UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS CLÁSSICAS E VERNÁCULAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOLOGIA E LÍNGUA PORTUGUESA

RENATA CEZAR DE MORAES ROSA

Percepção e ExProsodia®: correlação entre análise automática e a finalização de frases assertivas isoladas do Português Brasileiro

(versão corrigida)

São Paulo
2015

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS CLÁSSICAS E VERNÁCULAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOLOGIA E LÍNGUA PORTUGUESA

Percepção e ExProsodia®: correlação entre análise automática e a finalização de frases assertivas isoladas do Português Brasileiro

(versão corrigida)

Renata Cezar de Moraes Rosa

Dissertação apresentada ao Programa de Pós-Graduação de Filologia e Língua Portuguesa do Departamento de Letras Clássicas e Vernáculas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para obtenção do título de mestre em Língua Portuguesa.

Orientador: Prof. Dr. Waldemar Ferreira Netto
(X) DE ACORDO

São Paulo
2015

ROSA, Renata Cezar de Moraes. **Percepção e ExProsodia®: correlação entre análise automática e a finalização de frases assertivas isoladas do Português Brasileiro**. Dissertação apresentada à Faculdade de Filosofia, Letras e Ciências Humanas para obtenção de título de Mestre em Língua Portuguesa.

Aprovado em: 21/08/2015

Banca Examinadora:

Profa. Dra. Maressa de Freitas Vieira

Instituição: Instituto Federal de Ciências e Tecnologia de São Paulo - IFSP

Profa. Dra. Rosane de Sá Amado

Instituição: Universidade de São Paulo - USP

Aos que me deram a vida: Heloísa e Maurício!
Aos que a abrilhantam: Inaiá e Anderson!
Àquele que a complementa: Wagner!

Agradecimentos

Primeiramente agradeço a Deus por me permitir alcançar uma etapa importante da minha vida.

Agradeço ao Maurício, ser humano brilhante e inteligentíssimo, que plantou em mim a sede pelo conhecimento, a gana pelo saber, o amor pela pesquisa.

Agradeço à Heloísa, guerreira, que travou grandes batalhas para que eu e meus irmãos, por meio da educação, pudéssemos chegar onde quiséssemos.

Agradeço ao Anderson, pelos longos e maravilhosos debates, pela alegria fraternalmente dividida, pelas lições de vida.

Agradeço à Inaiá, pela presença marcante em minha vida, pela cumplicidade, pela doçura e agradeço ao Rogério, pela torcida, pelo apoio.

Agradeço à Luli, pela companhia e amizade desde os tempos pré-universidade.

Agradeço ao Wagner, pelo apoio incondicional, pela dedicação, pela vida compartilhada, pela presença constante, pelo amor.

Agradeço imensamente ao meu orientador, Professor Dr. Waldemar Ferreira Netto, pela oportunidade e pelos ensinamentos. Um verdadeiro mestre sempre disposto a ensinar e compartilhar conosco o apreço pela pesquisa.

Agradeço à Fernanda Consoni, "minha coorientadora," pela paciência e parceira em diversos estudos.

Agradeço ao Daniel Oliveira Peres, pela ajuda inestimável nas leituras de meus trabalhos, pelas dicas, pelas lições valiosas de linguística, pela gentileza.

Agradeço ao Marcus Vinícius Moreira Martins, pela enriquecedora convivência, pelas “aulas de estatística”.

Agradeço à Maressa de Freitas Vieira, pela cortesia e dedicação com a qual leu minha pesquisa e a revisou.

Agradeço a todos os colegas da “Turma do Waldemar.” Muito me honra pertencer ao grupo de pesquisa ExProsodia.

Agradeço ao Professor Dr. Paulo Chagas de Souza, pelas esclarecedoras aulas de fonologia, pelas leituras de meus trabalhos, pelos apontamentos em minha qualificação.

Agradeço à Professora Dra. Rosane de Sá Amado, pela leitura atenta do meu relatório, pelas valorosas observações.

Agradeço o apoio institucional da Universidade de São Paulo.

Agradeço a todos que me incentivaram e que acreditam em mim. Obrigada!

RESUMO

A entoação pode ser decomposta em componentes estruturadoras (finalização e sustentação) e, também, semântico-funcionais (foco/ênfase e acento lexical). A proposta assumida pelo grupo de pesquisa ExProsodia, baseada nos estudos de Xu e Wang (1997), permite a análise isolada de cada uma das componentes de F0. Neste estudo tratamos das componentes estruturadoras, especificamente, do estudo da finalização (F). O trabalho tem como objetivo, primeiro, analisar por meio de testes de percepção o que os sujeitos falantes do português brasileiro (doravante PB) julgam ser um final de frase assertiva. Depois, é nosso propósito analisar se os dados perceptivos fornecidos pelos ouvintes coincidem com aqueles disponibilizados pela análise automática obtida por meio da rotina ExProsodia®. É nosso intento, ainda, verificar se, no que tange à percepção dos procedimentos de segmentação frasal com propósito de finalização, é observada a sequência finalizadora proposta por Ferreira Netto (2008), acrescida da hipótese de Consoni *et al* (2009) e Consoni (2011). Este estudo visa integrar a coletânea de trabalhos desenvolvidos pelo Grupo de pesquisa ExProsodia - Análise automática da entoação na fala de língua portuguesa, linha de pesquisa na área de Percepção de Prosódia sob a orientação do Prof. Dr. Waldemar Ferreira Netto. Elaboramos dois testes. No primeiro teste foram apresentadas nove frases (três frases sem manipulação e seis manipuladas). As frases não manipuladas são, ao longo deste trabalho, mencionadas como *frases modelos.* As seis frases manipuladas foram subdivididas em dois conjuntos de três frases cada: a de frases monotonais e as frases manipuladas de modo que entre o tom médio (TM) e o tom de finalização (TF) houvesse uma distância de 7st. Nomeamos as primeiras frases como *frases neutras* e as segundas, *frases hipóteses*. No segundo teste foram apresentadas seis frases (somente as frases manipuladas, ou seja, as *frases neutras* e as *frases hipóteses*). Cada vez que o participante ouvia uma frase era solicitado a este que indicasse se a frase, de acordo com a sua percepção, estava ou não finalizada. Foram dadas duas opções de resposta aos ouvintes: **sim, a frase está finalizada** ou **não, a frase não está finalizada**. Os testes foram aplicados em 20 ouvintes de ambos os sexos, com idade entre 18 e 55 anos, todos com escolaridade entre nível médio e nível superior. Os resultados do teste 1 indiciam que as porcentagens dos ouvintes que consideram as *frases neutras* como não finalizadas sempre foram altas e mantiveram-se, em sua grande maioria, na casa dos 80%. Ademais, o produto da análise das *frases hipóteses* corrobora os

dados dos trabalhos já citados. (CONSONI; FERREIRA NETTO, 2008; CONSONI *et al.*, 2008; CONSONI *et al.*, 2009 e CONSONI, 2011; FERREIRA NETTO, 2006, 2008; ROSA 2009). A análise do teste 2 fora efetuada por meio da teoria da detecção de sinais (TDS) (GREEN; SWETS, 1966; MACMILLAN; CREELMAN, 2005; ADBI, 2007). Ao fazer uso da TDS, pretendemos estimar a força da sensibilidade perceptiva de cada participante do teste 2. O valor dessa força é dada pelo cálculo do *dee-prime*. Os resultados do teste 2 apontaram *dee-primes* de três categorias diferentes, a saber: alto; médio e baixo. As respostas da grande maioria dos participantes situaram-se entre uma facilidade extrema e média de discriminar as frases finalizadas e as não finalizadas.

Palavras-chave: percepção, prosódia, entoação; fonética

ABSTRACT

The intonation can be decomposed into structuring components (finalization and support component) and also semantic-functional components (focus / emphasis and lexical stress). The proposal assumed by ExProsodia rechearch group, based on the studies of Xu and Wang (1997), it allows separate analysis of each component of F0. In this study we treat structuring components, specifically, the study of finalization (F). The study aims, first, to analyze through perception tests which Brazilian Portuguese (henceforth PB) speakers think that is the ending of declarative sentences. It is our intent also verify if the finalization sequence proposed by Ferreira Netto (2008), plus the assumption of Consoni *et al.* (2009) e Consoni (2011) is observed. And lastly, our purpose is to inspect whether perceptual data match with those provided by automatic analysis obtained by routine ExProsodia®. This study aims to integrate a collection of work done by the research group ExProsodia - Automatic analysis of intonation in the speech of the Portuguese language, research line in the area of Perception of Prosody under the guidance of Prof. Dr. Waldemar Ferreira Netto. We developed two tests. In the first test, nine sentences were presented to the Brazilian Portuguese listeners (three sentences without manipulation and six manipulated). Non-manipulated sentences, throughout this work, as mentioned as *models sentences*. The six manipulated sentences were subdivided into two groups of three sentences each one: the monotonous sentences and manipulated sentences. The last one is manipulated so that between the medium tone (MT) and the finalization tone (FT) there is a distance of 7 semitones. We named first sentences as *neutral sentences* and the others *as hypothesis sentences*. At the second test, six sentences (only manipulated sentences, *i.e. neutral sentences* and *hypothesis sentences*) were presented to

Brazilian Portuguese listeners. Each time the participant heard a sentence, he was requested to indicate whether that sentence, according to their perception, or was not finalized. Two response options were given to listeners: **yes, the sentence is finished** or **not, the sentence is not finished**. 20 listeners of both sexes, aged between 18 and 55 years, all with education among middle and senior level have participated of the test. The test results indicate that percentages of listeners who consider *neutral sentences* as unfinished sentences were always high and remained, mostly at around 80%. Moreover, the product of the analysis of *hypothesis sentences* corroborates data of the work already cited. (CONSONI; FERREIRA NETTO, 2008; CONSONI *et al.*, 2008; CONSONI *et al.*, 2009 e CONSONI, 2011; FERREIRA NETTO, 2006, 2008; ROSA 2009). The analysis of the test 2 has carried out by the signal detection theory (SDT) (GREEN; SWETS, 1966; MACMILLAN; Creelman, 2005; ADBI, 2007). By making use of the TDS, we intend to estimate the strength of the perceptual sensitivity of each participant of the test 2. Results showed dee-primes in three different categories, namely: high; medium and low. The responses of the majority of participants were between one extreme and mean facility to discriminate the finished and unfinished sentences.

Keywords: perception ; prosody ; intonation; phonetic

RESUME

L'intonation peut être décomposée en éléments de structuration (finalisation et la composante de soutien), mais aussi sémantique fonctionnelle (focalisation / accent et l'accent lexical). La proposition assumée par le groupe de recherche ExProsodia, basé sur les études de Xu et Wang (1997), permet l'analyse séparée de chaque composante de F0. Dans cette étude, on traite les éléments de structuration, en particulier, l'étude de finalisation (F). L'étude vise, en premier lieu, d'analyser au moyen de tests de perception ce que les sujets qui parlent portugais brésilien (ci-PB) jugent comme un final de phrase assertive. C´est notre intention de vérifier également si la séquence de finalisation des phrases proposée par Ferreira Netto (2008) et les propositions de Consoni *et al* (2009) e Consoni (2011) sont observées. Et enfin, c´est notre but d'inspecter si les données perceptives correspondent à ceux fournis par l'analyse automatique obtenue par la routine ExProsodia®. Cette étude vise à intégrer une collection de travaux réalisés par le groupe de recherche ExProsodia - analyse automatique de l'intonation dans le discours de la langue portugaise, ligne de recherche dans le domaine de la perception de prosodie sous la direction du professeur Dr. Waldemar Ferreira Netto. Nous avons développé deux tests. Au cours du premier test, neuf phrases ont été présentées (trois phrases sans manipulation et six manipulées). Les phrases manipulées, au long de ce travail, sont connues comme *phrases modèles*. Les six phrases manipulées ont été subdivisés en deux groupes de trois phrases chacun: les phrases monotones et les phrases manipulées de telle sorte qu'entre le ton moyen (TM) et le ton de finalisation (TF) il y avait une distance de 7st. Nous avons nommé les premières phrases comme *phrases neutres* et les deuxièmes comme *phrases hypothèses.* Au deuxième test, six phrases (seulement les phrases manipulées, c'est-

à-dire, les *phrases neutres* et les *phrases hypothèses*) ont été présentées. Chaque fois que le participant entendait une phrase, on lui demandait si celle-là, en fonction de sa perception, si elle était terminée ou non. Deux options de réponse ont été donnés aux auditeurs: **oui, la phrase est terminée** ou **non, la phrase n'est pas terminée**. Les tests ont été administrés à 20 auditeurs des deux sexes, âgés de 18 et 55 ans, tous à l'éducation entre lycée et supérieur. Les résultats des tests indiquent que les pourcentages d'auditeurs qui considèrent les phrases qui ne sont pas terminées comme *neutre*, elles ont été toujours élevées (autour de 80%). En outre, le produit de l'analyse des phrases hypothèses corrobore les données de l'ouvrage déjà citée. . (CONSONI; FERREIRA NETTO, 2008; CONSONI *et al.,* 2008; CONSONI *et al.,* 2009 e CONSONI, 2011; FERREIRA NETTO, 2006, 2008; ROSA 2009). L'analyse de l'essai effectué par deux de la théorie de la détection du signal (TDS) (VERT; SWETS, 1966; MACMILLAN; CREELMAN, 2005; ADBI, 2007). En faisant l´usage de la TDS, nous avons l'intention d'évaluer la force de la sensibilité perceptive du test 2. Les résultats du test 2 de chaque participant a montré deux dee-primes dans trois catégories différentes, à savoir: haute; moyenne et basse. Les réponses de la majorité des participants étaient entre un extrême et moyenne facilité de discriminer les phrases terminées de celles qui ne sont pas terminées.

Mots-clés: perception ; prosodie ; intonation; phonetique

SUMÁRIO

PRIMEIRA PARTE

1 INTRODUÇÃO.....22
1.1 Prosódia: sentido e seus constituintes.....22
1.2 Modelos clássicos de análise entoacional.....24
1.3 Modelos de análise entoacional do PB.....29

2 DECLINAÇÃO DA FREQUÊNCIA FUNDAMENTAL.....33
2.1 A percepção de finalização de frases.....33
2.2 Padrão entoacional das finalizações de frases da língua portuguesa.....35

3 EXPROSODIA®: MODELO DE DECOMPOSIÇÃO DA ENTOAÇÃO.....38
3.1 A rotina de análise automática da entoação ExProsodia®.....38
3.2 Percepção de prosódia e análise automática produzida pela rotina ExProsodia®.....39
3.2.1 Primeiras hipóteses acerca da percepção da finalização de frases do PB.....39

4 JUSTIFICATIVA E OBJETIVO DESTE ESTUDO.....42

SEGUNDA PARTE

5 APRESENTAÇÃO DO CORPUS.....46
5.1 Delineamento inicial para os testes.....47

6 TESTE DE PERCEPÇÃO.....51
6.1 Teste 1.....51
6.1.1 Metodologia do teste 1 51

6.1.2 Análise dos dados e resultados do teste 1.....54

6.1.2.1 Análise do teste *two-alternatives forced choice tasks* e resultados do teste 1.....57

6.1.2.1.1 Análise descritiva da percepção da *frase modelo*.....60

6.1.2.1.2 Análise descritiva da percepção da *frase neutra*.....61

6.1.2.1.3 Análise descritiva da percepção da *frase hipótese*.....62

6.2 Teste 2.....63

6. 2.1 Metodologia do teste 2.....63

6.2.2 Teoria da detecção de sinais.....63

6.2.3 Análise dos dados e resultado do teste 2.....66

7. DISCUSSÕES FINAIS.....70

REFERÊNCIAS BIBLIOGRÁFICAS.....73

ANEXOS.....80

Lista de abreviaturas

F0 frequência fundamental

PB Português Brasileiro

TM tom médio

TF tom de finalização

F componente estruturadora de finalização

ms milissegundos

Hz Hertz

st semitons

LDT limiares de diferenciação tonal

AED análise exploratória dos dados

TDS Teoria da detecção de sinais

RMS *root mean square*

Lista de figuras

Figura 1- Esquema das características prosódicas gerais da fala humana. (HIRST; DI CRISTO, 1998, p. 7) 23

Figura 2- Exemplo da notação de Liberman. (LIBERMAN, 1975, p. 49).....27

Figura 3-Notação dos níveis tonais. (CAGLIARI, 1982, p. 161).....30

Figura 4- Exemplo de representação pictórica das variações melódicas de alguns tons. (CAGLIARI, 1982, p. 163).....30

Figura 5- Modelo de notação do INTSINT. (HIRST; DI CRISTO, 1998, p. 16).....31

Figura 6- Exemplo da notação do INTSINT (HIRST; DI CRISTO, 1998) usada por Moraes (1998) para caracterizar o padrão entoacional da focalização no PB (MORAES, 1998, p. 188).....31

Figura 7- Modelo de padrão entoacional (VAISSIÈRE, 2005, p. 247).....34

Figura 8- Tendência geral da curva de F0 para frases declarativas (à esquerda) e tendência em línguas como o inglês, dinamarquês, francês, japonês e espanhol (à direita) (VAISSIÈRE, 2002, p. 4).....34

Figura 9- Padrão entoacional de frases declarativas do PB (MORAES, 1998, p. 183).....36

Figura 10- Exemplo de gráfico produzido pela rotina ExProsódia®. Nesse gráfico são apresentados os dados de frequência fundamental (F0), os dados detalhados de tom médio (TM; TM superior e TM inferior) e os dados de pausa em milissegundos. O tom de

finalização (TF) é representado pelo ultimo ponto indicado na curva de declinação de F0.....38

Figura 11- Resultado da análise comparativa entre as marcações (os pontos) feitas pelos destros e as marcações feitas pelos canhotos. A linha indica forte correlação entre os dados ($R^2 = 0,88$) (retirado de ROSA, 2009a).....40

Figura 12- Os traços contínuos mais escuros mostram o TM e a Finalização. Os traços pontilhados mostram os limites acima, de 3 st, e abaixo, de 4 st, do TM. A escala horizontal mostra os valores em MIDI, referindo os graus da escala cromática diatônica ocidental. A linha curva no centro mostra a tendência à finalização da frase, e as linhas diagonais apontam para Proeminência e para a finalização (retirado de CONSONI, 2011, p. 102).....43

Figura 13- Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **A JULIANA ADORA COMER FAROFA**.....47

Figura 14- Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **MARIANA TEM MEDO DE ALTURA.**.....47

Figura 15- Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **TATIANA CONVERSA COM A MARISA**.....47

Figura 16- Gráfico dos valores de TM e TF da *frase modelo* "A Juliana adora comer farofa."..... 48

Figura 17- Gráfico dos valores de TM e TF da *frase modelo* "Mariana tem medo de altura."..... 49

Figura 18- Gráfico dos valores de TM e TF da *frase modelo* "Tatiana conversa com a Marisa."..... 49

Figura 19- Curva entoacional da *frase neutra* "A Juliana adora comer farofa".....52

Figura 20- Curva entoacional da *frase neutra* "Mariana tem medo de altura."....52

Figura 21- Curva entoacional da *frase neutra* "Tatiana conversa com a Marisa."....53

Figura 22- Curva entoacional da *frase hipótese* "A Juliana adora comer farofa." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st....53

Figura 23- Curva entoacional da *frase hipótese* "Mariana tem medo de altura." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st....53

Figura 24- Curva entoacional da *frase hipótese* "Tatiana conversa com a Marisa." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st....54

Figura 25- Fórmula do cálculo de qui quadrado (retirado de HACKBARTH NETO; STEIN, 2003)....56

Figura 26- Gráfico padrão de respostas **não, a frase não está finalizada** e gráfico padrão das respostas **sim, a frase está finalizada**. O eixo x representa o número de sujeitos que respondeu 'não' ou 'sim.' O eixo y representa as 9 frases....58

Figura 27- Gráfico *box plot* do padrão de resposta *'não' (à esquerda) e do padrão de resposta 'sim' (à direita)*....59

