

THE PERCEPTION OF NEUTRAL DECLARATIVES AND TOTAL QUESTIONS BY BRAZILIAN SPEAKERS FROM CURITIBA: IDENTIFICATION AND DISCRIMINATION TESTS

PERCEPÇÃO DIALETAL DE DECLARATIVAS NEUTRAS E INTERROGATIVAS TOTAIS PRODUZIDAS POR FALANTES CURITIBANOS: TESTES DE IDENTIFICAÇÃO E DE DISCRIMINAÇÃO

Pollianna MILAN

Universidade Federal do Paraná (UFPR)

Denise Cristina KLUGE

Universidade Federal do Paraná (UFPR)

RESUMO

O presente artigo investiga o reconhecimento ou não do dialeto de Curitiba (Paraná) a partir de dois testes de percepção – um de identificação e outro de discriminação. Primeiramente gravamos dois curitibanos e depois outros 20 curitibanos julgaram se os enunciados destes dois primeiros informantes seriam reconhecidos como do falar de Curitiba, em contraste (ou não) com variedades catarinenses. Os resultados revelaram que o teste de discriminação se mostrou mais difícil aos ouvintes do que o teste de identificação.

ABSTRACT

The aim of this research is to evaluate the recognition or not of the dialect from Curitiba (Paraná) based on two perception tests – an identification and a discrimination test. First, we recorded two people from Curitiba's dialect and, after, another 20 people of Curitiba judged if the statements of the two informants would be recognized as the dialect of Curitiba, in contrast (or not) with Santa Catarina varieties. The results indicated that, for the listeners, the discrimination test was harder than the identification test.

PALAVRAS-CHAVE

Percepção da prosódia curitibana. Assertivas neutras. Interrogativas totais. Identificação. Discriminação. Amper-Por.

KEYWORDS

Perception of prosody of people from Curitiba. Neutral statements. Neutral yes/no questions. Identification. Discrimination. Amper-Por.

Introdução

A variação da melodia faz parte de qualquer língua do mundo, pois nenhuma é falada em monotom. É dizer, por isso, que línguas das mais diversas origens usam a entoação de distintas maneiras a fim de marcar intenções. Para termos uma ideia do que isso significa, citamos Austin (1962, apud Moraes e Rilliard 2014) que encontrou mais de mil diferentes elocuições a partir de um estudo feito sobre os verbos performativos do inglês – sendo que muitas destas eram tipicamente entoacionais. Ou seja, são as alterações de tom, ora ascendentes ora descendentes, que levam o ouvinte a atribuir diferentes sentidos ao que ouve. Para Moraes e Rilliard (2014), “o contorno de *pitch* é tido como um elemento que é necessário, e frequentemente suficiente, para estabelecer a real força de elocução de

um enunciado”¹ (p.234). Esses tons usados pelo locutor para expressar suas intenções são classificados de distintas maneiras por estudiosos da prosódia, porém nem sempre com o mesmo entendimento.

Para Moraes (1984, p.13), as alternâncias de tom fazem parte das múltiplas funções da entoação, que poderiam ser divididas basicamente em dois grupos: o linguístico ou gramatical; e o expressivo, que corresponde à exteriorização de emoções do falante assim como de suas atitudes. No grupo gramatical estariam, por exemplo, as modalidades linguísticas (declarativas, interrogativas, exclamativas) e no grupo expressivo estariam as emoções e atitudes. Estas últimas classificadas, posteriormente, por Moraes e Rilliard (2014) e Aubergé (2002) em outros subgrupos.

Os autores supracitados explicam que o grupo das atitudes geralmente se refere a expressões através da prosódia de diferentes tipos de estados afetivos, controlados conscientemente pelo falante. Estas atitudes seriam subdivididas de duas maneiras: as proposicionais como a ironia, a incredulidade, a confirmação, o pedido; e as atitudes sociais como a arrogância, a polidez, a autoridade, a sedução. Já a expressão das emoções (muitas vezes reconhecida facialmente) é mais espontânea e universal, segundo os autores, tais como o medo, a raiva e a felicidade.²

Isso não quer dizer, porém, que estas funções são opostas entre si. No entendimento de Bolinger (1985, p.34), elas são compartilhadas, pois o locutor pode fazer uma pergunta (função gramatical, de modalidade) adicionando a esta sentença elementos de surpresa ou de queixa (função expressiva), o que quer dizer que as funções de modalidade e expressiva podem ser usadas na prosódia sobre o mesmo material segmental, em uma mesma sentença.

¹ The pitch contour is seen as an element that is necessary, and often sufficient, to establish the actual illocutionary force of the utterance (tradução nossa).

