontos-alvo codificados como T, a média e o intervalo de confiança estão dispostos no gráfico 3:

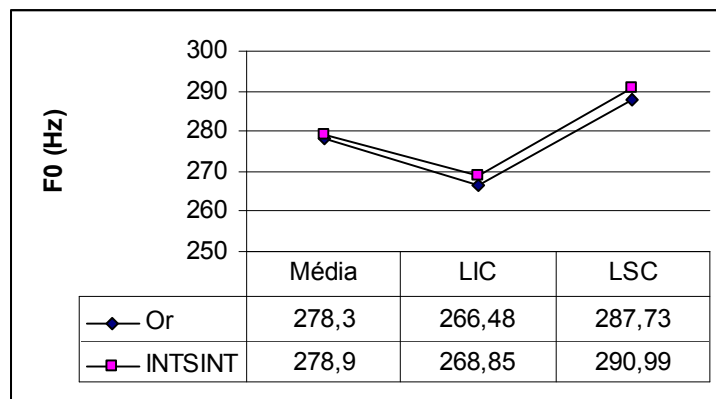


Gráfico 3: Média e intervalo de confiança para a F0 do sinal original e para a frequência predita pelo INTSINT referente ao T.

Legenda:

Or: sinal original

LIC: limite inferior de confiança

LSC: limite superior de confiança

Como é possível observar pelo gráfico acima, a diferença entre o sinal original e o INTSINT não chega a 3 Hz no que tange a valores relacionados à média. Esse é um bom indicador que aponta a codificação do INTSINT com valores de F0 praticamente idênticos aos valores do sinal original.

Mas, apesar de todas as informantes serem do sexo feminino, houve grande variação nos resultados, com desvio padrão de 40,02 Hz para o sinal original e 37,71 Hz para o predito pelo INTSINT. Nesse momento, fica clara a variação individual devido ao alto grau de dispersão dos valores que representam o ponto mais alto da tessitura de cada informante.

Foram codificados 24 pontos-alvo como M, com mínimo de 175 Hz e máximo de 295 Hz para o sinal original e mínimo de 178 Hz e máximo de 296 Hz para o INTSINT.

Quanto aos valores que se referem à média, os resultados do sinal original e do INTSINT encontram-se próximos (gráfico 4):

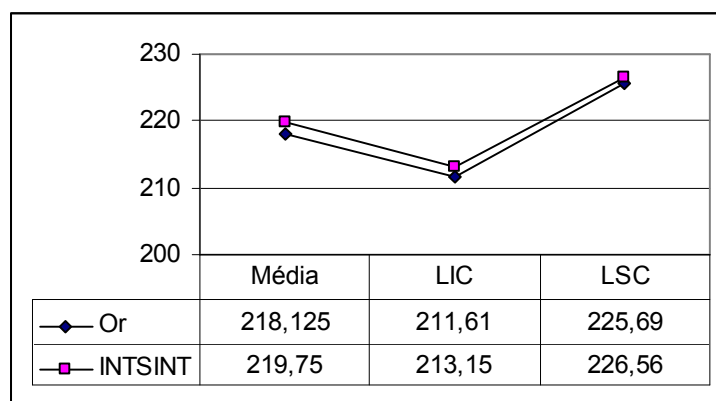


Gráfico 4: Média e intervalo de confiança para a F0 do sinal original e para a frequência predita pelo INTSINT referente ao M.

Legenda

F0 or: valor de F0 referente ao sinal original

F0 pr: valor de frequência predito pelo INTSINT

Dif F0 or e pr: diferença entre o valor de F0 referente ao sinal original e o predito pelo INTSINT

O gráfico 4 revela os valores relativos à média a maior variação entre o som original e o INTSINT é de 1,6 Hz. Mais uma vez é possível observar que os cálculos realizados com base no sinal original para estilizar a curva nos pontos-alvo do MOMEL e codificá-los através do INTSINT são confiáveis.

Porém, os valores encontrados para o desvio padrão (24,71 Hz para o sinal original e 24,50 Hz para o INTSINT) mostram que o grau de dispersão dos valores de F0 codificados como M são consideráveis.

O desvio padrão encontrado para os valores de B foi menor: 19,54 Hz e 19,17 Hz para o sinal original e INTSINT, respectivamente. Tal resultado mostra que a variação entre as frequências mais baixa e mais alta encontradas para B não tiveram grande variação (tabela 3).

Foram 28 pontos-alvo do MOMEL codificados como B, e os valores de frequência do sinal original estão bem próximos aos valores do INTSINT; chegaram a obter exatamente o mesmo valor (162 Hz). Os valores referentes à média encontram-se próximos (gráfico 5).