Figura 28- Gráfico de ROC dos sujeitos 3, 4, 8,16 e 19....67

Figura 29- Gráfico de ROC dos sujeitos 1, 2, 5, 6, 11, 14, 17, 18 e 20....68

Figura 30- Gráfico de ROC dos sujeitos 7, 9, 13 e 15....69

Figura 31 - Na figura, os círculos grandes representam os tons atratores, os círculos pequenos representam todas as variações que podem ocorrer na produção da entoação, as setas indicam as possíveis forças de atração para um ou outro tom (retirado de CONSONI, 2011, p. 101)....71

Lista de tabelas

Tabela 1- Padrão de resposta em número de sujeitos e em proporção.....55

Tabela 2-Valores de frequência observada e frequência esperada para cálculo de qui quadrado.....55

Tabela 3-Tabela de contingência com fórmula de correção de Yates baseada em tabela encontrada em (HACKBARTH NETO; STEIN, 2003, p. 6).....56

Tabela 4- Quadro-resumo dos valores de média e variâncias do padrão de resposta das 9 frases.....57

Tabela 5- Quadro-resumo dos valores de ANOVA.....57

Tabela 6-Padrão de respostas da Frase 1 - *frase modelo* "A Juliana adora comer farofa."....60

Tabela 7-Padrão de respostas da Frase 4 - *frase modelo* "Mariana tem medo de altura."....60

Tabela 8-Padrão de respostas da Frase 7 - *frase modelo* "Tatiana conversa com a Marisa."....60

Tabela 9-Padrão de respostas da Frase 2 - *frase neutra* "A Juliana adora comer farofa."....61

Tabela 10-Padrão de respostas da Frase 5 - *frase neutra* "Mariana tem medo de altura."....61

Tabela 11-Padrão de respostas da Frase 8 - *frase neutra* "Tatiana conversa com a Marisa."....61

Tabela 12-Padrão de respostas da Frase 3 - *frase hipótese* "A Juliana adora comer farofa."....62

Tabela 13-Padrão de respostas da Frase 6 – *frase hipótese* "Mariana tem medo de altura."....62

Tabela 14-Padrão de respostas da Frase 9 - *frase hipótese* "Tatiana conversa com a Marisa."....62

Tabela 15- Tabela das quatro possibilidades de resposta da TDS (ABDI, 2007, p. 2)....64

Tabela 16- Modelo de análise da sensibilidade perceptiva das finalizações de frases auferida por participante....65

Tabela 17- Cálculo do *dee-prime*. Sendo z o cálculo da inversa normal gaussiana (MACMILLAN; CREELMAN, 2005)....65

Tabela 18- Valores de *dee-prime* dos participantes do teste 2....66

Tabela 19- Proporção e *dee-prime* das respostas dos sujeitos 3,4,8,16 e 19....67

Tabela 20- Proporção e *dee-prime* das respostas dos sujeitos 1, 2, 5, 6, 11, 14, 17, 18 e 20....67

Tabela 21- Proporção e *dee-prime* das respostas dos sujeitos 7, 9, 10, 13 e 15....68

Primeira parte

1 INTRODUÇÃO

1.1 Prosódia: sentido e seus constituintes

É amplamente divulgado que a origem do vocábulo prosódia (προσ-ωδία) remonta à Grécia descrita na República de Platão. Nessa época, sua propriedade estava atrelada às características melódicas da versificação. Por meio da observação do histórico "daquilo que molda nossa enunciação" (BARBOSA, 2012, p. 11), é possível assegurar que a análise da prosódia concentrada, primordialmente, na sílaba (e na pronúncia) perdurou até o advento da Linguística Moderna. A evolução desse campo de estudo permitiu o desenvolvimento de um sentido particularizado dado à prosódia. Para Barbosa (2012), o *Dicionário enciclopédico da ciência das linguagens* (DUCROT; TODOROV, 1972) é o marco da definição especializada do termo "que molda nossa enunciação imprimindo a 'o que se fala' um 'modo de falar' que é dirigido intencionalmente ou não ao ouvinte" (BARBOSA, 2012, p. 14).

Crystal (1976) define a prosódia como um termo atrelado à fonética e fonologia suprassegmental que se refere, coletivamente, às variações de *pitch*, duração, tempo e ritmo. De acordo com Di Cristo (2000), a prosódia é o termo linguístico que comporta um aspecto fonético -descritivo - e outro fonológico -representativo- dos traços acústicos da fala. Cagliari (1992) decompõe os constituintes prosódicos de maneira a classificá-los em elementos da melodia da fala, da dinâmica da fala e da expressão da voz. Os primeiros associam-se à tessitura, tom e entoação. Os elementos de dinâmica atrelam-se à duração, pausa, acento, *arsis/tesis*, ritmo, mora, tempo. Os últimos tratam das características do registro, qualidade de voz, volume. Centraremos nosso estudo em um dos termos do elemento melódico da fala: a entoação. Esta pode ser definida por meio de seu correspondente perceptual: o *pitch* e, também, pelo seu correspondente acústico: o número de vibrações produzidas pelas cordas vocais em um espaço de tempo, ou seja, a frequência fundamental ou F0.

Diversos trabalhos ocupam-se em destrinchar a noção de prosódia. Outros se preocupam em determinar e elencar os chamados níveis entoacionais. A tarefa desses autores

não tem sido fácil. Hirst e Di Cristo (1998) já alertaram para essa dificuldade; afinal estes consideram que a tentativa de descrever o sistema entoacional de uma dada língua é uma meta particularmente difícil, pois a entoação é, ao mesmo tempo, um dos aspectos linguísticos mais universais e mais específicos da linguagem humana. De acordo com Hirst e Di Cristo (1998) *"the term intonation has often been used interchangeably in the literature with that of prosody. When a distinction is made between the two words, this is often not explicit"* (HIRST; DI CRISTO, 1998, p. 3). Na coletânea *Intonation System*, os autores propõem o uso do termo prosódia em seu sentido mais geral, abrangendo os sistemas cognitivos abstratos e parâmetros físicos. No nível fonológico, a prosódia consiste em um número de sistemas lexicais (tom, acento e quantidade) e um não-lexical, ou seja, a entoação - como seus parâmetros prosódicos tais quais a frequência fundamental, a intensidade, a duração e as características espectrais - (HIRST; DI CRISTO, 1998, p. 7).

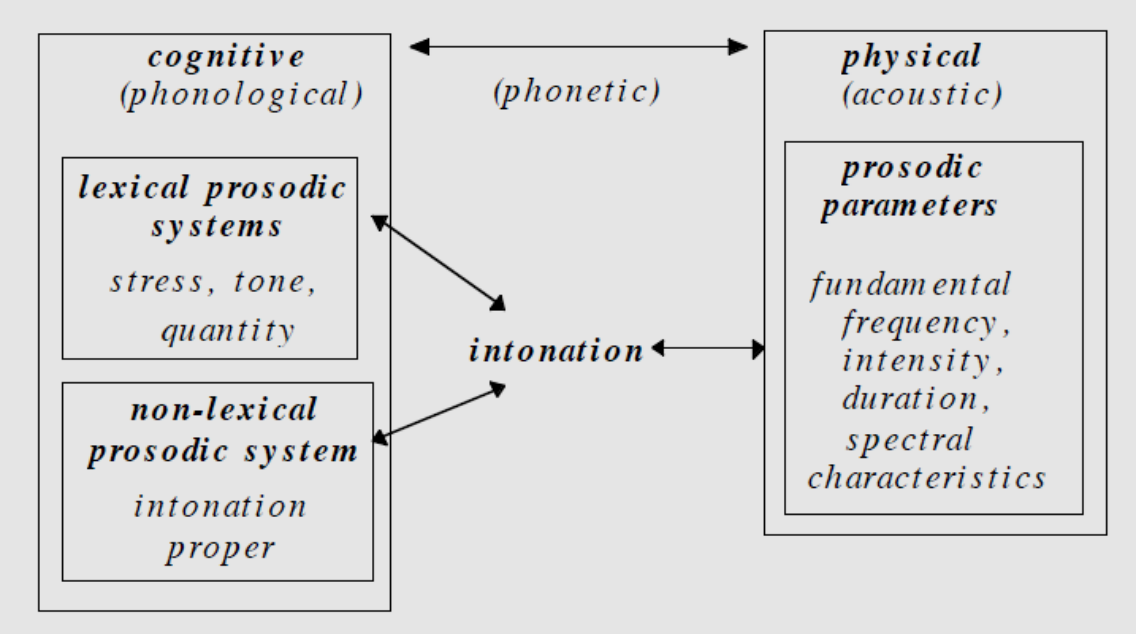


Figura 1 - Esquema das características prosódicas gerais da fala humana. (HIRST; DI CRISTO, 1998, p. 7).

1.2 Modelos clássicos de análise entoacional

Faremos, aqui, uma brevíssima apresentação de conhecidos modelos de análise entoacional. Escolhemos abordar apenas algumas das principais características de modelos bastante tradicionais como o de Troubetzkoy (1939) até os recentes trabalhos de fonologia não linear de Nespor e Vogel (1986). Na seção seguinte, elencamos, também, dois importantes estudos a respeito do Português Brasileiro (PB).

Troubetzkoy (1939), no início de sua obra, diferencia o estudo fonético do fonológico. Para o autor, os sons *du langage* que a fonética deve estudar possuem um grande número de particularidades acústicas e articulatórias que devem ser exaustivamente observadas pelo foneticista. Dentre essas propriedades, podemos citar as fônicas. Para o autor de *Principles de phonologie*[1], as particularidades fônicas que formam oposições distintivas dentro de diversas línguas podem ser repartidas em três classes: as particularidades vocálicas, consonantais e prosódicas. Ainda, segundo Troubetzkoy (1939), o processo fonatório da fala humana pode ser representado pelo seguinte esquema:

> quelqu'un siffle ou chante une mélodie à l'embouchure d'un tuyau sonore, et tantôt ferme, tantôt ouvre avec la main l'orifice de sortie de ce tuyau. Il est clair qu'on peut distinguer dans l'effet acoustique perçu par ce moyen trois sortes d'éléments : d'abord les sections comprises entre la fermeture et l'ouverture de l'orifice de sortie en tant que telles ; deuxièmement les sections comprises entre l'ouverture et la fermeture du même orifice en tant que telles ; troisièmement les segments de la mélodie sifflée ou chantée dans le tuyau en tant que tels. Les éléments du premier type correspondent aux consonnes, ceux du second type aux voyelles et ceux du troisième type aux unités prosodiques (TROUBETZKOY, 1939, p. 94).

De acordo com o autor do oitavo volume dos trabalhos do Círculo Linguístico de Praga, o terceiro tipo do processo fonatório proposto é o prosódico. As unidades rítmico-melódicas possuem três particularidades :

> l'unité prosodique doit être conçue comme une unité « musicale » (rythmique-mélodique), ou mieux comme un segment d'une unité « musicale », il s'en suit que les « caractéristiques prosodiques » se rapportent aux marques spécifiques de chacune des sections d'une mélodie (intensité, hauteur musicale) ou au mode de segmentation de la mélodie dans le processus phonatoire du discours humain. Le

[1] Neste estudo, fazemos menção à tradução de J. Cantineau da obra *Principles de Phonologie* de Troubetzkoy.

> premier type de particularités a pour résultat la différenciation rythmique-mélodique des unités prosodiques, le second type marque la liaison d'une unité prosodique donnée avec l'unité immédiatement voisine. En conséquence, les caractéristiques prosodiques peuvent être classées en particularités de différenciation et en particularités de mode de liaison (TROUBETZKOY, 1939, p. 99).

Pike (1964) apresenta-nos a entoação do inglês por meio de três perspectivas: partícula, onda ou campo. Na primeira perspectiva a entoação é considerada estática. Na segunda, a entoação seria dinâmica e o pico da onda representaria o núcleo do contorno entoacional, o final e espaço entre as ondas seriam interpretados como sendo as margens entre os contornos. Na perspectiva da entoação interpretada como um campo, esta representaria, entre outras coisas, contrastes entre a intersecção das categorias de forma e de significado.

É, portanto, com o auxílio da primeira abordagem, que o autor de *On the Grammar of Intonation* constrói seu conhecido sistema de notação em níveis. Para Pike (1964) haveria quatro padrões para as variações contrastivas de *pitch* no inglês americano. Trabalhos como o de Maeda (1976) surgiram na esteira da proposição dos *"four levels of contrastive pitch"* (PIKE, 1964).

> The particle perspective leads one to perceive intonation segments as if they were static bits. Through it one finds phonemes of intonation – if his assumptions leave room for any phonemes at all. From this implicit perspective come our four levels of contrastive pitch, which we shall symbolize as E(xtra-high), H(igh), M(id) and L(ow) (PIKE, 1964, p.106).

Unindo fonética e fonologia, Halliday (1970) discrimina a entoação em cinco níveis (alto, meio alto, médio, meio baixo, baixo). A entoação é analisada pelo autor de *A Course in a Spoken English Intonation* como um sistema complexo de três variáveis: tonalidade, tonicidade e o tom (dividido em primário e secundário). Um quarto elemento gravita entre essas três variáveis: o ritmo. Crystal (1976) apresenta os movimentos de *pitch* possíveis segundo modelo de Halliday (1970).

> Primary tone may be equate with other scholar´s nuclear pitch movements: Halliday recognize a fall, rise, 'sharp' fall-rise, fall-rise, low-rise, rise-fall, fall plus low rise and rise-fall plus low rise. Secondary tone exists in the pre-tonic stretch of utterance, and similar to other scholar´s type of head: Halliday distinguishes even (level, falling, rising) uneven (low 'spiky') and suspended ('listing'); high (level, falling, rising), and low (level, rising); mid (level) and low (level) sub-types (CRYSTAL, 1976, p. 43).

Liberman (1975), no início de seu trabalho, delimita alguns postulados, tais como: (i) entradas lexicais não são especificadas para características tonais, mas o tom é uma característica linguisticamente significante para o inglês; (ii) aspectos tonais não são inerentes ao segmento que compõe a representação fonológica do enunciado em inglês, mas têm uma identidade independente; (iii) representações tonais consistem em conjuntos bem ordenados de segmentos; (iv) os segmentos subjacentes de representações tonais são tons estáticos, definidos como *Low* e *High*. Tons cinéticos serão sempre analisados como sequências de tons estáticos; (v) a associação entre representações tonais e lexicais (não tonais) será estabelecida por regras linguísticas; (vi) padrões de acentos existem, mesmo a despeito de representações tonais, como propriedades do texto. Depois de esclarecidos os pressupostos, Liberman (1975) postula quatro questões às quais seu modelo teórico pretende responder:

> First, what are the underlying form by stress patterns, tonal patterns (tunes) and patterns of intonation phrasing? […] Second, by what rules is a given "tune" associated with its "text"? Why do change is stress pattern or intonational phrasing affect this association in the way that they do? Third, what is the nature of the phonetic representation of intonation, and by what rules is it related to underlying intonation form? […] Forth, how is the intonational system integrated into the theory of language as whole? (LIBERMAN, 1975, p. 28).

Em suma, a sugestão teórica proposta por Liberman (1975) é a de padrões métricos como 'árvores' cujos rótulos de nó sejam representados pelo símbolo *Strong* (forte) e pelo *Weak* (fraco). As 'árvores' devem possuir constituintes ordenados e devemos presumir que ela esteja 'enraizada' (*rooted*). Outra característica importante defendida por Liberman (1975) é a de que os 'galhos' dessas árvores métricas não devem possuir mais do que duas ramificações (ramificações binárias).

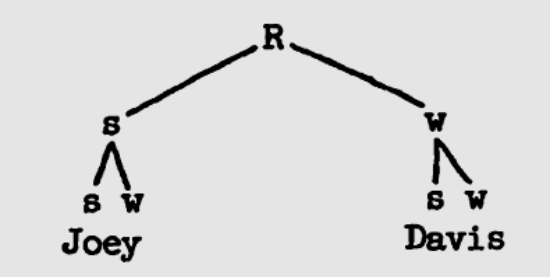


Figura 2 - Exemplo da notação de Liberman (LIBERMAN, 1975, p. 49).

Maeda (1976) propõe-se a identificar traços acústicos, fisiológicos, perceptuais e linguísticos da entoação do inglês americano. Para tanto, no capítulo 2 do estudo *A Characterization of American English Intonation,* Maeda (1976) analisa a frequência fundamental de 39 sentenças isoladas e 14 sentenças de um texto lido por diversos falantes. Por meio da análise dessas sentenças, foi encontrado um conjunto de símbolos denominados pelo autor de 'atributos.' Os 'atributos' *rise* (R) e *lowering* (L) caracterizariam as subidas e descidas do contorno da frequência fundamental; a baseline (BL) representaria a queda gradual da frequência fundamental ao longo da sentença. Há, ainda, em conjunto com os 'atributos' R e L, o Peak (P) que ocorre, geralmente, com um R e o *Rise on the plateau* (R1) (MAEDA, 1976, p. 17).

Uma das teses de Pierrehumbert (1980) é a de que a representação fonética da entoação, percebida como contorno da F0, tem particular importância, pois funciona como evidência da presença de representações subjacentes. Assim, é em razão do fracasso de métodos para acessar uma representação mais abstrata que a representação fonética tem sido útil, inclusive, em outros campos da linguísitca. Pierrehumbert (1980), portanto, deixa clara a importância do contorno do F0, enunciando que *"...all possible combinations of nuclear pitch accent, phrase accent, and boundary tone correspond to well-formed F0 contours in English"* (PIERREHUMBERT, 1980, p. 28). Nas discussões finais do aclamado *The Phonology and Phonetics of English Intonation,* Pierrehumbert (1980) anuncia que as representações fonológicas têm três componentes:

> First, there is a grammar which generates well-formed tunes for the intonation phrases. The tones are structured strings of H and L tones. (…) The second component of the phonological representation is a metrical grid for the text of the phrase. The third is a set of rules which align the tune with the text, on the basis of the structure of the tune and the metrical representation for the text (PIERREHUMBERT, 1980, p.236).

O modelo da linguista americana é o de apenas duas variáveis: *Low* e *High*. Para Pierrehumbert (1980), a teoria de dois tons dissolve diversos problemas da teoria dos quatro tons e ainda possui a vantagem de descrever a entoação como uma sequência de níveis tonais e, para tanto, ela elucida que *"notation system uses %, the symbol for the intonation phrase boundary, as a diacritic for the boundary tones. That is two boundary tones will be referred to as L% and H%"* (PIERREHUMBERT, 1980, p. 28).

Bolinger (1998[2]) supõe que o inglês americano não tem palavras tonais e nem tons segmentais. O autor postula, ainda, que não há, na referida língua, relação definida entre entoação e gramática já que a entoação transmite emoção e atitude. Outro postulado importante do autor é aquele segundo o qual a entoação é extremamente dependente de duas características da sílaba.

> The phonological apparatus by which intonation attaches itself to verbal locutions is heavily dependent on two features of the syllable: (a) whether or not the syllable is stressed within the word (by and large we can say that there is one stress per word), and (b) whether the vowel of the syllable is reduced or full (BOLINGER, 1998, p. 52).

Segundo Bolinger (1998), a frequência fundamental da língua inglesa em seus aspectos de produção (não estereotipados) possui três funções: formar acentos, finalizações e contornos. Os primeiros seriam mudanças bruscas de frequência usualmente combinadas com o aumento no comprimento e sonoridade na sílaba acentuada de uma dada palavra. As finalizações da língua oral são marcadas, na língua escrita, por sinais de pontuação: ponto final para a queda da frequência e ponto de interrogação para subida da frequência. Contornos (ou grupos entoacionais) são formas melódicas de enunciados, mais ou menos delimitados por finalizações (BOLINGER, 1998, p. 48).

Podemos citar, ainda, estudos recentes que trabalham com a fonologia não linear. Isso significa dizer que a *"phonology has assumed that tone is formally represented in the lexicon of a tone language as a sequence of tonal segments (H and L for example)..."* (HIRST; DI

[2] Com o intuito de apresentar breve quadro evolutivo das teorias entoacionais, citamos o capítulo Intonation in American English presente no livro *Intonation systems: a survey of twenty languages* de Hirst e Di Cristo de 1998, mas a teoria de Bolinger já aparece descrita na obra *Around the edge of language* de 1964.