² O trabalho de Fónagy (1993) aponta para a dificuldade de distinguir o que é atitude do que é emoção. O autor define atitude como um comportamento determinado conscientemente, controlado pelo falante; opondo-se à emoção, definida como uma descarga espontânea de uma tensão psíquica, não controlada pelo falante.

Neste estudo, porém, nos atemos apenas à função gramatical/de modalidade da entoação, especificamente das asserções neutras (que são ditas sem o comprometimento do locutor, quando não há juízo de valor) e das interrogativas totais (que comportam resposta sim/não). Estas duas modalidades também variam prosodicamente conforme as línguas e, ainda, numa mesma língua, podem variar segundo o dialeto.

Neste artigo, então, a abordagem perceptual está restrita ao papel da prosódia como pista para o reconhecimento dialetal – especificamente o de Curitiba – a partir da audição das modalidades de asserção neutra e interrogativa total.

No Brasil, pelo nosso conhecimento, existem poucos estudos sobre a percepção dialetal do português brasileiro (doravante PB) e, os que existem, ainda são incipientes. Citamos, contudo, três pesquisas que abordam o mesmo tema proposto neste artigo, apesar de utilizarem metodologias diferentes.

Peres (2011) estudou a percepção prosódica de três dialetos brasileiros: o de São Paulo (capital, especificamente do bairro da Mooca), o de Pelotas (RS) e o de Senador Pompeu (CE).

Ele aplicou dois testes de identificação perceptual para ouvintes das três localidades. No primeiro teste, o autor eliminou do sinal acústico os segmentos produzidos (sentenças curtas e longas), restando apenas as características prosódicas (enunciados “sem palavras”). No segundo, foram eliminadas as variações prosódicas, restando apenas os segmentos em uma curva monotônica (sem melodia).

Os estímulos dos testes foram criados a partir da gravação de enunciados de 3 informantes masculinos de cada uma das 3 localidades. Os informantes tinham entre 30 e 62 anos, com 2º grau completo ou ensino superior completo. Ao todo, foram selecionados 6 trechos de fala de cada informante (3 curtos – de até 10 segundos – e 3 longos – de 15 a 25 segundos).

Os dois testes de percepção contaram com 39 ouvintes (13 de cada localidade), idade variando entre 20 e 55 anos. No primeiro teste, os ouvintes escutaram 54 estímulos (com a mescla dos 3 dialetos aleatoriamente), resultando em 2.106 respostas. No segundo teste foram ouvidos 18 estímulos (com a mescla dos 3 dialetos aleatoriamente), resultando em 702 respostas.

Cada ouvinte deveria identificar, em ambos os testes propostos por Peres, os estímulos que pertenciam ao seu dialeto e os que pertenciam a outros dialetos brasileiros. Os resultados dos dois testes mostraram um alto índice de identificação dialetal, sendo que o segundo teste teve um percentual de acertos ainda maior que o primeiro, em razão de o primeiro ser mais difícil no reconhecimento do dialeto por não ter segmentos nos enunciados (apenas informação prosódica).

O segundo teste teve uma média de acertos de 75%. O primeiro teste apresentou os seguintes resultados: ouvintes de Pelotas (RS) reconheceram 67% (468 acertos) dos estímulos dos pelotenses; ouvintes de São Paulo reconheceram 69% (483 acertos) dos estímulos dos paulistas; e ouvintes de Senador Pompeu (CE) identificaram 63% (443 acertos) dos estímulos dos senadorenses. Testes estatísticos aplicados por Peres mostraram que o número de acertos dos paulistas é significativamente maior que o de acertos dos pelotenses e senadorenses.

Constantini (2014) estudou as características prosódicas de 18 dialetos brasileiros e, além da análise de produção, que não nos interessa para efeito de comparação com nosso estudo, ela também aplicou um teste de percepção dialetal. A finalidade era saber se ouvintes seriam capazes de identificar a origem dos falantes de acordo com seus sotaques.

As falas usadas para a elaboração do teste perceptual de Constantini são de um banco de dados da Polícia Federal de Brasília, que tem como intenção promover mais estudos na área de fonética forense no Brasil. As gravações foram feitas por 35 policiais federais, todos homens, idade média de 35 anos, ensino superior completo, de 7 regiões brasileiras: 6 de São Paulo (1 da capital, 2 de Piracicaba, 1 de Jundiaí, 1 de Ribeirão

Preto e 1 de Campinas), 6 de Minas Gerais (todos de Belo Horizonte), 6 do Rio de Janeiro (5 da capital e 1 de Niterói), 4 do Paraná (2 de Curitiba, 1 de Ponta Grossa e 1 de São Mateus do Sul), 5 do Distrito Federal, 6 da região Nordeste (2 de Natal, 2 de Fortaleza, 1 de João Pessoa e 1 de Aracaju) e 2 da Região Norte (1 de Manaus e 1 de Porto Velho).