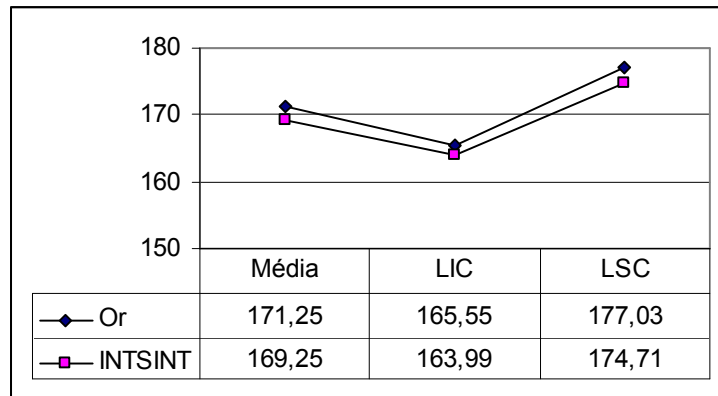


Gráfico 5: Média e intervalo de confiança para a F0 do sinal original e para a frequência predita pelo INTSINT referente ao B.

Ao compararmos os três códigos, T, M e B, vemos que o símbolo mais codificado foi o T, como mostra o gráfico 6:

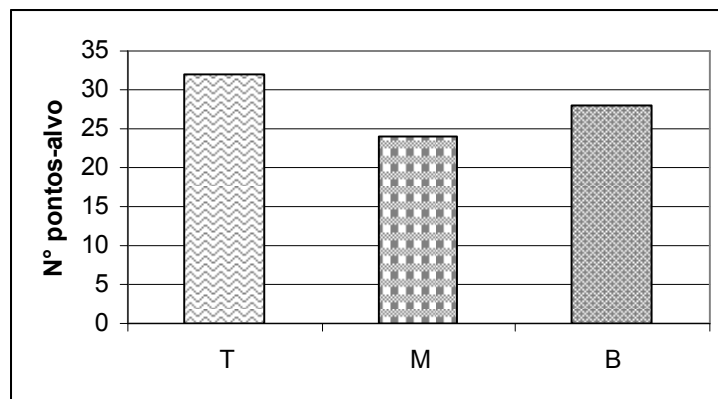


Gráfico 6: N° de pontos-alvo codificados como T, M e B.

Outra observação realizada na comparação dos três códigos foi que a média da diferença entre a frequência do sinal original e da predição do INTSINT foi menor para o código B, como mostra o gráfico 7:

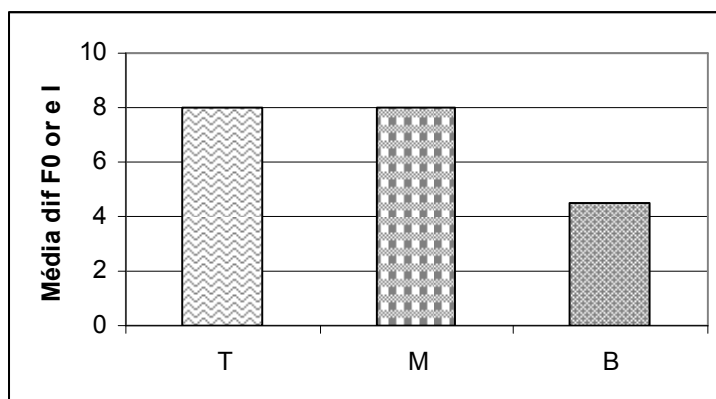


Gráfico 7: Média da diferença entre o sinal original e o INTSINT nos códigos T, M e B.

Legenda:

Dif F0 or e I: Diferença da F0 do sinal original e do INTSINT

A diferença evidenciada no gráfico acima nos leva a questionar se os programas de análise entonativa MOMEL e INTSINT lidam de forma mais apropriada com dados com amplitude de frequência mais baixa. Estudos que envolvessem informantes do sexo masculino forneceriam dados importantes sobre tal questão.

Retomando as hipóteses levantadas para o presente estudo, pode-se observar nos resultados que, apesar da variabilidade encontrada individualmente, os registros propostos pelo INTSINT no nível fonológico de superfície permitem uma análise que relaciona forma e função com suporte de diferentes teorias, uma vez que há separação clara entre forma e função. Os estudos aqui descritos que têm como base a relação prosódia e sintaxe são um exemplo. Outro exemplo seria a própria utilização, nesta pesquisa, das unidades entonativas baseadas em Halliday (1970) e Pierrehumbert (1987).

Existem outros pontos que deverão ser abordados em próximos estudos a fim de ampliar e melhor entender os níveis necessários para um estudo completo da entonação e como tais níveis se relacionam e são estruturados mentalmente pelos falantes. Estudos que utilizam modelos semi-automáticos para o estudo da entonação, como o INTSINT, têm sido desenvolvidos (RODERO, 2011).