CRISTO, 1998, p. 8). Como exemplo disso, podemos citar a Fonologia métrica (SELKIRK, 1984; NESPOR; VOGEL, 1986; 2007).

O recorte de um dos traços prosódicos, a saber, o tom/entoação; permitiu, por exemplo, a abordagem fonológica da língua. Surgia, entre outras teorias, então, a Fonologia Prosódica e a Teoria dos Domínios. Por meio de um estudo que articulasse fonologia e sintaxe, Nespor e Vogel (1986, 2007) aventuram-se a apresentar a categorização da prosódia em domínios (*Teoria Relation-Based*). Segundo Wheeldon (2000), *"each prosodic constituent serves as the domain of application of specific phonological rules and phonetic process"* (WHEELDON, 2000, p.254).

De acordo com a *Teoria Relation-Based*, a prosódia apresenta sete domínios (sete níveis de constituintes prosódicos): a sílaba (syllable); o pé (foot); a palavra fonológica ou prosódica (Pwd), o grupo clítico (C); a frase fonológica (Pph); a frase entoacional ou o sintagma entoacional (I) e, por fim, o enunciado fonológico (U). Segundo a pesquisadora brasileira Luciana Tenani (2006), é possível, por exemplo, fazer uma análise da variação de F0 em termos de eventos tonais, relacionando *a Teoria Relation-Based* de Nespor e Vogel (1986; 2007) e a análise da entoação em níveis tonais como proposto por Pierrehumbert (1980).

1.3 Modelos de análise entoacional do PB

Fazendo uso de modelos clássicos, autores brasileiros como Cagliari (1982) e Moraes (1998) também se propõem a encontrar padrões entoacionais do PB**.** Na obra *Elementos de fonética do português brasileiro*, Cagliari (1982) apresenta as características entoacionais mais importantes do PB. Para tanto, o professor justifica a escolha de fazer a análise do PB usando o modelo proposto por Halliday (1970).

> O modelo descritivo de Halliday incorpora parte da descrição do ritmo da língua, como base para a descrição entoacional, marca de maneira simplificada, porém completa, todas as características mais importantes da língua, usa uma notação de fácil escrita e leitura e, sobretudo mantém laços muito mais estreitos com fatos gramaticais do que a maioria de outros modelos (CAGLIARI, 1982, p. 155).

Figura 3 - Notação dos níveis tonais (CAGLIARI, 1982, p. 161).

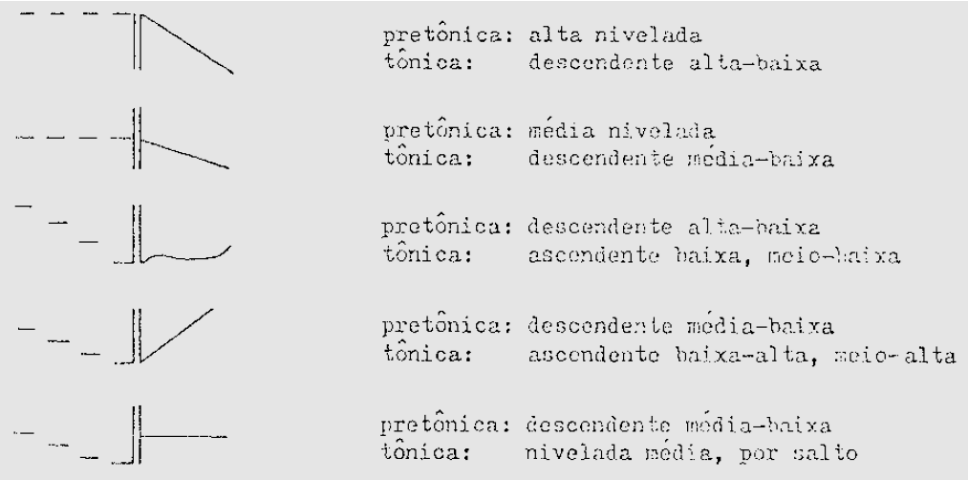


Figura 4 - Exemplo de representação pictórica das variações melódicas de alguns tons (CAGLIARI, 1982, p. 163).

Em trabalho intitulado *Intonational Phonology of Brazilian Portuguese*, Moraes (2007) sintetiza a gramática da entoação do PB da seguinte maneira:

> (i) phonological utterances (U) are made up of intonational phrases (IP) (…) (ii) there is always a nuclear accent in the final position of an IP, even if the utterance focus is anticipated, that is, if it is located in a non final position (…) (iii) nuclear accents are formed by two types of tonal events, pitch accents and boundary tones. (iv) boundary tones are associated basically with the right edge of intonational phrases and are realized on the final post stressed syllable(s), or on the final part of the last stressed syllable, if there aren't post stressed ones; there are no phrase accents. (v) pitch accents (PA) are associated with stressed syllables; they are only bitonal accents, and present always a leading tone, followed by the starred tone (MORAES, 2007, p.1).

Aliada à urgência de modelos entoacionais, vê-se a necessidade da criação de sistemas de transcrição. Moraes (1998), na tentativa de descrever os padrões entoacionais da língua portuguesa falada no Brasil, usou o *International Transcription System for Intonation* (INTSINT) desenvolvido por Hirst e Di Cristo (1998). A ideia da criação de um sistema como o INTSINT surge a partir da premissa de que *"an international transcription system needs to embody a certain number of hypotheses about what possible variations of prosodic features are significant across languages"* (HIRST; DI CRISTO, 1998, p. 14). O INTSINT usa como notação os seguintes símbolos: *top (T), bottom (B), higher (H), lower (L), same (S), upstep (U), downstep (D).*

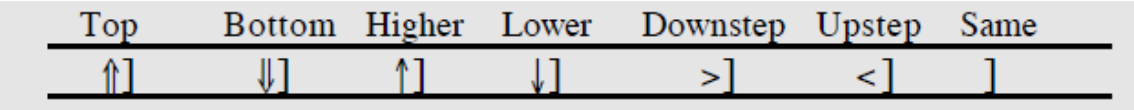


Figura 5 - Modelo de notação do INTSINT (HIRST; DI CRISTO, 1998, p. 16).

Na tentativa de elucidar os padrões entoacionais do PB, Moraes (1998) apresenta-nos exemplos de padrões entoacionais de frases não enfáticas; padrões da modalidade interrogativa e, entre outros padrões, aquele atrelado à focalização (tema - rema). Para o autor de *Intonational Phonology of Brazilian Portuguese*, o falante do PB usa a entoação para identificar se a informação já fora dada (o tema, ou seja, a informação já presente na consciência do ouvinte) ou é nova (o rema, uma nova informação). De acordo com Moraes (1998), o rema é caracterizado melodicamente por movimentos da frequência na sílaba tônica final. O padrão melódico do tema varia de acordo com a sua posição no enunciado. Moraes (1998) ilustra a partir da notação do INTSINT (HIRST; DI CRISTO, 1998).

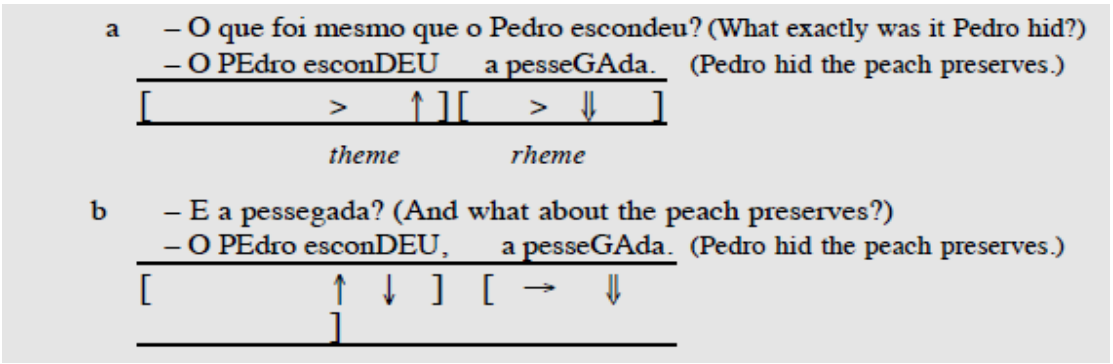


Figura 6 - Exemplo da notação do INTSINT (HIRST; DI CRISTO, 1998) usada por Moraes (1998) para caracterizar o padrão entoacional da focalização no PB (MORAES, 1998, p. 188).

De acordo como Hirst e Di Cristo (1998), o advento do INTSINT possibilitou o surgimento do sistema proposto por Silverman et al (1992), este último baseado nos trabalhos de Pierrehumbert (1980). O ToBI (*Tone and Break Indices)* consiste em níveis paralelos que representam os múltiplos componentes da prosódia. Os autores fazem questão de salientar que a camada que mais se assemelha à análise entoacional tradicional é a do nível tonal. Este é transcrito como uma sequência linear de frequências. O ToBi usa o sistema de "*break indices*" (SILVERMAN et al, 1992, p. 868) . Para a transcrição de tons altos e baixos são usados os símbolos *(H)igh e (L)ow*. De acordo com o entendimento de Consoni (2011), "a anotação ToBI combina representações da forma prosódica (H e L) com representações da função prosódica (-*%)" (CONSONI, 2011, p. 19).

2 DECLINAÇÃO DA FREQUÊNCIA FUNDAMENTAL

2.1 A percepção de finalização de frases

É notória a premissa de que a declinação representa uma queda no contorno do F0 (PIKE, 1945; VAISSIÈRE, 1971; MAEDA, 1976; PIERREHUMBERT, 1979, LADD, 1984, T´HART, COLLIER & COHEN, 1990, entre outros). No trabalho *The Perception of Fundamental Frequency Declination*, Pierrehumbert (1979) oferece uma proposta de análise da declinação frasal não em termos fisiológicos, mas do ponto de vista de representação mental. A partir dessa abordagem 'representativa' da declinação frasal, Pierrehumbert (1979) chega à conclusão de que *"speakers normalize for declination in judging the relative height of peaks in the intonation contour"* (PIERREHUMBERT, 1979, p. 368).

O efeito da declinação *"or the tendency of pitch to drift downwards over the course of an intonation group"* (PIERREHUMBERT, 1979, p. 363) pôde ser observado em diversos idiomas por meio da análise feita por Vaissière (1971) para o francês; Hirvonen (1970) para o finlandês; Silverstein (1976) e Welmers (1973) T´Hart e Cohen (1973) para o holandês e Maeda (1976) para o inglês (PIERREHUMBERT, 1979, p. 393).

Os estudos de Pierrehumbert (1979) forneceram subsídios para que Vaissière (2005) pudesse anunciar duas importantes premissas. A primeira de que *"a fall is judged to convey an impression of finality only if it restores the pitch to the level of the lower declination line"* (VAISSIÈRE, 2005, p. 247) e a segunda de que *"listeners seem to normalize their perception of the overall contour in terms of expected declination"* (VAISSIÈRE, 2005, p. 249).

A figura abaixo representa o arquétipo básico dos padrões entoacionais observados por Vaissière (2005) – inclusive o da declinação.

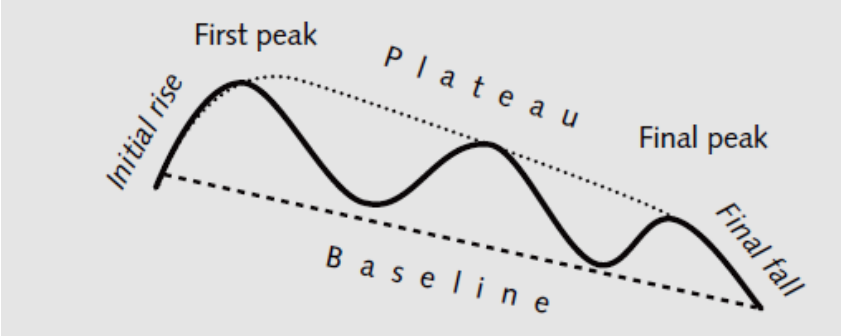


Figura 7 - Modelo de padrão entoacional (VAISSIÈRE, 2005, p. 247).

Em suma, para Vaissière (2002) a frequência fundamental tende a aumentar nas primeiras sílabas da frase e depois a diminuir (ora rapidamente, ora lentamente) até que esta seja finalizada na nota mais baixa da frase. Na figura abaixo pode ser observada a tendência geral da curva de F0 para frases declarativas (à esquerda) e essa mesma tendência em línguas como o inglês, dinamarquês, francês, japonês e espanhol (à direita).

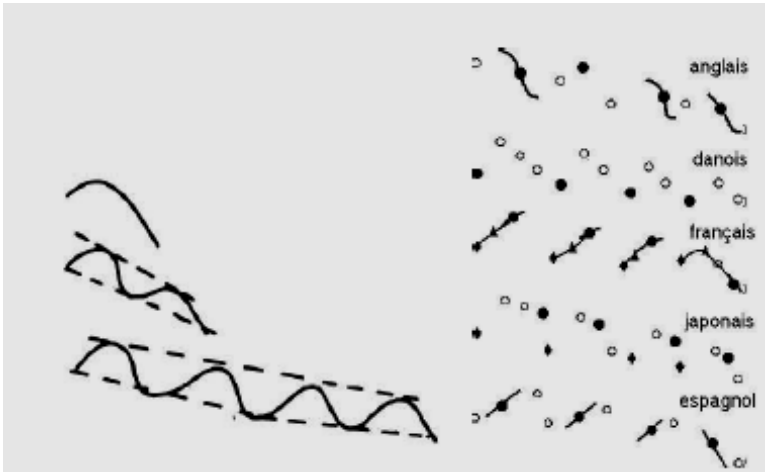


Figura 8 - Tendência geral da curva de F0 para frases declarativas (à esquerda) e tendência em línguas como o inglês, dinamarquês, francês, japonês e espanhol (à direita) (VAISSIÈRE, 2002, p. 4).

T´Hart (1998), ao elencar os principais resultados acústicos e fonéticos de seu estudo a respeito da língua holandesa, apresenta três principais categorias entoacionais do referido idioma: (i) características acústicas dos contornos entoacionais, (ii) a relação entre os contornos e os padrões encontrados e, por fim, (iii) o fenômeno da declinação. Em relação à

primeira característica, T´Hart (1998) menciona que os padrões entoacionais do holandês podem ser descritos a partir de 10 contornos de frequência perceptivelmente relevantes. Ainda, segundo T´Hart (1998), por meio da análise que relaciona os contornos e os padrões da entoação do holandês, foi possível construir uma 'gramática da entoação.' Ao observar o fenômeno da declinação no holandês, T´Hart (1998) chega à conclusão de que esta representa um queda gradual do F0 ao longo de um enunciado.

2.2 Padrão entoacional das finalizações de frases da língua portuguesa

O padrão entoacional das finalizações da língua portuguesa também fora objeto de estudo. Para o português europeu, Falé e Faria (2006) aplicaram um teste perceptivo na tentativa de determinar os contornos entoacionais que diferenciam uma frase declarativa simples de uma interrogativa.

> as frases declarativas simples terminam, habitualmente, com uma descida abrupta dos valores de F0 que se inicia na vogal pré--tónica final, tem continuidade na última vogal tónica e finda na última vogal da frase; por seu lado, as interrogativas globais terminam com uma grande subida final de F0 que começa na última vogal tónica, prosseguindo até ao fim do último elemento vozeado da frase (FALÉ; FARIA, 2006, p. 241).

Como exemplo de estudo da entoação do PB, podemos citar o trabalho *F0 Declination In Brazilian Portuguese In Read And Spontaneous Speech* , de Moraes (1999). O referido autor estudou o fenômeno da declinação frasal por meio do estudo comparativo entre a fala espontânea em comparação à leitura.

De acordo com Moraes (1998) as frases declarativas (sem ênfase) do PB são caracterizadas pela queda da frequência fundamental e o final do enunciado (mais precisamente, na sílaba final tônica).

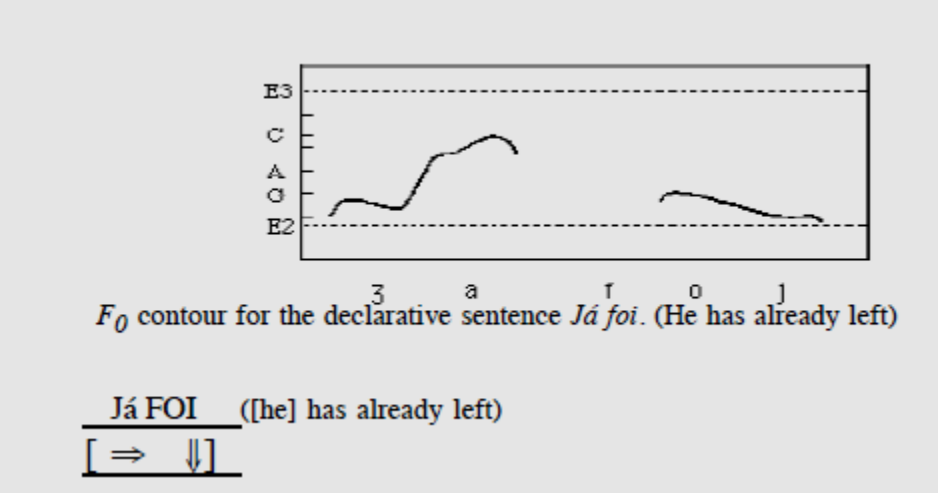


Figura 9 - Padrão entoacional de frases declarativas do PB (MORAES, 1998, p. 183).

Consoni, Ferreira Netto e Peres (2010) propuseram, também, uma observação comparativa entre a análise automática e a percepção da finalização de frases na leitura e na fala espontânea seguindo preceitos da notação musical. Os resultados do estudo dos linguistas brasileiros indicaram que existe uma tendência de que a finalização da frase seja representada por uma cadência maior perfeita (dominante-tônica).

Como nosso objetivo é analisar a finalização de frases assertivas, cumpre fazer algumas observações. É preciso salientar que partimos de dois preceitos retirados de um dos trabalhos norteadores do nosso estudo, qual seja, *Function of Sentence Intonation* (DANEŠ, 1960). O primeiro preceito é de que frase é um conjunto de unidades comunicativas moldadas pela entoação. O segundo é de que a afirmação é uma categoria não marcada.

> A função fundamental da entoação é transformar palavras, como unidades de chamamento, em unidades comunicativas, isto é, em frases. Cada palavra ou sucessão de palavras torna-se automaticamente uma frase quando e pronunciada em uma certa forma entoacional (DANEŠ, 1960, p. 43)[3].

Consideramos que a percepção dos procedimentos de segmentação frasal, com propósito de finalização, pressupõe que tanto o falante quanto o ouvinte tenham estabelecido

[3] Utilizamos tradução de Waldemar Ferreira Netto. Trecho original: "The fundamental function of intonation is to transform words, as appellative units, into communicative units, i.e. into utterances. Each word or succession of words automatically becomes an utterance when it is pronounced in a certain intonational form" (DANEŠ, 1960, p. 43).

um valor ou um procedimento fixo que caracterize a finalização da frase e que o reconheçam durante sua execução. Ainda, de acordo com Daneš (1960):

> O ouvinte que conhece a língua na qual ele está falando (a língua naturalmente prevê apenas esses ouvintes) também conhece as funções da entoação na língua. Na percepção da fala fluente, o ouvinte tem de reconhecer se a frase está terminando ou se ela ainda vai continuar. Isso significa que sua atenção está focalizada nos traços entoacionais que têm uma função distintiva no sistema de oposições entoacionais, isto é, aqueles que desempenham funções esperadas (DANEŠ, 1960, p. 42)[4].

Outros linguistas brasileiros examinaram o fenômeno da declinação frasal. No estudo *Variação tonal na finalização de frases de PB,* Baz e Ferreira Netto (2011) analisam comparativamente fala espontânea e leitura. Os autores, ao averiguarem a consistência da informação de que a finalização de orações assertivas em português brasileiro pode ser reconhecida pelo seu tom baixo, chegaram à conclusão de que há uma preferência de finalizar frases abaixo de 4st do tom médio (TM).

[4] Idem item 3. Trecho original: "The listener who knows the language is which he is spoken to (the language naturally anticipates only such listeners) also knows the functions of intonation in the language. In the perception of connected discourse the listener must recognize whether the utterance is ending or whether it will continue. This means that his attention is focused on those intonational features that have a distinctive function in the system of intonation oppositions, i.e. that perform expected functions "(DANEŠ, 1960, p. 42).