O teste de identificação foi montado com 75 estímulos e contou com a participação de 20 ouvintes, resultando em 1.500 respostas. Ele funcionou da seguinte maneira: um carioca, por exemplo, ouvia os 12 sotaques e deveria dizer de onde era cada um deles, se do Paraná, de Minas Gerais, do Norte e assim por diante. A autora não forneceu mais detalhes sobre a metodologia do teste.

Os resultados mostram que o maior índice de acertos foi do Rio de Janeiro (77,9%) e os menores do Norte (10%) e do Paraná (14,5%). Isso pode ter acontecido, contudo, não por uma dificuldade de se reconhecer dialetos como o do Paraná e do Norte do país, mas, sobretudo, porque nenhum sujeito destas duas regiões participou da pesquisa de percepção como ouvintes, o que pode, inclusive, colocar em xeque todos os dados obtidos nos testes de percepção de Constantini, visto que o reconhecimento dos dialetos ficou, de certa forma, tendencioso (certamente os cariocas – que participaram do teste de percepção – reconheceriam mais o seu dialeto do que o do Paraná).

Além de Rio de Janeiro, tiveram praticamente metade dos dialetos identificados, no estudo de Constantini, os sotaques do Nordeste (55,8%), São Paulo (53,7%) e Minas Gerais (55,9%). Os estímulos do Distrito Federal tiveram, em média, 10,8% de acertos.

Outro estudo prosódico que trabalhou com percepção, além da análise de produção, é o de Nunes (2015). A autora aplicou 8 testes perceptuais. Todos foram realizados com florianopolitanos, na cidade de Florianópolis, e com aracajuanos, na Grande Aracaju. Participaram dos 7 primeiros testes 25 florianopolitanos e 25 aracajuanos, com idade entre 19 e 60 anos. No último teste (oitavo), participaram 35 florianopolitanos e 35 aracajuanos, também com idade entre 19 e 60 anos.

As gravações usadas nos testes perceptuais foram feitas seguindo orientações do projeto Atlas Multimídia Prosódico do Espaço Românico do Português (Amper-Por), ao qual a pesquisa de Nunes está filiada (a explicação sobre a obtenção do *corpus*, neste projeto, está na seção 2 deste artigo).

Os dois primeiros testes de Nunes tiveram como objetivo a classificação das sentenças entre interrogativas neutras e não neutras (de confirmação e de dúvida), apenas com informações do pré-núcleo (com o sintagma nominal, ou seja, o que vem antes do verbo). Os testes 3 e 4 pretendiam classificar as sentenças entre interrogativas neutras e não neutras (de confirmação e de dúvida) apenas com informações do núcleo (todos os segmentos localizados depois do verbo). Os experimentos 5 e 6 testaram se os ouvintes identificavam modalidades apenas com informações tonais (enunciados sem os segmentos). Os testes 7 e 8 trataram de identificação dialetal. Estes dois últimos nos interessam sobremaneira, por isso apenas eles serão detalhados neste artigo.

O teste 7 tinha como finalidade fazer com que os ouvintes identificassem seu dialeto apenas com arquivos tonais (sem palavras). Participaram 50 juízes das duas capitais (Florianópolis e Aracaju), ouvindo 20 estímulos (10 de cada capital), totalizando 1.000 respostas. Os ouvintes florianopolitanos deveriam dizer, ao ouvir cada estímulo, se se tratava de um falante de Florianópolis ou de outro dialeto. Da mesma maneira, os ouvintes aracajuanos deveriam dizer se o estímulo era de um falante de Aracaju ou de outro dialeto.

Os florianopolitanos acertaram 56,8% dos estímulos produzidos por locutores de Florianópolis. Já os aracajuanos acertaram 58% dos estímulos produzidos por locutores de Aracaju.

O teste 8 tinha por objetivo apurar se florianopolitanos e aracajuanos identificavam suas próprias variedades em comparação a outras do mesmo estado. Por isso, o teste contou com estímulos de outras cidades tanto de Santa Catarina como de Sergipe. Os ouvintes de Florianópolis escutaram os estímulos de Santa Catarina e deveriam dizer se se tratava

de um falante de Florianópolis, de Blumenau, de Chapecó ou de Lages. Da mesma maneira, os de Aracaju ouviram os estímulos de Sergipe e deveriam dizer se se tratava de um falante de Aracaju, de Aquidabã, de Estância ou de Lagarto.

Os ouvintes florianopolitanos acertaram 84% dos estímulos de Florianópolis, 40% dos estímulos produzidos por lageanos, 37,7% dos chapecoenses e 37,1% dos blumenauenses.

De Sergipe, aracajuanos acertaram 43,3% das produções de Aracaju; 22,8% dos estímulos produzidos por lagartenses, 22,1% dos aquidabãenses e 17,1% dos estancianos.