Outros estudos que provavelmente esclareceriam muitas questões sobre o funcionamento da prosódia e seus níveis de análise são aqueles que tratam especificamente da leitura, já que perceptivamente os falantes fazem uma distinção natural entre fala espontânea e leitura.

4. Considerações finais

As nossas primeiras constatações sobre a utilização do INTSINT referem-se à dificuldade encontrada de estabelecer quais unidades utilizaríamos para a aplicação. Sabendo que o INTSINT é aplicado sobre os pontos-alvo do MOMEL e que este, por sua vez, divide os dados com base no nível físico (pausas

superiores a 250ms), restou-nos o impasse de definir a unidade mais apropriada para avaliarmos o desempenho do programa.

Procedemos, então, com base em Halliday (1970) e Pierrehumbert (1987), à divisão das unidades entonativas, e a partir delas buscamos analisar de forma consistente as tendências que o INTSINT permitiria investigar. Para suprir essas dificuldades, selecionamos três frases: uma frase com duas unidades entonativas, uma frase declarativa com apenas uma unidade entonativa e uma frase interrogativa que também continha apenas uma unidade entonativa.

Nesta etapa vimos um limite previsto, mas importante, do programa INTSINT. Ele não fornece qualquer unidade de análise, portanto, o desenvolvimento do estudo depende de uma teoria prosódica.

Tal fato, que, a princípio, nos parece uma limitação, é também uma vantagem do programa. Com os resultados do INTSINT em mãos, o pesquisador tem a liberdade de escolher qual teoria de análise prosódica utilizará como base para seus estudos. Isso quer dizer que, tendo o nível fonológico de superfície, a representação do nível fonológico subjacente não ficará restrita ou obrigatoriamente relacionada ao nível anterior.

Para as modalidades declarativas e interrogativas, o início alto foi comum em ambos, fato representado pelo primeiro ponto-alvo, codificado como T, ou segundo ponto-alvo, codificado por M, mas seguido de H, T ou U. No entanto, elas diferem no final da unidade: nas frases interrogativas, uma subida acentuada representada pelos símbolos T e H pode ser observada logo antes da queda final. Já nas frases declarativas vemos, sim, uma subida seguida de descida, mas não tanto e nem tão sistematicamente quanto nas frases interrogativas.

Longe de encerrar as possibilidades de investigação nas unidades entonativas e nas modalidades escolhidas para este estudo, chegamos ao nosso objetivo inicial de verificar o comportamento do programa INTSINT frente a variações entonativas bem delimitadas. Percebemos que o programa INTSINT é capaz de transmitir tendências e também o padrão melódico geral da curva de F0.

5. Agradecimentos

Agradecemos ao Cnpq, que financiou parte da pesquisa aqui apresentada.

6. Referências

Araujo FM. A entonação de sentenças clivadas e pseudo-clivadas no Português Brasileiro [Dissertação de mestrado]. Universidade Federal de Santa Catarina; 2012.

Auran C, Bouzon C, Hirst D. The Aix-MARSEC project: an evolutive database of spoken British-English. IN: Proceedings of I Speech Prosody on Japan; 2004. Japan, March 23-26, 2004.

Boersma P, Weenink D. Praat: doing phonetics by computer. 1992-2012. Disponível em: www.praat.org