3 EXPROSODIA®: MODELO DE DECOMPOSIÇÃO DA ENTOAÇÃO

3.1 A rotina de análise automática da entoação ExProsodia®

A rotina ExProsodia® é uma ferramenta de análise automática da entoação construída em linguagem *Visual Basic 6.3* para o *Microsoft Excel*. A ExProsodia® é salutar na medida em que possibilita a análise isolada dos dados das componentes de F0. Para Ferreira Netto (2006, 2008), as componentes de F0 subdividem-se em estruturadoras (finalização e sustentação) e semântico-funcionais (foco/ênfase e acento lexical). Segundo Baz e Ferreira Netto (2011):

> Na proposta do projeto ExProsodia, parte-se do TM (tom médio) para a decomposição da série temporal de F0, que decorre das intenções prosódicas do falante. O TM é estabelecido pela sustentação do alvo do movimento rítmico tonal durante a produção da fala. É definido a partir das médias acumuladas no tempo (BAZ ; FERREIRA NETTO, 2011).

Os gráficos obtidos por meio da rotina ExProsodia® ilustram os pontos de determinada série temporal, como pode ser observado na figura abaixo.

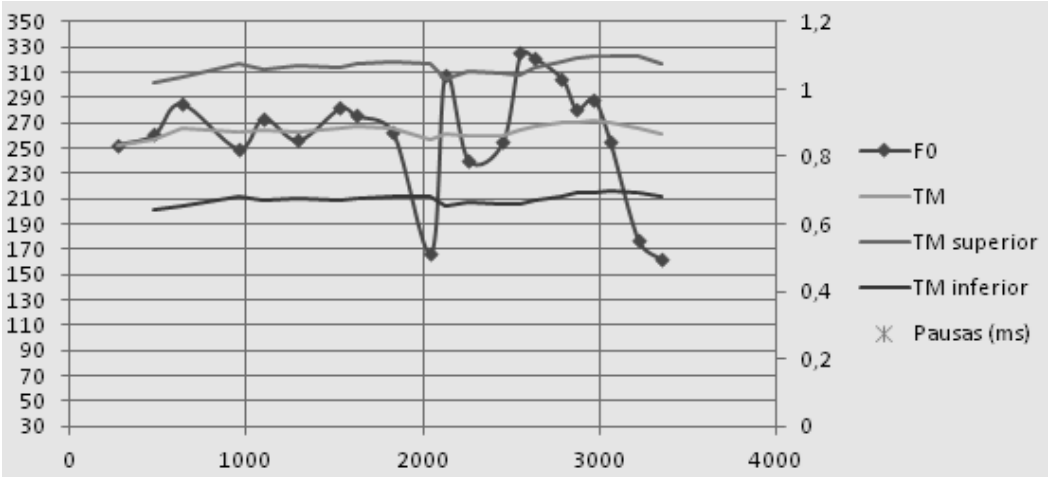


Figura 10 - Exemplo de gráfico produzido pela rotina ExProsódia®. Nesse gráfico são apresentados os dados de frequência fundamental (F0), os dados detalhados de tom médio (TM; TM superior e TM inferior) e os dados de pausa em milissegundos. O tom de finalização (TF) é representado pelo ultimo ponto indicado na curva de declinação de F0.

São os procedimentos da rotina: (i) gravação digital em formato .wav. ou .mp3; (ii) conversão, processamento e extração, na forma de texto, dos arquivos por meio do programa *Speech Filing System* (SFS) (HUCKVALE et al., 1987); (iii) inserção, na rotina, dos dados obtidos no SFS.

Os dados obtidos pela ExProsodia® são calculados por intermédio da observação de determinados parâmetros dos candidatos a pico silábico: (i) limiar de frequência entre 50 e 350 Hz; (ii) limiar de duração entre 20 ms (4 frames) e 300 ms (60 frames); (iii) limiar de intensidade 2000 RMS.

Neste estudo, utilizamos os valores do tom médio (TM), do tom de finalização (TF) e os gráficos da série temporal fornecidos pela rotina ExProsodia®.

3.2 Percepção de prosódia e análise automática produzida pela rotina ExProsodia®

3.2.1 Primeiras hipóteses acerca da percepção da finalização de frases do PB

Com o intuito de verificar a abrangência e a pertinência dos dados fornecidos pela rotina ExProsodia®, Consoni *et al* (2008) propuseram um estudo a respeito da correlação entre análise automática das finalizações prosódicas e a separação intuitiva de frases em textos longos. A verificação foi feita por testes de percepção de prosódia nos quais era apresentado um trecho de fala espontânea para que os ouvintes sinalizassem, nas transcrições, quais pontos eles reconheciam como final de frase. O texto transcrito foi apresentado em papel, com letras maiúsculas, sem qualquer sinal de pontuação. A comparação foi feita entre duas variáveis (tons graves em sílaba tônica e tons graves em sílabas tônicas finais) e a quantidade de marcas feitas pelos sujeitos em cada uma das posições as quais eles julgaram ser finais de sentenças. Os autores chegaram a resultados preliminares que indicavam uma forte correlação entre as marcações feitas pelos participantes do teste e as variáveis tom grave

em sílaba final e tom grave em sílaba tônica final (respectivamente, 0,84 e 0,79) (CONSONI *et al*, 2008).

Utilizando-se da mesma metodologia adotada por Consoni *et al* (2008), mas com *corpus* diferente, Rosa (2009a) sugeriu mais uma pesquisa. Baseando-se em trabalhos a respeito da lateralização do cérebro (LAGUITTON *et al*, 1997; HELLIGE, 2002 FRIEDERICI; ALTER, 2003; TREMBLAY *et al*, 2004), a autora buscou analisar se o fato de o sujeito ser destro ou canhoto poderia influenciar nos resultados de um estudo que envolvesse percepção de prosódia.

A autora decidiu, então, aplicar testes que teriam por objetivo verificar se, no que tange à percepção das finalizações de frases do PB, haveria diferença na qualidade perceptiva entre destros e canhotos; e se, na tentativa de correlacionar o que os ouvintes consideram como final de frase em textos longos e os dados propostos pela rotina Exprosodia® (FERREIRA NETTO, 2008), faria diferença os sujeitos serem destros ou canhotos (ROSA, 2009a).

Os resultados do estudo *Destros, canhotos e a percepção de procedimentos de segmentação frasal com propósito de finalização* indicaram que não havia divergência entre destros e canhotos no que se referente à percepção das finalizações, podendo ambos receber o mesmo tratamento numa análise automática da entoação (ROSA, 2009a).

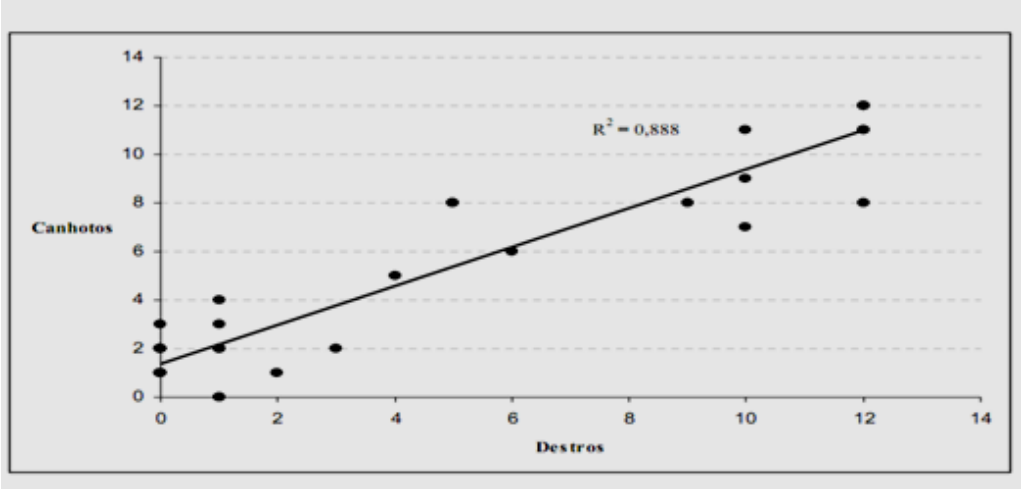


Figura 11 - Resultado da análise comparativa entre as marcações (os pontos) feitas pelos destros e as marcações feitas pelos canhotos. A linha indica forte correlação entre os dados ($R^2 = 0,88$). (retirado de ROSA, 2009a).

Os dois experimentos indicaram (CONSONI et al., 2008; ROSA, 2009a), *a priori*, a presença da variação tonal predominante nas finalizações assertivas de frases (BAZ e FERREIRA NETTO, 2011). Essas primeiras pesquisas assinalaram, também, a relevância dos dados fornecidos pela rotina de análise automática ExProsodia®. Os referidos trabalhos, ainda, abriram caminho profícuo para a possibilidade de avaliação perceptiva da componente estruturadora de finalização de frases assertivas em fala espontânea da língua portuguesa falada no Brasil.

4 JUSTIFICATIVA E OBJETIVO DESTE TRABALHO

Encontramos em Silva (2009) a informação de que diversos autores já estudaram a fala espontânea e leitura (grego: Nicolaidis, 2001; russo, búlgaro, polonês e tcheco: Barry e Andreeva, 2001; francês: Basset et al, 2001; finlandês: Lennes et al, 2001 etc), mas a pesquisa no Brasil a respeito desse assunto ainda é incipiente.

Como já exposto (*ver seção 3.2.1*), no que tange à análise da componente estruturadora de finalização em fala espontânea, Consoni *et al* (2008) verificaram a relação entre a análise automática de finalizações prosódicas e a separação intuitiva de frases em textos longos. A comparação foi feita entre duas variáveis (tons graves em sílaba tônica e tons graves em sílabas tônicas finais) e a quantidade de marcas feitas pelos sujeitos destros em cada uma das posições as quais eles julgaram ser finais de sentenças. Os autores chegaram a resultados preliminares que indicavam uma forte correlação.

À luz das proposições de T'Hart (1981), Consoni *et al* (2009) propuseram um estudo a respeito da sensibilidade para a variação de F0 no Português Brasileiro. O estudo "*Sensitivity to f0 variation in Brazilian Portuguese"* fora baseado nas pesquisas feitas por T'hart (1981) para a língua holandesa. O referido autor argumentava que, na língua holandesa, diferenças de 3 semitons são percebidas em situações comunicativas. Consoni *et al* (2009) fizeram dois experimentos nos quais foram manipuladas palavras da língua portuguesa do Brasil seguindo critérios propostos por Pierce (1983). O primeiro experimento avaliou a sensibilidade para a variação em semitons ascendentes de F0. O segundo observou a sensibilidade para a variação em semitons descendentes de F0. Os resultados apontaram para uma variação perceptivelmente relevante de 3 semitons ascendentes e de 4 semitons descendentes (CONSONI *et al*, 2009 e CONSONI, 2011). De acordo com Baz e Ferreira Netto (2011), o tom médio atua diretamente no princípio *Perceptual Magnet Effect* (IVERSON; KUHL 1995; KUHL, 2000 ; KUHL, TSAO, LIU; ZAHNG; DE BOER, 2001) e o 'poder magnético' do TM estaria atrelado à variação perceptivelmente relevante para o PB proposta por Consoni *et al* (2009) e Consoni (2011). Martins (2012) também observou os limiares de diferenciação tonal (LDT) do PB. Segundo o autor de *Aspectos da percepção e do*

controle entoacional do português brasileiro, o LDT é um "valor z a partir do qual o individuo é capaz de notar qualquer diferença entre dois tons seguidos" (MARTINS, 2012. p. 13). A fim de observar a variação perceptivelmente relevante para dada língua, T´Hart (1981) e Consoni (2011) manipularam a palavra. A proposta de Martins (2012) centra-se na observação do LDT no nível frasal.

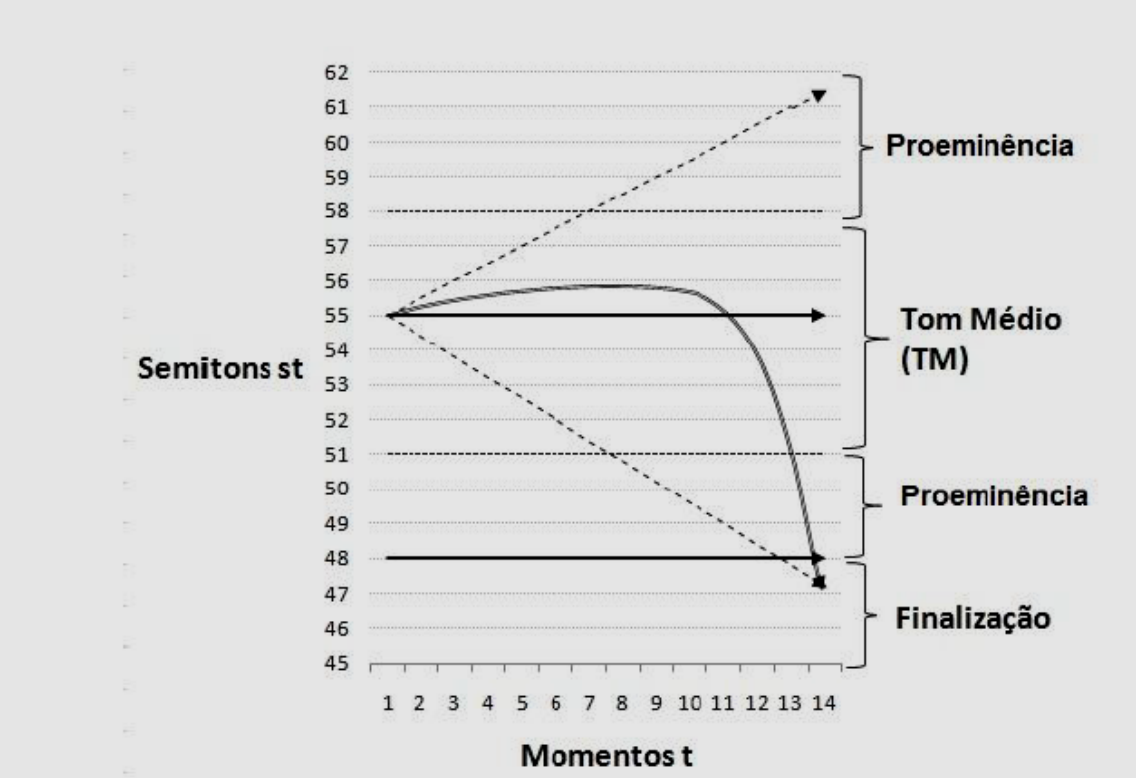


Figura 12 - Os traços contínuos mais escuros mostram o TM e a Finalização. Os traços pontilhados mostram os limites acima, de 3 st, e abaixo, de 4 st, do TM. A escala horizontal mostra os valores em MIDI, referindo os graus da escala cromática diatônica ocidental. A linha curva no centro mostra a tendência à finalização da frase, e as linhas diagonais apontam para Proeminência e para a finalização (retirado de CONSONI, 2011, p. 102).

Acreditamos que o estudo da seleção dos lugares das marcas de finalização é salutar na medida em que possibilita a averiguação de quais elementos prosódicos são utilizados para marcar as fronteiras prosódicas finais de uma frase a partir dos referenciais dos limiares de diferenciação tonal propostos por Consoni *et al* (2009) e Consoni (2011).

Este trabalho tem, portanto, como objetivo, primeiro, analisar por meio de testes de percepção da fala espontânea o que os sujeitos falantes do português brasileiro (doravante PB) julgam ser um final de frase assertiva. Depois, é nosso propósito analisar se os dados

perceptivos fornecidos pelos ouvinte coincidem com aqueles disponibilizados pela análise automática obtida por meio da rotina ExProsodia®. É nosso intento, ainda, verificar se, no que tange à percepção dos procedimentos de segmentação frasal com propósito de finalização, é observada a sequência finalizadora proposta por Ferreira Netto (2008), acrescida da hipótese de Consoni *et al* (2009) e Consoni (2011). Este estudo visa integrar a coletânea de trabalhos desenvolvidos pelo Grupo de pesquisa ExProsodia - Análise automática da entoação na fala de língua portuguesa, linha de pesquisa na área de Percepção de Prosódia sob a orientação do Prof. Dr. Waldemar Ferreira Netto.

Segunda parte

5 APRESENTAÇÃO DO *CORPUS*

Na esteira de trabalhos anteriores do grupo ExProsodia (COSTA, 2010; ROSA, 2010; FERREIRA NETTO *et al,* 2012), utilizamos, neste estudo, estímulos gravados por uma única voz feminina. Costa (2010) desenvolveu uma pesquisa na qual comparava a entoação da língua portuguesa falada por mulheres guatós e mulheres não-índias. Rosa (2010) analisou a existência ou não de diferenças significativas entre a percepção das fronteiras prosódicas finais de textos lidos e trechos de fala espontânea. Ferreira Netto *et al.* (2012) observaram a entoação do PB por meio da comparação entre a análise automática feita com auxilio da ferramenta ExProsodia® e testes perceptivos. Todos os trabalhos mencionados buscavam analisar comparativamente os dados da componente entoacional estruturadora de finalização obtida por intermédio de análise automática e aqueles colhidos por meio de testes de percepção. Dada a complexidade da voz e a necessidade de obter resultados apurados, o grupo de pesquisa Exprosodia optou por utilizar gravações de vozes femininas. O referido grupo justifica a escolha do uso de gravações de vozes do gênero supramencionado para estudos da entoação do PB com base no fato de que vozes femininas variam menos em diferentes falantes (ANDRADE, 2003).

Neste estudo foram gravadas três frases: “A Juliana adora comer farofa;” “Mariana tem medo de altura” e “Tatiana conversa com a Marisa.” Denominamos as gravações originais como *frase modelo,* ou seja, frases nas quais não há nenhuma manipulação. Era esperada a presença da LDT de 4st descendentes para caracterizar o início da possibilidade da percepção da finalização. Pode-se observar, nas figuras abaixo, o espectrograma com a curva entoacional sobreposta da gravação das *frases modelos* obtido por meio da ferramenta *Praat* (BOERSMA; WEENINK, 2011).

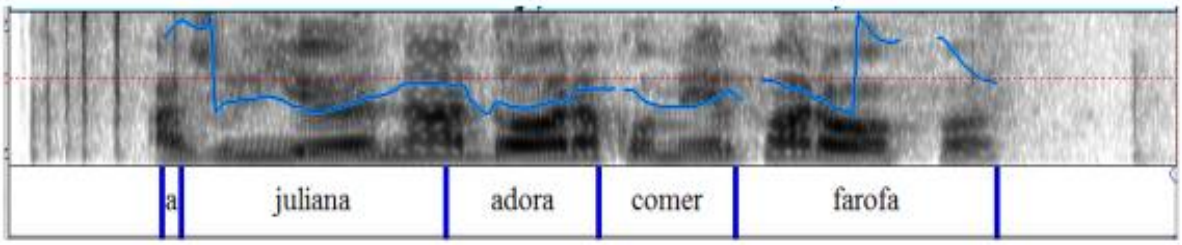


Figura 13 - Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **A JULIANA ADORA COMER FAROFA.**

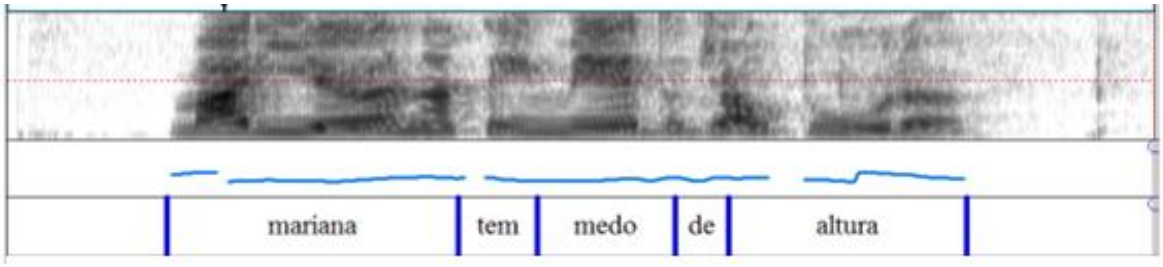


Figura 14 - Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **MARIANA TEM MEDO DE ALTURA.**

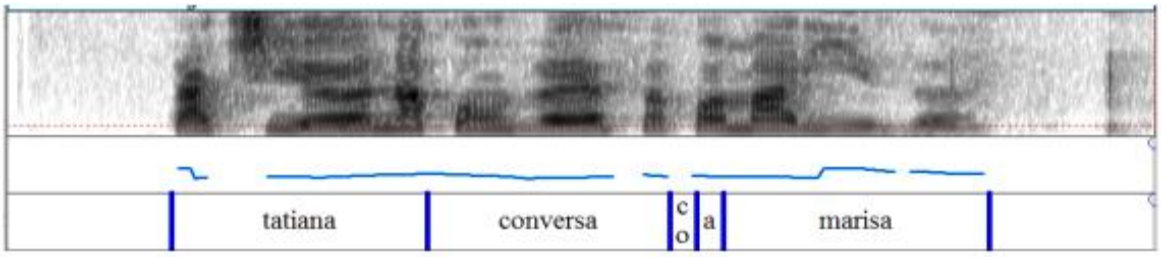


Figura 15 - Espectrograma com curva entoacional sobreposta da gravação sem nenhuma manipulação da frase **TATIANA CONVERSA COM A MARISA.**

5.1 Delineamento inicial para os testes

A análise dos dados foi feita com o auxílio da rotina ExProsodia® (FERREIRA NETTO, 2010). Ela baseia-se na hipótese de que as variações de entoação, percebidas pelos ouvintes no nível frasal, decorrem da coordenação entre as componentes estruturadoras e semântico-funcionais. Admite-se a proposição de que entoação do português brasileiro (PB)

pode ser decomposta em 5 tons (CAGLIARI,1982). Esses tons seriam estabelecidos como sendo 2 bandas acima ou abaixo do tom médio com uma escala de 3 semitons entre cada banda. A escala de 3 semitons foi defendida por T´Hart (1981) como sendo a variação tonal perceptivelmente relevante para os ouvintes do holandês. Para o PB, como já exposto, a variação perceptivelmente relevante é de 3 semitons ascendentes e de 4 semitons descendentes (CONSONI *et al.*, 2009). Fizemos, também, análises de dados utilizando a escala perceptiva de semitons e verificamos, primeiro, se os valores médios e final das sentenças são estáveis em cada sujeito e, depois, se esses valores obedecem à sequência finalizadora proposta por Ferreira Netto (2006, 2008).

A primeira etapa para o delineamento dos testes foi a obtenção dos valores do TM, do TF e dos gráficos da série temporal das frases a partir do processamento da rotina ExProsodia®. A versão da rotina utilizada neste trabalho dá valores semitons plotados em MIDI.

A *frase modelo* "A Juliana adora comer farofa" apresentou um TM de 54 MIDI e um TF de 50 MIDI. Nota-se uma diferença entre eles de 4st. Os valores podem ser observados no gráfico abaixo.

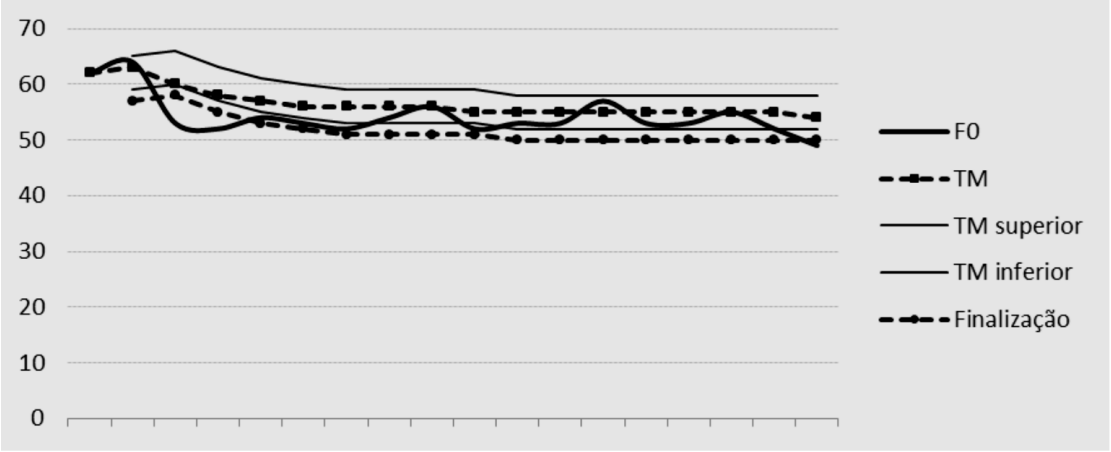


Figura 16 - Gráfico dos valores de TM e TF da *frase modelo* "A Juliana adora comer farofa."

O processamento dos dados da rotina ExProsodia®, apresentou para a *frase modelo* "Mariana tem medo de altura" um TM 52 MIDI de e um TF de 48 MIDI. Novamente, a diferença entre eles é de 4st.

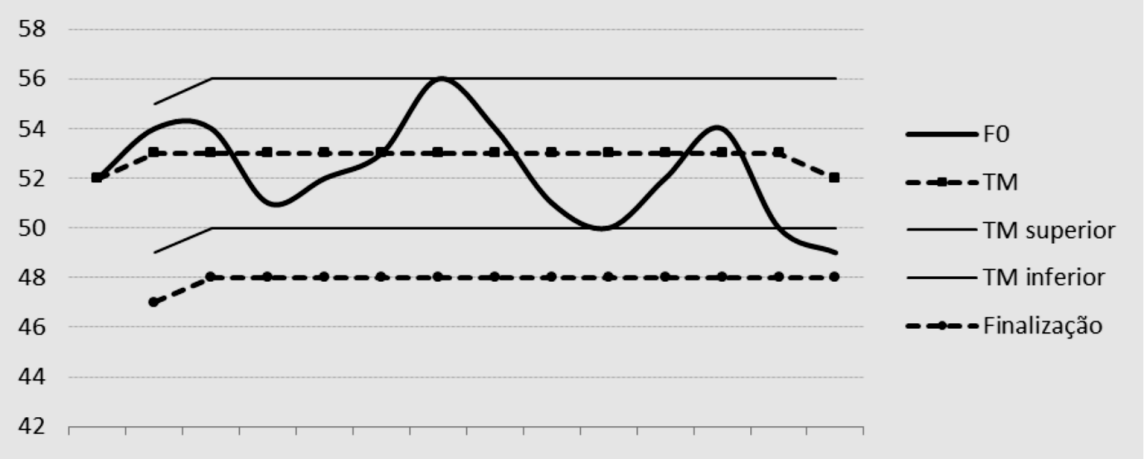


Figura 17 - Gráfico dos valores de TM e TF da *frase modelo* "Mariana tem medo de altura."

E, por fim, a análise dos dados da *frase modelo* "Tatiana conversa com a Marisa" forneceu um TM de 51 MIDI e TF de 47 MIDI. A diferença é de 4st.

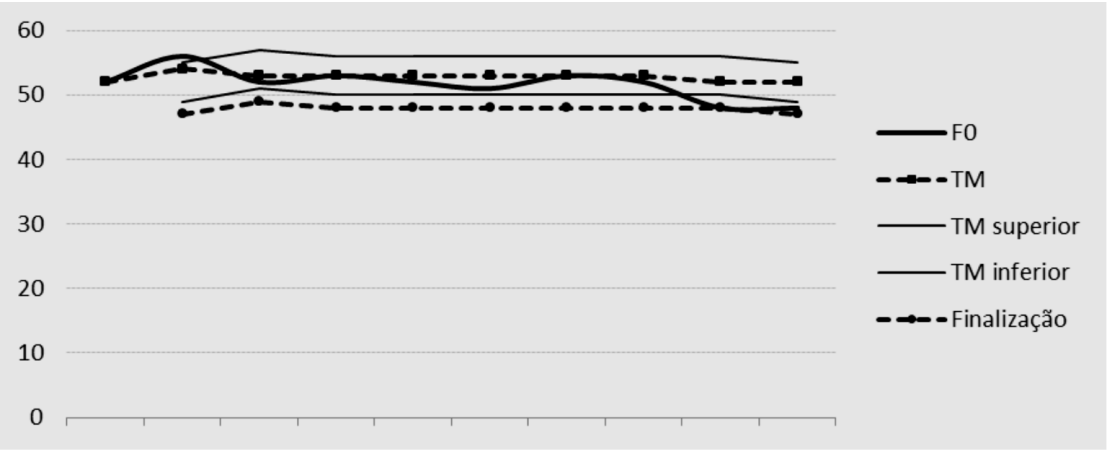


Figura 18 - Gráfico dos valores de TM e TF da *frase modelo* "Tatiana conversa com a Marisa."

Como pode ser verificado por meio dos dados obtidos com o auxílio da ferramenta ExProsodia®, há, de fato a presença da variação inicial perceptivelmente relevante para o reconhecimento de finalizações de frases, qual seja, 4st para o PB. Essa informação é pertinente na medida em que possibilitará que as *frases modelos* sejam testadas em oposição às frases que não obedecem à hipótese de finalização de Ferreira Netto (2008) somada aos pressupostos de Consoni (2011).

6 TESTE DE PERCEPÇÃO

6.1 Teste 1

6.1.1 Metodologia do teste 1

Elaboramos um teste no qual foram apresentadas nove frases. Uma *frase modelo* e duas manipuladas de cada uma das três gravações. Os sujeitos que aceitaram participar voluntariamente do teste *(ver anexo I)* foram alocados, individualmente, em lugar extremamente silencioso juntamente com a aplicadora do teste.[5] A aplicadora do teste era a responsável por iniciar a execução das gravações. As frases eram apresentadas de maneira randômica. Foi dado um intervalo entre 5 e 10 segundos entre cada audição. O teste foi aplicado em 20 ouvintes de ambos os sexos, com idade entre 18 e 55 anos, todos com escolaridade entre ensino médio e nível superior. Cada vez que o participante ouvia uma frase era solicitado a este que indicasse se a frase, de acordo com a sua percepção, estava ou não finalizada. Foram dadas duas opções de resposta aos ouvintes: ***sim, a frase está finalizada*** ou ***não, a frase não está finalizada.***

As escolhas metodológicas dos testes foram baseadas nas palavras de Daneš (1960), a saber:

> A análise auditiva precisa ser refinada, sobretudo pelo uso de vários procedimentos experimentais. Não é possível se contentar com audição casual uma frase nem apenas com o uso pessoal; como a próxima etapa, alguns tipos de frases precisam ser procurados deliberadamente, de acordo com um plano pré-estabelecido. Os procedimentos mais importantes são: (a) repetição da mesma frase pela mesma pessoa e por diferentes pessoas, uma vez que a comparação de varias realizações de uma mesma frase elimina muito dos traços casuais e idiossincráticos; (b) omissão, substituição e transformação — procedimentos que são úteis em outras áreas da pesquisa linguística, e que nos permitem atingir traços fonológicos relevantes e verificar as relações entre padrões entoacionais [...] (DANES, 1960, p. 38)[6]

[5] A aplicadora de todos os testes foi a mestranda Renata Cezar de Moraes Rosa, autora deste trabalho.

[6] Idem item3. Trecho original: " The auditory analysis needs to be refined, above all by using various experimental procedures. One cannot be content with casually overheard utterances or with personal usage only; as the next step, certain types of utterances must be sought out deliberately, according to a fixed plan. The most

É notória a premissa de que a análise do contorno entoacional e os resultados obtidos podem ser influenciados pela duração das palavras e a posição das sílabas. Assim, neste estudo, a escolha das frases obedeceu a alguns critérios. Entre eles podemos citar: o número de sílabas da última palavra; a duração das últimas palavras e das frases como um todo e o lugar da atribuição do acento na palavra. As três frases selecionadas têm, praticamente, a mesma duração. A última palavra de cada frase é, propositalmente, uma paroxítona. Ainda, as últimas palavras também possuem a mesma duração.

Utilizamos a ferramenta *Praat* (BOERSMA; WEENINK, 2011) para fazer, nas frases, duas manipulações diferentes. A primeira é a transformação da frase tonal em monotonal. Para tanto, foi necessária a captura do arquivo sonoro *sound phrase* e a transformação deste no arquivo *manipulation phrase* por meio dos ícones *manipulation*, *to manipulate*, O arquivo *manipulation phrase* fora estabilizado em 6 semitons com a ajuda dos ícones *pitch* e *stylise pitch*. À primeira manipulação, demos o nome de *frase neutra*. Abaixo, apresentamos as curvas entoacionais das *frases neutras.*

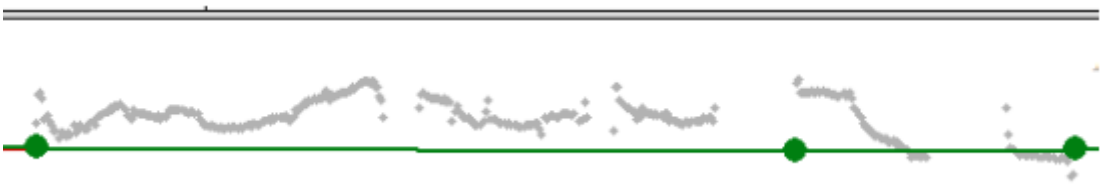
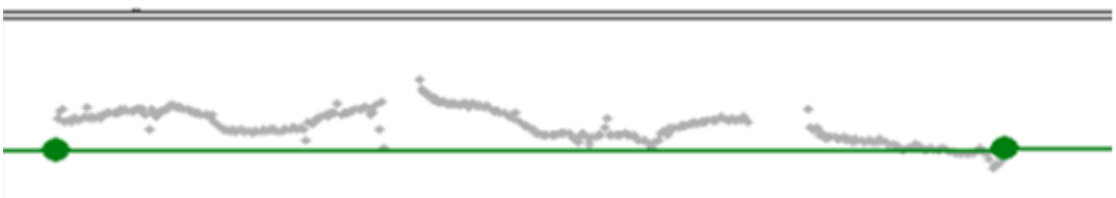


Figura 19 - Curva entoacional da *frase neutra* "A Juliana adora comer farofa."



important procedures are: (a) repetition of the same utterance by the same person and by different persons, since the comparison of various realizations of one and the same utterances eliminates most of the casual and idiosyncratic features; (b) omission, substitution and transformation — procedures which are useful in other areas of linguistic research as well since they enable us to arrive at relevant phonological features and to ascertain the phonological relations between intonational patterns." [...] (DANES, 1960, p. 38)

Figura 20 - Curva entoacional da *frase neutra* "Mariana tem medo de altura."

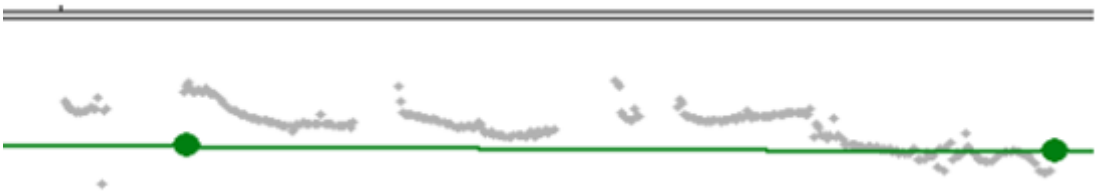


Figura 21 - Curva entoacional da *frase neutra* "Tatiana conversa com a Marisa."

A segunda manipulação foi feita de modo que obtivéssemos frases cujo intervalo entre a pretônica e o conjunto tônica/pós-tônica fosse de 7st, ou seja, exatamente a hipótese finalizadora proposta por Ferreira Netto (2008), acrescida da colaboração de Consoni et alii (2011). O processo para a criação da segunda manipulação difere da primeira somente na última etapa. Ao invés de usarmos o ícone *stylise pitch*, utilizamos *stylise pitch (2st)*. A segunda manipulação recebeu a alcunha de *frase hipótese*. Abaixo, as curvas entoacionais das *frases hipóteses*.

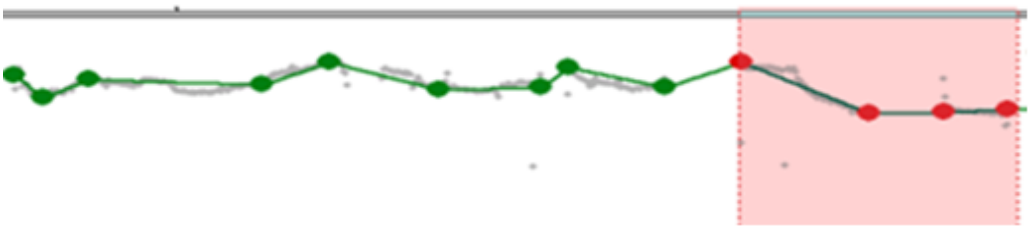


Figura 22 - Curva entoacional da *frase hipótese* "A Juliana adora comer farofa." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st.

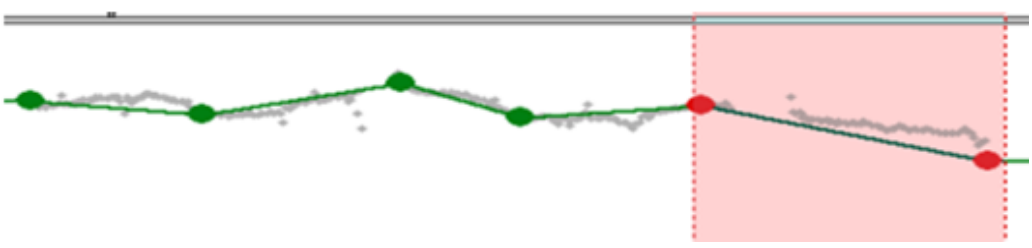


Figura 23 - Curva entoacional da *frase hipótese* "Mariana tem medo de altura." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st.

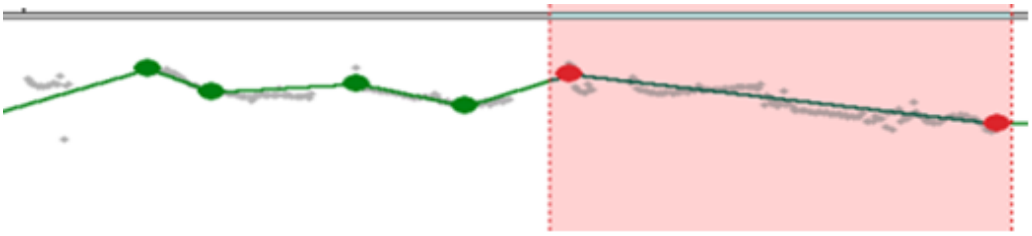


Figura 24 - Curva entoacional da *frase hipótese* "Tatiana conversa com a Marisa." A demarcação indica a manipulação seguindo a hipótese do intervalo de 7st.

6.1.2 Análise dos dados e resultados do teste 1

Na aplicação dos testes não havia nenhuma sinalização de que as frases fossem iguais ou diferentes entre si. Não havia igualmente nenhuma marcação ou numeração nas frases. Ainda, estas não foram passadas em ordem, pois a cada audição elas eram randomizadas. Com o intuito de viabilizar a eficácia dos testes, nós numeramos as frases e, por esse motivo, aqui, neste trabalho, elas aparecem assim descritas:

Frase 1: *frase modelo* " A Juliana adora comer farofa."
Frase 2: *frase neutra* "A Juliana adora comer farofa."
Frase 3: *frase hipótese* "A Juliana adora comer farofa."
Frase 4: *frase modelo* "Mariana tem medo de altura."
Frase 5: *frase neutra* "Mariana tem medo de altura."
Frase 6: *frase hipótese* "Mariana tem medo de altura."
Frase 7: *frase modelo* "Tatiana conversa com a Marisa."
Frase 8: *frase neutra* "Tatiana conversa com a Marisa."
Frase 9: f*rase hipótese* "Tatiana conversa com a Marisa."