- Brierley C. Prosody Resources and Symbolic prosodic features for automated phrase break prediction [thesis]. University of Leeds (School of computing); 2011.
- Cagliari LC. O sistema entonacional do português do Brasil. In: Elementos de fonética do português brasileiro. Campinas: Unicamp; 1981.
- Cemin J, Andrade MEG, Nunes VG. Uma contribuição da prosódia na desambiguação de sentenças semanticamente ambíguas. *Work. Pap. Linguíst.*, 2011 12(2): 69-94.
- Di Cristo A. Une approche intégrative des relations de l'accentuation au phrasé prosodique du français. *Journal of French Language Studies*; 2011 21:73-95.
- Celeste LC, MOMEL e INTSINT: uma contribuição à metodologia do estudo prosódico do Português Brasileiro (dissertação de mestrado). Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2007.
- Efron B. Bootstrap methods: another look at jackknife. *Ann. Statist.* 1979 7(1): 1-26.
- Fernandes NH. Contribuição para uma análise instrumental da acentuação e intonação do português. (Dissertação de mestrado). São Paulo: Universidade de São Paulo; 1976.
- Gonçalves CAV. Transcrevendo a entonação. *Veredas: revista de estudos linguísticos*; 2000 3(2): 9-19.
- Gussenhove C. *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press; 2004.
- Halliday MAK, Hassan R. *Cohesion in English*. London: Longman; 1976.
- Hart TJ, Collier R. Integrating Different Levels of Intonation Analysis. *Journal of Phonetics*; 1975 3: 235-255.
- Hart TJ, Collier R, Cohen A. *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press; 1990.
- Hirst, D. The analysis by synthesis of speech melody: from data to models. *Journal of Speech Sciences*; 2011 1(1):55-83.
- Hirst D. *Form and Function in the Representation of Speech Prosody*. Aix-en-Provence: Université de Provence; 2005.
- Hirst D. Automatic analysis of prosody for multilingual speech corpora. In: Keller, E. Bailly, G. Terken, J. e Huckvale, M. (eds) *Improvements in Speech Synthesis*. Chichester, UK: Wiley; 2001 p. 1-9. Disponível em: <http://aune.lpl.univ-aix.fr/~hirst/articles/2001%20hirst.pdf>.
- Hirst, D. A symbolic coding of segmental duration and tonal alignment: an extension to the Intsint system. In: *Proceeding of Sixth European Conference on Speech Communication and Technology Eurospeech on Budapest*; 1999 Sept 5-9; Hungary. Budapest; 1999 p. 1639-42.
- Hirst D, Di Cristo AE, Espesser R. Levels of representation and levels of analysis for intonation. In: Horne, M. (ed) *Prosody : Theory and Experiment*. Dordrecht: Kluwer Academic Press; 2000.
- Hirst D, Di Cristo A. A survey of intonation systems. IN: Hirst e Di Cristo (eds). *Intonation Systems : A Survey of Twenty Languages*. Cambridge: Cambridge University Press; 1998.
- Ladd RD. *Intonation phonology*. Cambridge: Cambridge University Press; 1996.
- Lou J, Barnard E. Automatic intonation modeling with INTSINT. In: *Proceeding of the 15 th Annual Symposium of the Pattern Recognition Association of South Africa*. Grabow, November; 2004.
- Lucente L, Barbosa PA. Estudo-Piloto de uma notação entoacional para o português brasileiro: ToBI or not ToBI? In: *Anais do 6º encontro CELSUL - Círculo de Estudos Lingüísticos do Sul*. Florianópolis:

UFSC, 2004. Disponível em: <<http://www.celsul.org.br/Encontros/06/Coordenadas/27.pdf>>. Acesso em 20 de setembro de 2012.

Maciel FJ, Rothe-Neves R. Investigações experimentais da entonação no português brasileiro: um revisão de literatura. In: Anais do IX Congresso Nacional de Fonética e Fonologia e III Congresso Internacional de Fonética e Fonologia. Brasil: Belo Horizonte; 2007.

Mata AI. A questão da entonação na interrogação em português: “Isso é uma pergunta?”. In: Pereira, I. Estudos em prosódia. Coleção de estudos linguísticos, 2. Lisboa: Colibri; 1992.

Moraes JA. Recherchers sur l’intonation modale du portugais brésilien parlé à Rio de Janeiro. (Thèse de doctorat de troisième cycle). Université de la Sorbonne Nouvelle, Paris III; 1984.

Moraes JA, Abraçado J. A descrição prosódica do Português do Brasil no AMPER. IN: Géolinguistique 3/2005 (Projet AMPER), 2005; 337-345.

Odebrecht Jr M, Seara IC. Conversão de Contorno de Pitch para Aplicação em Sistemas de Conversão de Voz. Anais do XXVII Simpósio Brasileiro de Telecomunicações (SBrT 2009) pp. 1-6, setembro; 2009.

Odebrecht Jr M. Conversão do contorno de pitch por divisão de componentes para aplicação em sistemas de conversão de voz (Dissertação de mestrado). Universidade Federal de Santa Catarina; 2012. Disponível em: <http://repositorio.ufsc.br/xmlui/handle/123456789/9457>.

Pierrehumbert JB. The phonology and phonetics of the English intonation (PhD thesis). Indiana University Linguistics Club; 1987.

Reis C. L’interaction entre l’accent, l’intonation et le rythme en portugais brésilien (Thèse de doctorat). Université de Provence; 1995

Ribeiro JA. Inferência sobre os hiperparâmetros dos modelos estruturais usando Bootstrap (Dissertação de mestrado). Universidade Federal de Minas Gerais; 2006

Rodero E. Intonation and Emotion: Influence of Pitch Levels and Contour Type on Creating Emotions. Journal of Voice; 2001 January 25(1):25-34.

Silva MCF, Seara IC. Mais sobre a entonação de sentenças com ordem SV. Revista letras; 2008 maio/dez 75-76:171-181.

Vaz da Silva AMC, Moutinho LC, Teixeira A. Contributos para um Modelo Prosódico do Português. IN: Duarte, I. Leiria, I. (org.). Actas do XX Encontro da Associação Portuguesa de Linguística, Lisboa: APL/Colibri; 2005: 831-842.

Zendron da Cunha K. Sentenças exclamativas em português brasileiro: padrão entonacional e sintaxe [Dissertação de mestrado]. Universidade Federal do Paraná; 2012.