As respostas esperadas eram a de que as *frases modelos (frase 1, frase 4 e frase 7)* e as *frases hipóteses (frase 3, frase 6 e frase 9)* fossem consideradas finalizadas. As primeiras

por possuírem um intervalo entre o TM e o TF de 4st, ou seja, o LDT inicial para o reconhecimento da finalização. E as últimas por terem sido manipuladas de modo a apresentarem a sequência finalizadora de 7st (FERREIRA NETTO, 2008). Conjecturava-se também que as *frases neutras (frase 2, frase 5 e frase 8)* fossem julgadas como não finalizadas, afinal estas são monotonais e não seguem a série da componente estruturadora de finalização. As tabelas abaixo ilustram a quantidade de sujeitos que responderam a todas as frases ***sim, a frase está finalizada*** e aqueles que responderam ***não, a frase não está finalizada.***

Padrão de resposta	**Sim, a frase está finalizada**	**Não, a frase não está finalizada**
Quantidade de sujeitos	106	74
TOTAL: 180		

Padrão de resposta em proporção	**Sim, a frase está finalizada**	**Não, a frase não está finalizada**
Quantidade de sujeitos	58%	41%
TOTAL: 100%		

Tabela 1 - Padrão de resposta em número de sujeitos e em proporção.

As tabelas *(ver tab. 1)* acima enfatizaram que o número de respostas 'sim' não coincidiu com o valor de respostas 'não'. Fez-se urgente, então, testar se essas respostas eram ou não aleatórias. Conhecida é a premissa de que testes de hipótese não paramétricos servem a tal propósito. Com o escopo de testar a já mencionada hipótese, empregamos o denominado *teste qui quadrado*. Segundo Hackbarth Neto & Stein (2003), esse tipo de teste é utilizado quando estão em comparação dois ou mais grupos de variáveis. Logo, é exequível aferir com qual frequência um evento se desvia da frequência esperada. Isto posto, propusemos duas hipóteses, a saber:

Hipótese nula – as respostas são aleatórias

Hipótese alternativa – há um padrão nas respostas

Padrão de resposta	Frequência observada (Oi)	Frequência esperada (Ei)
SIM	106	90
NÃO	74	90

Tabela 2 - Valores de frequência observada e frequência esperada para cálculo de qui quadrado.

$$Qui_{Cal} = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Figura 25 - Fórmula do cálculo de qui quadrado (retirado de HACKBARTH NETO; STEIN, 2003).

Os cálculos evidenciaram um $X^2 = 5,86$ para um $\alpha = 5\%$ e $v= 1$; $X1^2 = 3,84$. O valor crítico de X^2 é tal que $P(X^2 > 3, 84) = 5\%$. O quadro-resumo abaixo (tab. 3) exemplifica esses dados.

α = 0,05	**n= 180**
colunas (k) = 2	
linhas (r) = 2	
grau de liberdade = 1	
Fórmula com Correção de Yates	
Qui calculado (X^2) = 5,86	
Qui tabelado ($X1^2$) = 3,84	

Tabela 3 - Tabela de contingência com fórmula de correção de Yates baseada em tabela encontrada em Hackbarth Neto & Stein (2003, p. 6)

Como $X^2 = 5,86$ e $5,86 > 3,84$, rejeitamos H0, isto é, 5% de significância, podemos afirmar que há um padrão de respostas. Era preciso verificar, assim, qual seria esse padrão.

6.1.2.1 Análise do teste *two-alternatives forced choice tasks* e resultados do teste 1

É preciso sondar se os padrões de respostas das 9 frases diferem entre si, de fato. É nosso intento, neste trabalho, efetuar essa verificação por intermédio do teste estatístico inferencial *ANOVA*. Para tanto, formulamos duas hipóteses:

Hipótese nula – não há diferença entre as variáveis;

Hipótese alternativa – há diferença entre as variáveis.

Para um $\alpha = 0{,}05$ e um $\beta = 0{,}95$, encontramos um $p \leq 0{,}05$ (para $n = 180$) ANOVA F= 12,08318 para um F crítico = 1,9929. Esse dado representa a rejeição da hipótese nula e a assunção de que a hipótese alternativa seja, *a priori*, verdadeira.

RESUMO			
Grupo	*Contagem*	*Média*	*Variância*
Frase 1	20	1,2	0,168421
Frase 2	20	1,85	0,134211
Frase 3	20	1,15	0,134211
Frase 4	20	1,25	0,197368
Frase 5	20	1,8	0,168421
Frase 6	20	1,2	0,168421
Frase 7	20	1,25	0,197368
Frase 8	20	1,85	0,134211
Frase 9	20	1,2	0,168421

Tabela 4 - Quadro-resumo dos valores de média e variâncias do padrão de resposta das 9 frases.

ANOVA						
Fonte da variação	*SQ*	*gl*	*MQ*	*F*	*valor-P*	*F crítico*
Entre grupos	15,8	8	1,975	12,08318	.00000	1,9929
Dentro dos grupos	27,95	171	0,16345			
Total	43,75	179				

Tabela 5. Quadro-resumo dos valores de ANOVA.

Aplicamos, também, *teste-t* a fim de comparar as variáveis entre si. Os resultados indicaram $p \leq 0{,}05$ para todas as comparações, sendo que todos os valores de significância foram da ordem $p < 0{,}001$, para n=180.

Nesta seção nortear-nos-emos a partir dos dizeres de Dancey e Reidy (2006) que indicam, também, a necessidade de uma análise exploratória dos dados (AED), ou seja, a exploração gráfica dos dados. A técnica utilizada, aqui, é a do histograma, pois esta dá-nos visão bastante abrangente de padrões de respostas.

Nas figuras abaixo *(ver fig. 26)*, podemos visualizar o histograma de frequência dos padrões de resposta **não, a frase não está finalizada** e **sim, a frase está finalizada**.

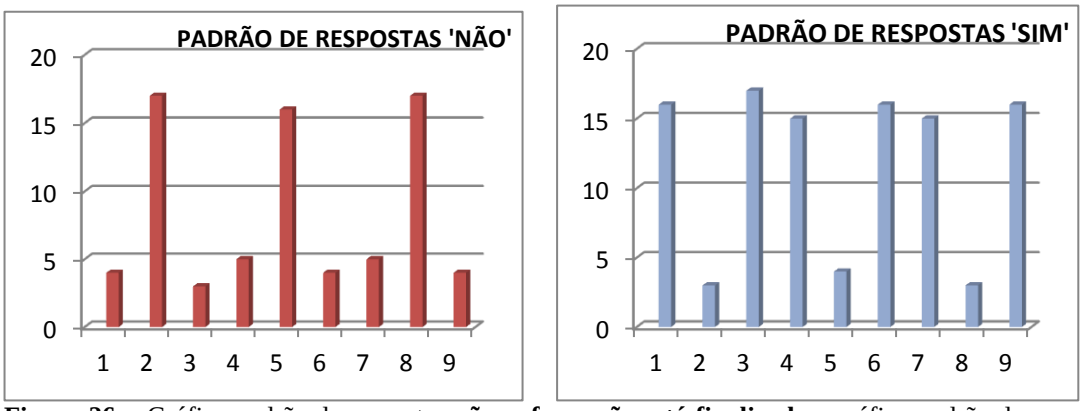


Figura 26 - Gráfico padrão de respostas **não, a frase não está finalizada** e gráfico padrão das respostas **sim, a frase está finalizada**. O eixo x representa o número de sujeitos que respondeu 'não' ou 'sim.' O eixo y representa as 9 frases.

Observando a fig. 26, é possível reparar que as *frases 2, 5 e 8* possuem as frequências de respostas 'não' mais altas. A modal do padrão de respostas 'não' encontra-se espelhada na *frase 2* com 17 sujeitos respondendo **não, a frase não está finalizada**. O valor da mediana do padrão de respostas 'não' está atrelado à *frase 4.*

No padrão de respostas **sim, a frase está finalizada**; as frequências mais altas situam-se nas *frases 1, 3, 4, 6, 7 e 9.* As *frases 1, 4 e 7* são as *frases modelos*. As *frases 3, 6 e 9* são as *frases hipóteses*. A modal do padrão de respostas 'sim' encontra-se na *frase 3*, ou seja, uma *frase hipótese*.

Outra ferramenta importante para a AED é o gráfico de caixa e bigodes ou gráfico de *box plot*. O gráfico *de box plot* fornece uma indicação clara de valores extremos. Por meio do valor da mediana, acrescido dos valores do primeiro (Q1) e do terceiro quartis (Q3), podemos calcular a amplitude interquântica, ou seja, a distância entre dois quatis (h). Estes elementos numéricos propiciam as medidas dos limites máximos e mínimos *outlier* e facilitam a construção do gráfico.

Os quartis do padrão de resposta ‘não’ são, respectivamante, 4 e 16. Os do padrão de resposta ‘sim’, são Q1= 4 e Q3 = 16 . Logo, a distância interquântica do padrão de respostas ‘não’ e o padrão de ‘sim’ é h= 12. O valor mínimo *outlier* do padrão ‘não’ é 3 e o valor máximo *outlier* é 17. O valor mínimo *e* máximo *outlier* repetem-se, ou seja, 3 e 17, respectivamente.

Ilustramos abaixo os gráficos de *box plot* do padrão de respostas ***não, a frase não está finalizada*** e, também, do padrão de respostas ***sim, a frase está finalizada***. Em nenhum dos gráficos, há valores atípicos. A análise do gráfico *box plot* denota uma grande dispersão entre os padrões de resposta ‘não’ e de resposta ‘sim.’

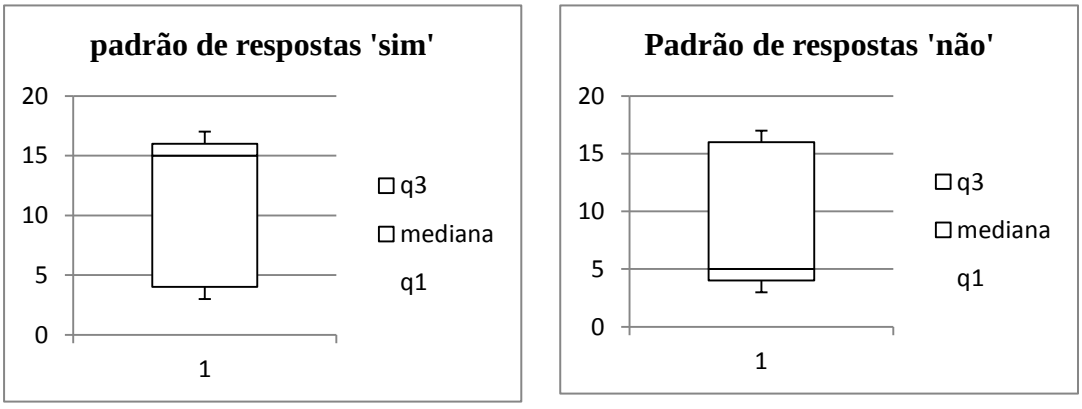


Figura 27 - Gráfico *box plot* do padrão de resposta *‘não’ (à esquerda) e do padrão de resposta ‘sim’ (à direita).*

6.1.2.1.1 Análise descritiva da percepção da *frase modelo*

Fazendo uma análise descritiva, ratificamos a premissa de Consoni (2011), qual seja, um LDT de 3 st ascendentes e 4st descendentes para a percepção inicial da finalização de uma dada frase. Escolhemos, por ora, não abordar a proposição de Martins (2012). O autor revisa o LDT do PB encontrado por Consoni (2011) e ultima sua dissertação anunciando que a variação perceptivelmente relevante para o PB é o de 3 st ascendentes e 3 st descendentes. Nas considerações finais de seu trabalho, Martins (2012) faz uma importante ressalva a respeito da diferença do LDT do PB encontrado por ele e por Consoni (2011). O autor usou voz masculina e Consoni (2011), voz feminina. Nós, na esteira do trabalho de Consoni (2011), empregamos voz feminina. Nossa escolha é de nos atermos à sequência finalizadora aventada por Ferreira Netto (2008) e ao LDT do PB de 3 st ascendentes e 4st descendentes (CONSONI, 2011).

É viável observar, nas tabelas abaixo, altas porcentagens dos sujeitos que supuseram que as *frases 1*, *4* e *7* fossem finalizadas. Esse alto índice (respectivamente, 80%, 75% e 75%) aponta para o fato de que o padrão de respostas para as *frases modelos* é o de que elas são finalizadas.

Frase 1	**sujeitos**	**proporção**
Sim	**16**	80%
Não	**4**	20%

Tabela 6 - Padrão de respostas da Frase 1 - *frase modelo* "A Juliana adora comer farofa."

Frase 4	**sujeitos**	**proporção**
Sim	15	75%
Não	5	25%

Tabela 7 - Padrão de respostas da Frase 4 - *frase modelo* "Mariana tem medo de altura."

Frase 7	**sujeitos**	**proporção**
Sim	15	75%
Não	5	25%

Tabela 8 - Padrão de respostas da Frase 7 - *frase modelo* "Tatiana conversa com a Marisa."

6.1.2.1.2 **Análise descritiva da percepção da *frase neutra***

Como já aduzido ao longo deste trabalho *(ver seção 6.1.1)*, a *frase neutra* é monotonal. As teses de Ferreira Netto (2008) e Consoni (2011) são testadas na medida em que na manipulação da *frase neutra* não foi respeitada a premissa de variação perceptivelmente relevante do intervalo 7st entre o TM e o TF. Nossa hipótese de que essa frase seria reconhecida como não finalizada é corroborada na medida em que entre 85% e 75% dos ouvintes as perceberam como tal. As tabelas abaixo indiciam essa elevada proporção.

Frase 2	**sujeitos**	**proporção**
Sim	3	15%
Não	17	75%

Tabela 9 - Padrão de respostas da Frase 2 - *frase neutra* "A Juliana adora comer farofa."

Frase 5	**sujeitos**	**proporção**
Sim	4	20%
Não	16	80%

Tabela 10 - Padrão de respostas da Frase 5 - *frase neutra* "Mariana tem medo de altura."

Frase 8	**sujeitos**	**proporção**
Sim	3	15%
Não	17	85%

Tabela 11 - Padrão de respostas da Frase 8 - *frase neutra* "Tatiana conversa com a Marisa."

6.1.2.1.3 **Análise descritiva da percepção da *frase hipótese***

A análise da percepção da finalização da *frase hipótese* nas três audições ratifica a teoria exposta por Ferreira Netto (2006, 2008). Os resultados preliminares da primeira audição indicam que 85% dos ouvintes consideraram essa frase como finalizada. Nas audições seguintes, esse número decresce, mas mantém-se em um patamar elevado (80% e 80%, respectivamente).

As três audições possuem resultados preliminares satisfatórios, posto que confirmam nossa hipótese de que o padrão perceptivo de respostas desse tipo é claramente aquele atrelado à frase finalizada.

Frase 3	**sujeitos**	**proporção**
Sim	17	85%
Não	3	15%

Tabela 12 - Padrão de respostas da Frase 3 - *frase hipótese* "A Juliana adora comer farofa."

Frase 6	**sujeitos**	**proporção**
Sim	16	80%
Não	4	20%

Tabela 13 - Padrão de respostas da Frase 6 – *frase hipótese* "Mariana tem medo de altura."

Frase 9	**sujeitos**	**proporção**
Sim	16	80%
Não	4	20%

Tabela 14 - Padrão de respostas da Frase 9 - *frase hipótese* "Tatiana conversa com a Marisa."

A análise descritiva global dos três tipos de frase permitiu-nos considerar previamente que os padrões de respostas estão atrelados ao tipo de frase apresentada para a audição. Há, portanto, uma tendência geral de perceber as *frases modelo* e *hipótese* como finalizadas e de inferir que as *frases neutras* não são finalizadas.

Consideramos ser importante uma apreciação das respostas individualizadas a fim de medir a sensibilidade de discriminação quanto à audição desses participantes (PERES, 2011).

6.2 **Teste 2**

6.2.1 Metodologia do teste 2

A fim de cotejar as respostas individuais somente das frases manipuladas *(ver capítulos 5 e 6),* solicitamos que os mesmos 20 ouvintes do teste 1 *(ver seção 6.1)* ouvissem 6 frases e respondessem quais estavam finalizadas e quais não estavam. Não fora informado aos participantes que se tratavam das mesmas frases do teste 1. Novamente tomamos o cuidado de apagar qualquer sinalização que indicassem, previamente, semelhanças ou igualdade em relação à percepção da finalização das frases. As frases eram randomizadas. Testamos a sensibilidade dos ouvintes para a percepção das finalizações de frases.

6.2.2 **A teoria da detecção de sinais**

Segundo Abdi (2007), a teoria da detecção de sinais (TDS) é usada para analisar dados provindos de experimentos nos quais é preciso categorizar estímulos ambíguos gerados por meio de sinais ou ruídos. Os sinais e ruídos podem ser tomados como ‘figuras’ para, por exemplo, a “discriminação de pares de estímulos” (PERES, 2011, p. 78). O objetivo da TDS é, então, entre outras coisas, estimar o *dee-prime (d´)*, ou seja, o valor indicativo da força de um determinado sinal, no caso deste trabalho, a força da sensibilidade perceptiva de cada

participante do teste 2. A TDS possui diferentes tipos de respostas ligadas à ausência ou presença de sinais: *hit; false alarm; correct rejection* e *miss*.

	DECISION: (PARTICIPANT'S RESPONSE)	
REALITY	Yes	No
Signal Present	Hit	Miss
Signal Absent	False Alarm (FA)	Correct Rejection

Tabela 15 - Tabela das quatro possibilidades de resposta da TDS (ABDI, 2007, p. 2).

Peres (2011) usou a mesma teoria para aferir a sensibilidade da discriminação de variantes regionais brasileiras. No delineamento de seu modelo de análise, as quatro respostas da TDS estavam assim discriminadas: "(i) *hit* é o reconhecimento da variedade pelo participante; (ii) *miss* é o não reconhecimento da variedade pelo participante; (iii) *false alarm* é o reconhecimento indevido de uma variedade e (iv) *correct rejection* é o não reconhecimento de uma variedade do qual o indivíduo não faz parte" (PERES, 2011, p. 77). Com base no que fora proposto por Peres (2011) e as possibilidades de respostas para a TDS atreladas à presença ou ausência de finalização nas frases, propusemos as seguintes hipóteses:

(i) *Hit* – resposta *sim* para presença de sinal. O participante respondeu sim, a frase está finalizada para quando a resposta esperada era a de que a frase estivesse finalizada;

(ii) *Miss* – resposta *não* para presença de sinal. O participante respondeu não, a frase não está finalizada para quando a resposta esperada era a de que a frase estivesse finalizada;

(iii) *False alarm* – resposta *sim* para ausência de sinal. O participante respondeu sim, a frase está finalizada para quando a resposta esperada era a de que a frase não estivesse finalizada;

(iv) *Correct rejection* – resposta *não* para ausência de sinal. O participante respondeu não, a frase não está finalizada quando a resposta esperada era a de que a frase não estivesse finalizada.

A tabela abaixo ilustra o delineamento do nosso modelo de análise.

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

.

Tabela 16 - Modelo de análise da sensibilidade perceptiva das finalizações de frases auferida por participante.

O *dee-prime* é dado pelo valor da inversa normal gaussiana do *hit* menos o valor da inversa normal gaussiana da taxa de *false alarm*.

$$d' = z(hit) - z\ (false\ alarm)$$

Tabela 17 - Cálculo do *dee-prime*. Sendo z o cálculo da inversa normal gaussiana (MACMILLAN; CREELMAN, 2005).

O valor obtido por meio desse cálculo permite a observação de dois parâmetros. O primeiro e mais evidente é a diferença entre as taxas de acertos (*hit*) e a taxa de alarmes falsos (*false alarm*). Uma elevada porcentagem de hit, *a priori*, ilustra que o participante conseguiu discriminar os estímulos. O segundo parâmetro é a estratégia do participante. Um participante que sempre responder não, nunca cometerá falsos alarmes. O participante que sempre responder sim, terá os acertos garantidos. Ao primeiro tipo dá-se o nome de *liberal*. Ao segundo, dá-se a alcunha de *conservador* (ABDI, 2007).

6.3.3 Análise dos dados do teste 2

Por meio da análise da tabela 18, pudemos classificar os *dee-primes* obtidos de acordo com a resposta dos sujeitos em altos, médios e baixos. Podemos, então, conjecturar que *dee-primes* altos indiciam que os sujeitos tiveram êxito em reconhecer as finalizações de frases. Os *dee-primes* médios apontam uma certa dificuldade na distinção entre frases finalizadas ou não e um *dee-prime* baixo assinala uma grande inexatidão na identificação dos tipos de frases.

Participantes	***dee-prime***	**Participantes**	***dee-prime***
Sujeito 1	0,41	**Sujeito 11**	0,41
Sujeito 2	0,41	**Sujeito 12**	2,05
Sujeito 3	2,05	**Sujeito 13**	-0,15
Sujeito 4	2,05	**Sujeito 14**	0,41
Sujeito 5	0,41	**Sujeito 15**	-0,15
Sujeito 6	0,41	**Sujeito 16**	2,05
Sujeito 7	-0,15	**Sujeito 17**	0,41
Sujeito 8	2,05	**Sujeito 18**	0,41
Sujeito 9	-0,15	**Sujeito 19**	2,05
Sujeito 10	-0,15	**Sujeito 20**	0,41

Tabela 18. Valores de *dee-prime* dos participantes do teste 2.

Peres (2011) ilustra os resultados de seu estudo por meio da curva de ROC (receiving operating characteristic), ou seja, um gráfico de sensibilidade que relaciona acertos e alarmes falsos. À luz do trabalho *O papel da prosódia na identificação das variedades regionais do português brasileiro*, nós exemplificaremos os dados do teste 2, também, por via dos gráficos de ROC.

Os sujeitos 3, 4, 8, 12, 16 e 19 obtiveram um alto valor de *dee-prime (ver anexo II)*. A tabela 19 sinaliza a totalidade de acerto por conta desses sujeitos. É patente que quanto maior o *dee-prime*, maior é a força da sensibilidade perceptiva para as finalizações de frases.

	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

Tabela 19 - Proporção e *dee-prime* das respostas dos sujeitos 3,4,8, 12, 16 e 19.

Como pode ser observado no gráfico abaixo *(ver fig. 28)*, o eixo das abscissas é nulo. Não há falsos alarmes. O eixo das ordenadas da curva ROC expõe um *hit rate* de 100%.

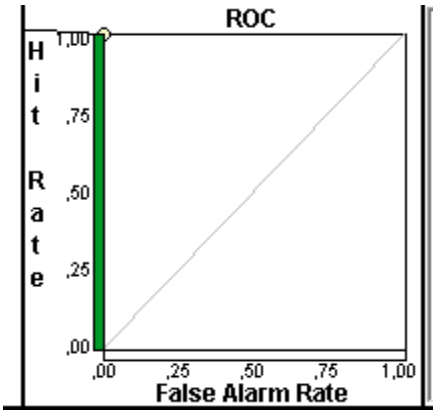


Figura 28 - Gráfico de ROC dos sujeitos 3,4,8,16 e 19.

Os sujeitos 1, 2, 5, 6, 11, 14, 17, 18 e 20 possuem um *dee-prime* de 0,41 *(ver anexo II)*. Isso indica que apesar de a proporção de acertos ter sido alta (0,99), a de rejeição correta também é (0,77). A proporção de alarmes falsos é média (0,33). Há um discernimento razoável dos dois tipos de frases.

	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

Tabela 20 - Proporção e *dee-prime* das respostas dos sujeitos 1, 2, 5, 6, 11, 14, 17, 18 e 20.

O gráfico abaixo *(ver fig. 29)* ilustra que o eixo x tem uma proporção mediana, na ordem de 0,33. O eixo y é alto (0,99). A taxa de acertos é elevada, mas o valor moderado de falsos alarmes faz com que o *dee-prime* caia.

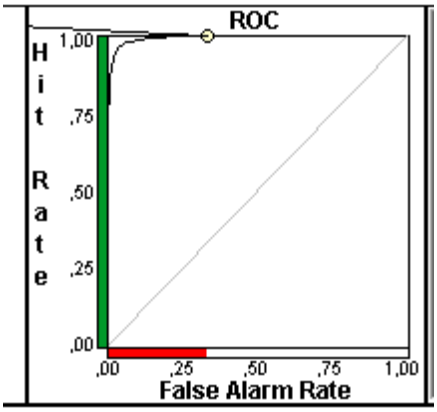


Figura 29 - Gráfico de ROC dos sujeitos 1, 2, 5, 6, 11, 14, 17, 18 e 20.

Os sujeitos 7, 9, 13 e 15 não obtiveram uma *performance* satisfatória *(ver anexo II).* O *dee-prime* negativo indica uma dificuldade, por parte desses sujeitos, em reconhecer e discriminar as finalizações de frases.

	Yes	No	d´
Signal present	0,77	0,33	-0,15
Signal absent	0,33	0,77	

Tabela 21 - Proporção e *dee-prime* das respostas dos sujeitos 7, 9, 10, 13 e 15.

O gráfico abaixo *(ver fig. 30)* torna claras as proporções de acertos (0,77) e de falsos alarmes (0,33). O eixo das abscissas mostra a proporção de acertos alta e o eixo das ordenadas demonstra a proporção módica de falsos alarmes.

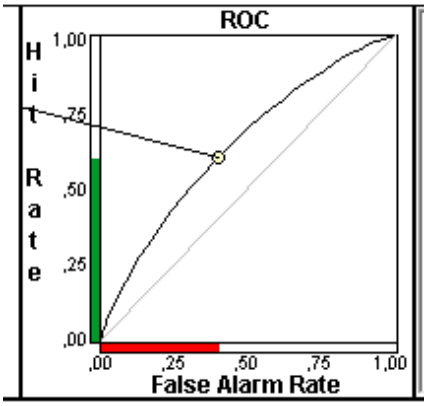


Figura 30 - Gráfico de ROC dos sujeitos 7, 9, 13 e 15.

7 DISCUSSÕES FINAIS

Os resultados do teste 1 apontam para o fato de que os parâmetros da escala de semitons percentualmente relevante formulada por Consoni *et al*. (2011) são pistas salutares para o reconhecimento das finalizações de frase assertivas isoladas do PB. Ademais, pudemos observar que a premissa formulada por Ferreira Netto (2008) é levada em conta pelo ouvinte no momento em que este julga se determinada frase está ou não finalizada, assim como enuncia Ferreira Netto (2007).

> [...] a fala tem de ser pré-programada quanto à integração entre sua duração e sua variação de freqüências componentes e delimitativas, isto é, o falante tem de prever a duração de sua frase, iniciando com tom suficientemente alto para que possa concluí-la em tom baixo (FERREIRA NETTO, 2007, p.1).

As porcentagens dos ouvintes que consideram as *frases neutras* como não finalizadas sempre foram altas e mantiveram-se, em sua grande maioria, na casa dos 80%. O produto da análise (e a elevada porcentagem dos ouvintes que consideraram essas frases finalizadas) das *frases hipóteses* corroboram os dados dos trabalhos já citados (CONSONI; FERREIRA NETTO, 2008; CONSONI *et al*., 2008, 2009; FERREIRA NETTO, 2006, 2008; ROSA, 2009).

A análise dos dados do teste 2 possibilitou que criássemos três categorias de *dee-prime*. Categorizamos a força da sensibilidade da percepção das finalizações de frases em três níveis: alto, médio e baixo. Dos 20 participantes, 75% destes permaneceram entre a categoria média e alta. Há indícios fortes para afirmarmos que houve, por parte dos participantes do teste 2, uma extrema e mediana facilidade de discriminar frases finalizadas daquelas que não são. É preciso salientar, entretanto, que um impasse acometeu 25% dos participantes do teste 2. A falta de unanimidade na percepção das finalizações descendentes de frases assertivas já vem sendo estudada há algum tempo pelo grupo de pesquisa ExProsodia. Algumas hipóteses já foram suscitadas.

Rosa (2009b) propôs um estudo no qual visava determinar se havia coincidência entre o que ouvintes julgam como final de frases em textos longos e o tom descendente da última sílaba. A referida verificação foi feita por meio de testes de percepção. Foram utilizados os dados de segmentação de frase propostos pela rotina ExProsodia®. Os resultados mostraram que o tom da última sílaba com tom grave nem sempre é um fator preponderante para o reconhecimento das terminações das frases. (ROSA, 2009b)

Em sua tese de doutorado, Consoni (2011) apresentou a repetição de alguns padrões tonais relacionados à percepção da proeminência em PB. A autora expôs o fruto de sua análise por meio de 4 níveis formados por tons de atração como ilustra figura abaixo:

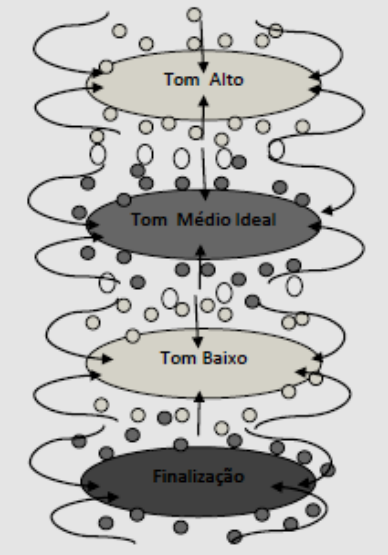


Figura 31 - Na figura, os círculos grandes representam os tons atratores, os círculos pequenos representam todas as variações que podem ocorrer na produção da entoação, as setas indicam as possíveis forças de atração para um ou outro tom. (retirado de CONSONI, 2011, p. 101).

Segundo Consoni (2011) a proeminência de 7 st seria percebida pelos falantes de PB como sendo aquela atrelada à finalização da frase, mas com a ressalva de que a duração e a intensidade também poderiam ser usadas para dar ênfase à marca de proeminência. Assim, ao analisar falta de unanimidade na percepção dos ouvintes no que concerne à componente estruturadora de finalização, deveríamos observar atentamente a informação de que o "aumento da duração na sílaba proeminente em PB, por exemplo, poderia ser um recurso para tornar mais evidente a ascensão ou queda tonal." (CONSONI, 2011, p. 104). À luz dessa informação,

podemos conjecturar que a presença de uma pausa subsequente ao tom descendente poderia estabelecer uma percepção unânime de finalização.

Em síntese, é conveniente dizer, portanto, que, por meio destes testes, obtivemos resultados promissores que, primeiro, corroboram as hipóteses de Ferreira Netto (2006, 2008) e Consoni (2011) e que nos auxiliam sobremaneira na árdua missão de analisar as diversas (e possíveis) noções impressionísticas da entoação (Daneš, 1960) e, depois, que esses resultados poderão ser úteis para o aprimoramento da análise automática da entoação.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDI. H. Signal Detection Theory (SDT) In: NEIL SALKIND (Ed.). *Encyclopedia of Measurement and Statistics.* Thousand Oaks (CA): Sage, 2007. 1416 p.

ANDRADE, L.M de O. *Determinação dos limiares de normalidade dos parâmetros acústicos da voz.* Dissertação (Mestrado) - Universidade de São Paulo, 2003.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Rev . Est. Ling .*, Belo Horizonte, v . 20, n. 1, p. 11-27, jan./jun. 2012.

BARRY, J.; ANDREEVA, B. Cross-language similarities and differences in spontaneous speech patterns. In: KOHLER, K.J. ; SIMPSON, A.P. (edt). *Journal of the International Phonetic Association*, vol. 31, p. 127 – 138, Cambridge University Press, 2001.

BASSET, J.; AMELOT, A.; VAISSIÈRE, J.; ROBEAU, B. Nasal airflow in French spontaneous speech. In: KOHLER, K.J. ; SIMPSON, A.P. (edt). *Journal of the International Phonetic Association*, vol. 31, p. 87 – 100, Cambridge University Press, 2001.

BAZ, D.G.M; FERREIRA NETTO, W. *Variação tonal na finalização de frases de PB.* Texto apresentado no III COLÓQUIO BRASILEIRO DE PROSÓDIA DA FALA. UFMG: Belo Horizonte, 2011. Disponível em http://www.academia.edu/2272642/Varia%C3%A7%C3%A3o_tonal_na_finaliza%C3%A7%C3%A3o_de_frases_de_PB . Acesso em 12 Nov. 2013.

BOERSMA, P.; WEENINK, D. *Praat:* doing phonetics by computer (Version 5.3.57) [Computer program], 2011. Disponível em <http://www.praat.org). Acesso em: 25 Out. 2013.

BOLINGER, D. Around the edge of language: intonation. *Harvard Education Review*, Cambridge, v. 34, n.2, p. 282-293, 1964.

_____. Intonation in American English. In: *Intonation systems: a survey of twenty languages.* HIRST, D. and DI CRISTO, A. (ed.). London: Cambridge University Press, p. 45-55, 1998.

CAGLIARI, L. C. *Elementos de fonética do português brasileiro.* 1982. Tese (Livre-Docência) – IEL, UNICAMP, Campinas, 1982.

_____. Prosódia: algumas funções dos supra-segmentos. *Cadernos de Estudos Linguísticos 23*: 137-151, Campinas, 1992.

CONSONI, F.; FERREIRA NETTO, W. A percepção de variação em semitons ascendentes para falantes de Português Brasileiro em palavras isoladas. *X Anais do Congresso Nacional de Fonética e Fonologia/IV Congresso Internacional de fonética e Fonologia*, Niterói, 2008.

CONSONI, F. ; ROSA, R. C. M. ; FERREIRA NETTO, W . A questão da correlação entre a análise automática das finalizações prosódicas e a separação intuitiva de frases em textos longos. In: IV CONGRESSO INTERNACIONAL DE FONÉTICA E FONOLOGIA, 2008, Niterói: UFF *Caderno de Resumos*, p. 14-14, 2008.

CONSONI, F.; FERREIRA NETTO, W.; PERES, D. O.; LASSAK, A. de L.: ROSA, R. C. M. Sensitivity to f0 variation in Brazilian Portuguese. *Book of abstracts of 40th Poznan Linguistic meeting*, Poznan, Polônia, 2009.

CONSONI, F.; FERREIRA NETTO, W.; PERES, D.O. *Finalisation des phrases en lecture et en parole spontanée: le cas du Portugais brésilien.* Texto apresentado no XXVIII JOURNÉES DES ÉTUDES SUR LA PAROLE. JEP, p. 217-220, 2010.

CONSONI, F. *Aspectos da percepção da proeminência tonal em português brasileiro.* Tese (Doutorado) - Universidade de São Paulo, 2011.

COSTA, N.S.A. *Variações entoacionais na língua portuguesa falada por mulheres guatós.* Tese (Doutorado) – Universidade de São Paulo, 2010.

CRYSTAL, D. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press, 1976. 390 p.

DANEŠ, František. Function of sentence intonation. *Word.* v. 16, n. 1, p. 34-54. 1960.

DANCEY, C.P.; REIDY, J. *Estatística sem matemática para Psicologia.* Porto Alegre: Artmed – Bookman, 2006. 608 p.

DI CRISTO, A.. *Interpreter la prosodie*. XXIIIEMES JOURNEES D'ETUDE SUR LA PAROLE, Aussois, 19-23 juin 2000.

DUCROT, O.; TODOROV, T. *Dictionnaire encyclopédique des sciences du langage.* Paris: Éditions du Seuil, 1972.

FALÉ, I.; FARIA, I. H. Percepção categorial de contrastes entoacionais em Português Europeu. In: CONGRESSO DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA, XXI. *Actas...* Porto, 2006, p. 341-348.

FERREIRA NETTO, W. *Variação de freqüência e constituição da prosódia da língua portuguesa.* Tese (Livre-Docência) – Universidade de São Paulo, 2006.

_____. *Declinação frasal na língua portuguesa.* Texto apresentado no 55° SEMINÁRIO GEL, Franca, São Paulo, 2007.

_____. *Decomposição da entoação frasal em componentes estruturadoras e semântico-funcionais.* Texto apresentado no X CONGRESSO NACIONAL DE FONÉTICA E FONOLOGIA/IV CONGRESSO INTERNACIONAL DE FONÉTICA E FONOLOGIA, Niterói, RJ, 2008.

_____. ExProsodia. Proc. 8992-2. *Revista da Propriedade Industrial - RPI*, n. 2038, p. 167, 26 de janeiro de 2010.

FERREIRA NETTO, W. ; MARTINS, M.V.M. ; PERES, D. O. ; ROSA, R. C. M. ; MEDEIROS, J. . Brazilian Portuguese intonation: A comparison between automatic and perceptual analyses. *The Journal of the Acoustical Society of America*, v. 132, p. 2004, 2012.

FRIEDERICI, Angela D; ALTER, Kai. Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, n. 2, v. 89, p. 267-276, 2003.

GREEN, D.M.; SWETS, J.A. *Signal Detection Theory and Psycophysics*. New York: Wiley, 1966. 521p.

HALLIDAY, M.A.K. *A course in spoken English*: Intonation. London: Oxford University Press, 1970. 142 p.

HACKBARTH NETO, A.A.; STEIN, C.E. *Uma abordagem dos testes não-paramétricos com utilização do excel.* Mini curso de estatística. Instituto de Matemática UFRGS, 2003. Disponível em http://www.mat.ufrgs.br/~viali/estatistica/mat2282/material/textos/artigo_11_09_2003.pdf. Acesso em 07 Dez. 2013

HELLIGE, Joseph B . Spécialisation hémisphérique: données récentes. *Revue de Neuropsychologie*, n. 1, v. 12, p. 7-49, 2002.

HIRST, D. & DI CRISTO, A. *Intonation Systems A Survey of Twenty Languages.* (ed.). London: Cambridge University Press. 1998. 500 p.

HIRVONEN, P. Finnish and English communicative intonation. Turk: University of Turku, 1970.

HUCKVALE, M.A.; BROOKES, D.M; DWORKINK, L.T.; JOHNSON, M. E.; PEARCE, D.J.; WHITAKER, L. *The SPAR Speech Filing System.* In: EUROPEAN CONFERENCE ON SPEECH TECHNOLOGY, 1987. Endinburgh. Disponível em: http://www.phon.ucl.ac.uk/home/mark/papers/sparsfs87.pdf . Acesso em: 15 de Out.2013.

KUHL, P.K.; IVERSON, P. Linguistic experience and the "perceptual magnet effect". In: STRANGE, W. (ed.). *Speech perception and linguistic experience: issues in cross-language research.* Baltimore: York Press, p. 121-154 , 1995.

KUHL, P. K. A new view of language acquisition. *Proceedings of the National Academy of Sciences*, n. 97, v. 22, p. 1850 -1857, 2000.

KUHL, P.K.; TSAO, F.-M.; LIU, H.-L.; ZAHNG, Y.; DE BOER, B. Language/Culture/Mind/Brain. Progress at the margins between disciplines. *Annals of New York Academy of Sciences*, n. 935, p. 136-74, 2001.

LADD, R. *Declination:* A Review and Some Hypotheses. Cambridge University Press Vol. 1, p. 53-74, 1984.

LAGUITTON, Virginie *et al.* Pitch perception: a difference between right and left handed listeners. *Neuropsychologia*, n. 3, v. 36, p. 201-207, 1997.

LENNES, M.; ALAROUTO, N.; VAINIO, M. Is the phonetic quality of unaccented words unpredictable? An example from spontaneous Finnish. In: KOHLER, K.J. ; SIMPSON, A.P. (edt). *Journal of the International Phonetic Association*, vol. 31, p. 127 – 138, Cambridge University Press, 2001.

LIBERMAN, M.Y. *The intonation system of English.* Doctoral Dissertation, MIT, 1975.

MACMILLAN, N.A.; CREELMAN, C.D. *Detection Theory*: a user´s guide. New Jersey: LEA, 2005. 512 p.

MAEDA, S. *A characterization of American English intonation.* Doctoral Dissertation, MIT, 1976.

MARTINS, M.V.M. *Aspectos da percepção e do controle entoacional do Português Brasileiro.* Dissertação (Mestrado) – Universidade de São Paulo, 2012.

MORAES, J. A. Intonation in Brazilian Portuguese. In: *Intonation systems: a survey of twenty languages***.** HIRST, D. and DI CRISTO, A. (ed.). London: Cambridge University Press, 1998, p. 179-194.

_____. F0 Declination In Brazilian Portuguese In Read And Spontaneous Speech. In: *Proceedings of the XIV International Congress of Phonetic Sciences*, Rio de Janeiro: ICPh, p. 2323-2326, 1999.

_____. *Intonational Phonology of Brazilian Portuguese*, paper apresentado no Workshop on INTONATIONAL PHONOLOGY: UNDERSTUDIED OR FIELDWORK LANGUAGES.ICPHS, Satellite Meeting, Saarbrücken, 2007. Disponível em http://www.linguistics.ucla.edi/people/jun/Workshop2007ICPhS/Moraes-BR.pdf. Acesso em: 18 Out. 2013.

NESPOR, M.; VOGEL, I. *Prosodic Phonology.* Dordrecht: Foris Publications, 1986.

_____. *Prosodic Phonology*, 2nd edition. Berlin/New York: Mouton de Gruyter, 2007.

NICOLAIDIS, K. An electropalatografia study of Greek spontaneous speech. In: KOHLER, K.J. ; SIMPSON, A.P. (edt). *Journal of the International Phonetic Association,* vol. 31, p. 67 – 86, Cambridge University Press, 2001.

PERES, D. O. *O papel da prosódia na identificação das variedades regionais do português brasileiro.* Dissertação (Mestrado) - Universidade de São Paulo, 2011.

PIERCE, John R. *The science of musical sound.* New York: Scientific American Books,1983. 270 p.

PIERREHUMBERT, J. The perception of fundamental frequency declination. *Journal of the Acoustical Society of America*, 66, p. 363–369, 1979.

_____. *The Phonology and Phonetics of English Intonation.* Tese (Doutorado) - Massachusetts Institute of Technology, 1980.

PIKE, K.L. On the Grammar of Intonation. *Proccedings of the Fifth International Congress of Phonetic Sciences*. Separatum. Münster, 1964.

ROSA, R. C. M. *Percepção de prosódia:* finalização de frases. Texto apresentado no 17° SIICUSP, São Paulo, SP, 2009.

_____. *Destros, canhotos e a percepção de procedimentos de segmentação frasal com propósito de finalização*. Texto apresentado no GEL 57, 2009, Ribeirão Preto. Seminário do GEL 57, 2009a.

_____. *Estudo da seleção de lugares para marcas de finalização.* Texto apresentado no THE 7TH BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY / I WORKSHOP DE INICIAÇÃO CIENTIFICA EM TECNOLOGIA DA INFORMAÇÃO E DA LINGUAGEM HUMANA, 2009, São Carlos. CD-ROM Stil, 2009b.

_____. *Leitura e fala espontânea: fronteiras prosódicas finais de sentenças gravadas com voz feminina.* Texto apresentado no SEMINÁRIO DO GEL, 58., 2010, São Carlos (SP): GEL, 2010. Disponível em: <http://www.gel.org.br/?resumo=6991-10>. Acesso em: 10 jun. 2013.

SELKIRK, E. O., 1984. *Phonology and Sintax:* The Relation between Sound and Structure. Cambridge: The M.I.T. Press, 1984.

SILVA, C. L. M. *Fala espontânea e leitura oral no português do Brasil: comparação* por meio de análise acústica. Dissertação (Mestrado) - USP, 2009.

SILVERMAN, K.; BECKAMN, M.; PITRELLI, J.; OSTENDORF, WIGHTMAN, C.M.; PRICE, P.; PIERREHUMBERT, J.; HIRSCHEBERG, J. Tobi: a standard for labeling English prosody. *Proceedings ICSLP*, n. 2, p. 867-870 , 1992.

SILVERSTEIN, R. *A Strategy for utterance production in hausa*. Stud. Afr. Linguist.; 1976.

TENANI, L. Domínios prosódicos no Português Brasileiro: evidências rítmica, entoacional e segmental. *Revista Estudos Lingüísticos XXXV*, p. 118-131, 2006.

T´HART, J. Differential sensitivity to pitch distance, particulary in speech. *Journal of Acoustical Society of America*, n. 69, v. 3, p. 811-821, 1981.

T ´HART; J.; COLLIER, R; COHEN A. *A perceptual study of intonation.* Cambridge: Cambridge University Press, 1990. 232 p.

T´HART, J. Intonation in dutch. In: *Intonation systems: a survey of twenty languages.* HIRST, D. e DI CRISTO, A. (ed.). London: Cambridge University Press, 1998, p. 96-111.

TREMBLAY, Tania *et al*. Phonological processing of words in right- and left-handers. *Brain and Cognition*, n. 5, v. 55, p. 427-432, 2004.

TROUBETZKOY, N.S. *Principes de phonologie.* Trad. de J. Cantineau. Paris: Klincksieck, 1976 (1939). 396 p.

VAISSIÈRE, J. *Contribution à la synthèse par règlesdu français.* Tese (Doutorado) – Université des Langues et Lettresde Grénoble, 1971.

_____.Sur les universaux de substance prosodiques. In: WAUQUIER, Sophie (Ed.) *Les universaux sonores.* Nantes: Presses Universitaires de Rennes, 2002. Disponível em: www.personnels.univparis3.fr/users/vaissier/pub/ARTICLES/index_fichiers/2005. pdf›. Acesso em: dez. 2013.

_____. Perception of intonation. In: *The handbook of speech perception*. PISONI, D.; REMEZ, R. Blackwell publisinhg, 2005. 724 p.

WELMERS, W.E. African Languages Structures. University of California, Berkeley, 1973.

WHEELDON, L. Generating prosodic structure. In: *Aspects of language production.* Hove, UK: Psychology Press, 2000. 431 p.

XU, Yi; WANG, Q. E. *Component of intonation: what are linguistic, what are mechanical/physiological?* Presented at INTERNATIONAL CONFERENCE ON VOICE PHYSIOLOGY AND BIOMECHANICS, Evanston Illinois, 1997.

ANEXO I

Questionário de identificação

Leia o texto abaixo e preencha o questionário:

Aceito participar, voluntariamente, de pesquisa referente à dissertação de mestrado na área de fonética e fonologia (FFLCH/USP) da mestranda Renata Cezar de Moraes Rosa. A participação se resume a responder testes de percepção de prosódia do Português Brasileiro.

Inciais do nome:_____ Idade: _____

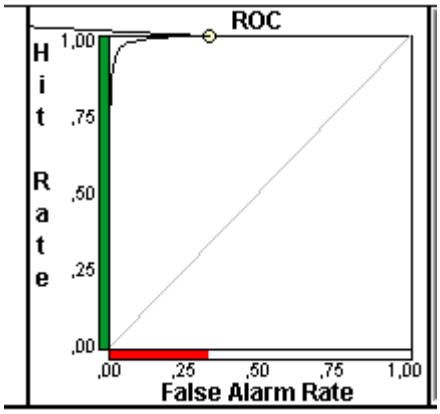
Sexo: ☐ feminino ☐ masculino

Nacionalidade: ☐ Brasileira ☐ Outra_____

Escolaridade: ☐ ensino médio ☐ ensino superior

ANEXO II

Gráfico de ROC e proporção de *dee-prime* das respostas de todos os participantes do teste 2

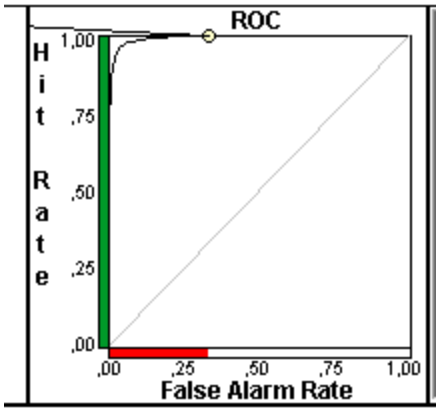


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 1

Sujeito 1	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

SUJEITO 1

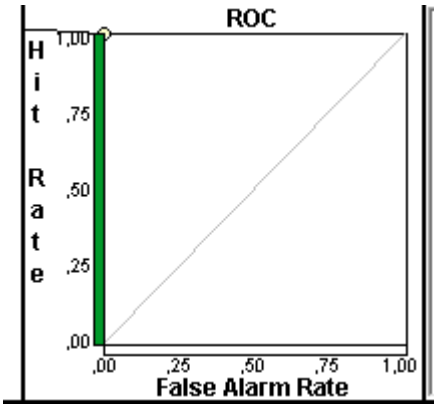


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 2

Sujeito 2	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

SUJEITO 2

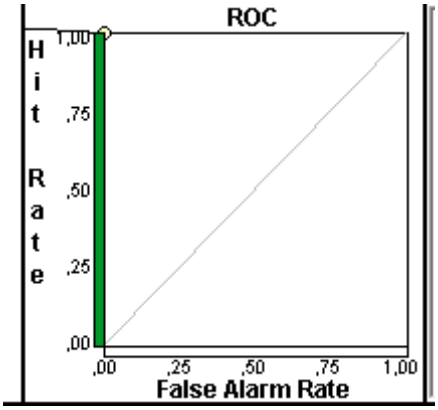


SUJEITO 3

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 3

Sujeito 3	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

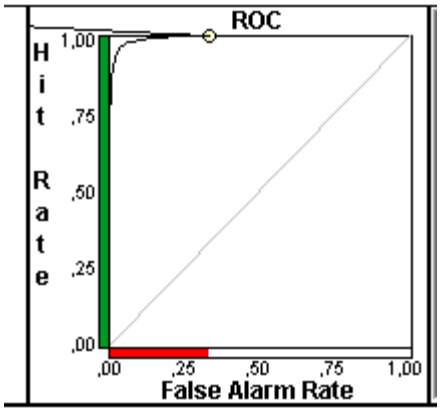


SUJEITO 4

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 4

Sujeito 4	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

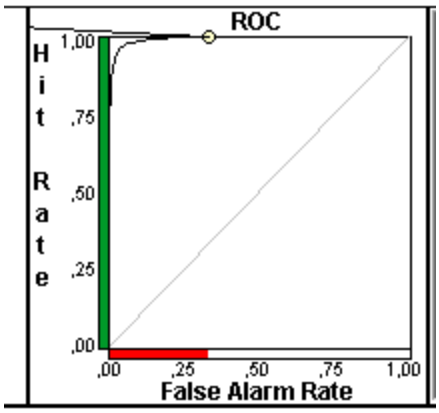


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 5

Sujeito 5	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

SUJEITO 5

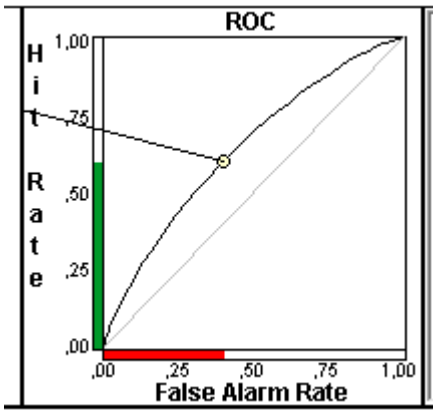


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 6

Sujeito 6	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

SUJEITO 6

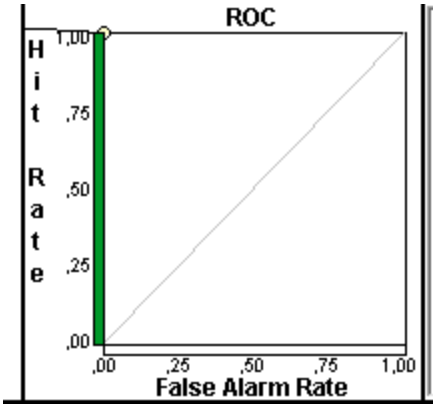


SUJEITO 7

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 7

Sujeito 7	Yes	No	d´
Signal present	0,77	0,33	-0,15
Signal absent	0,33	0,77	

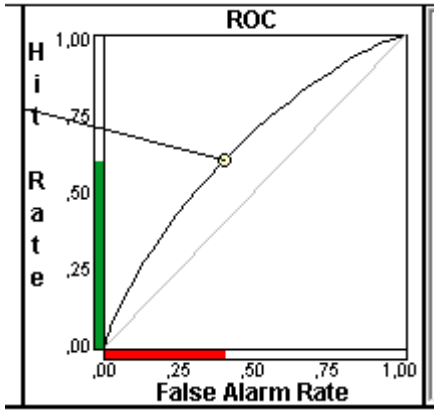


SUJEITO 8

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 8

Sujeito 8	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

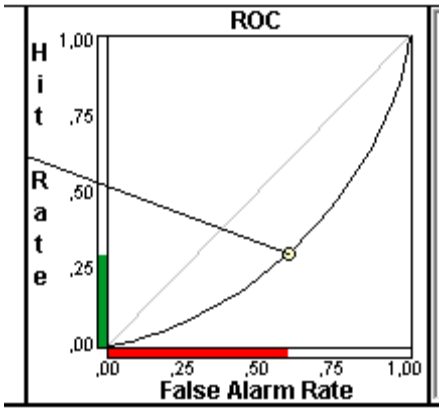


SUJEITO 9

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 9

Sujeito 9	Yes	No	d´
Signal present	0,77	0,33	-0,15
Signal absent	0,33	0,77	

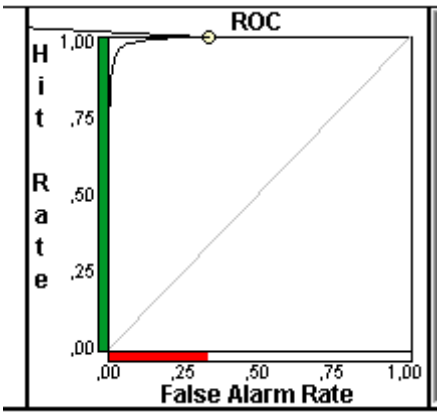


SUJEITO 10

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 10

Sujeito 10	Yes	No	d´
Signal present	0,33	0,01	---
Signal absent	0,77	0,99	

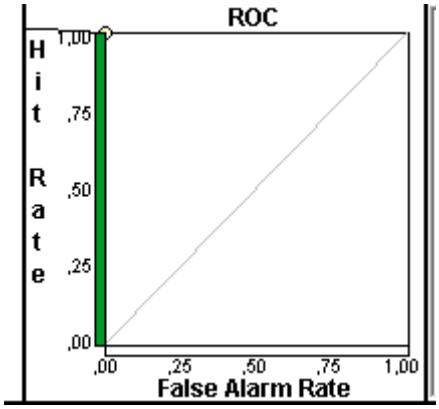


SUJEITO 11

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 11

Sujeito 11	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

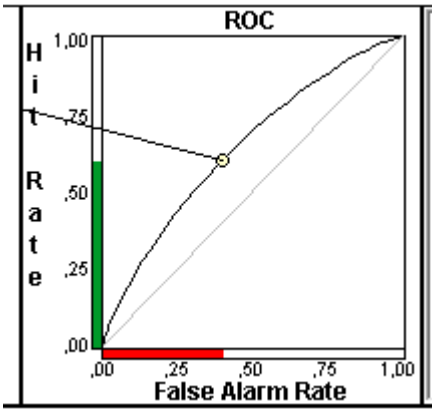


SUJEITO 12

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 12

Sujeito 12	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

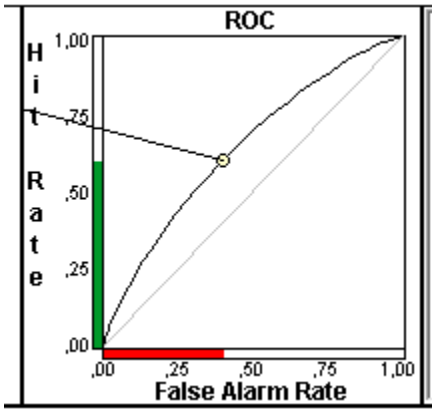


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 13

Sujeito 13	Yes	No	d´
Signal present	0,77	0,33	-0,15
Signal absent	0,33	0,77	

SUJEITO 13

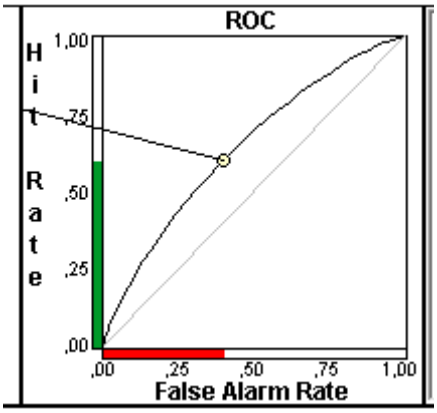


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 14

Sujeito 14	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

SUJEITO 14

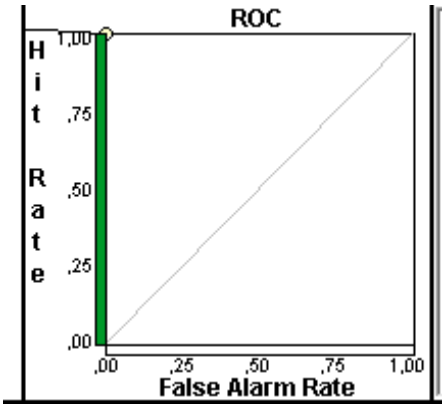


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 15

Sujeito 15	Yes	No	d´
Signal present	0,77	0,33	-0,15
Signal absent	0,33	0,77	

SUJEITO 15

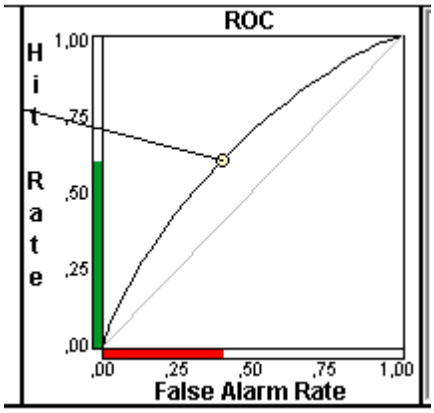


Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 16

Sujeito 16	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	

SUJEITO 16

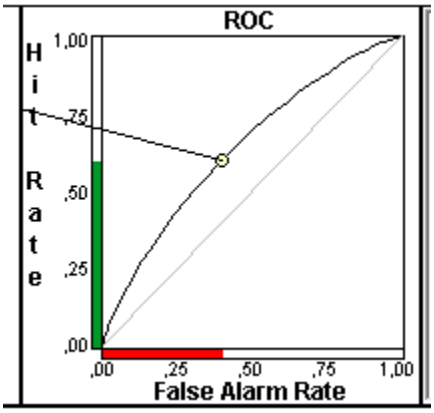


SUJEITO 17

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 17

Sujeito 17	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

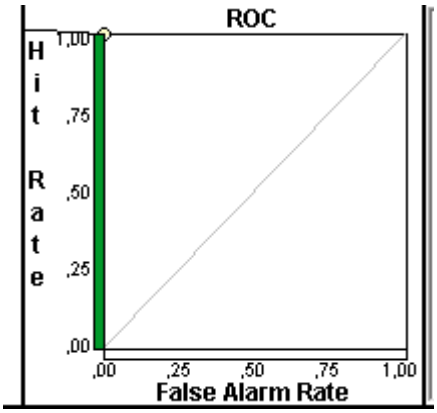


SUJEITO 18

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 18

Sujeito 18	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	

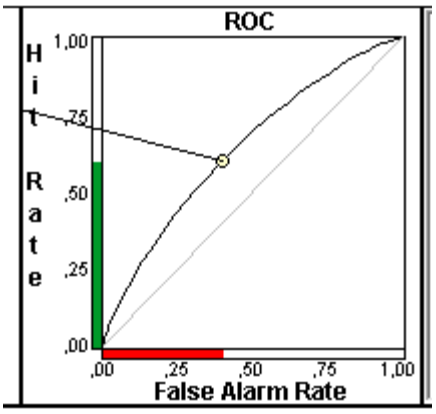


SUJEITO 19

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 19

Sujeito 19	Yes	No	d´
Signal present	0,99	0,01	2,05
Signal absent	0,01	0,99	



SUJEITO 20

Realty	yes	No
Signal Present	Hit	Miss
Signal Absent	False alarm	Correct rejection

Proporção sujeito 20

Sujeito 20	Yes	No	d´
Signal present	0,99	0,01	0,41
Signal absent	0,33	0,77	