Os resultados das pesquisas de Peres (2011), Constantini (2014) e Nunes (2015) mostram que testes de identificação podem ser eficientes no reconhecimento de dialetos. Os dados dos três estudos serão úteis para a comparação com os nossos resultados do teste de identificação. Nosso teste de discriminação dialetal não dialogará com outros estudos porque não encontramos pesquisas dialetais com este recorte.

Antes de adentrarmos na nossa pesquisa especificamente, sentimos a necessidade, ainda, de definirmos melhor o que chamamos neste artigo de dialeto, termo que também foi usado como sinônimo de falar e de variedade. Para isso, recorreremos à Sociolinguística.

Dialeto, falar e variedade são usados normalmente por pesquisas que estudam a diversidade da língua em comunidades de fala. Beline (2011) define estas comunidades como aquelas formadas por falantes que “compartilham traços linguísticos que distinguem seu grupo do de outros; comunicam-se relativamente mais entre si do que com outros; compartilham normas e atitudes diante do uso da linguagem” (p.128). Ou seja, são agrupamentos de falantes com características linguísticas em comum. O profissional que normalmente se ocupa de estudar isso, então, é o dialetólogo que, segundo Beline, tenta “localizar e descrever, regional e socialmente, os dialetos de uma língua e os diferentes falares que ela pode apresentar” (p.125).

Castilho (2010) define dialeto como a variedade linguística especificada por sua distribuição geográfica. O autor faz esta classificação conforme alguns dialetos (não todos) do PB, como os do Norte (amazônico e paraense), do Nordeste (pernambucano e baiano), do Sudeste (caipira e carioca), do Centro-oeste (cuiabano) e do Sul (paranaense, catarinense e gaúcho). Nesta linha de raciocínio, estudar as características regionais dos falares (por causa da localização geográfica destes falantes) é analisar um fenômeno da variação linguística. Sobre isso, Bagno (2007) afirma que “a língua varia de um lugar para outro, assim, é possível investigar, por exemplo, a fala característica das diferentes regiões brasileiras, de diferentes estados, e de diferentes áreas geográficas dentro de um mesmo estado” (p. 43), a chamada variação diatópica.

Para Bagno (2007, p.43), o termo variedade linguística “se refere a um dos muitos modos de falar uma língua”, correlacionados a fatores sociais como lugar de origem, idade, sexo, classe social, grau de instrução, entre outros. Então, umas das variedades linguísticas existentes seria o dialeto, de acordo com Bagno, “termo usado há muitos séculos, desde a Grécia antiga, para designar o modo característico de uso da língua num determinado lugar. Muitos linguistas empregam o termo dialeto para designar o que a Sociolinguística prefere chamar de variedade” (2007, p. 47 e 48).

1. Questões prosódicas

No português brasileiro, é a entoação a responsável por marcar a diferença das asserções neutras e interrogativas totais, por isso necessitamos desta melodia – perceptível auditivamente – para saber a intenção do locutor. Por meio dela, também, muitas vezes é possível reconhecer dialetos. Por trás deste sistema de reconhecimento auditivo operam parâmetros prosódicos ligados à percepção, que são denominados de *pitch*, *loudness* e quantidade.

De Angulo (1929, apud Crystal 1969) foi um dos primeiros estudiosos a se referir a estes aspectos suprasegmentais num sentido técnico, mas ainda de maneira geral. “Ele sugere o uso do termo prosódia para incluir todas estas mudanças que são frequentemente agrupadas como diferenças acentuais. Elas envolvem três fatores independentes: *pitch* (ou tom), duração (ou comprimento, quantidade) e amplitude (ou volume, stress, *loudness*)”³ (Crystal, 1969, p.56). Ainda segundo Crystal (1969), contemporâneos a De Angulo, Jakobson e Halle (1956) distinguiram o que é traço prosódico do que é traço inerente ao som, classificando-os em três tipos: “tom, força e quantidade correspondem aos três atributos da sensação (*voice-pitch*, *voice-loudness* e duração). E seus principais correlatos físicos são as dimensões de frequência fundamental, intensidade e tempo”⁵ (p.60).

A diferença é que alguns autores, como Lehiste (1970), estabelecem como suprasegmental o *pitch*, o acento e a quantidade e outros, como Ladd (2008), definem estes mesmos aspectos como frequência fundamental (F0), intensidade e duração. Jakobson e Halle (1956) explicam que esta segunda definição se refere às propriedades físicas e fonéticas dos sons (no âmbito da produção) e a primeira definição (*pitch*, acento e quantidade) aos seus correlatos psicofísicos (no âmbito perceptual).

Esta pesquisa tem como objetivo descrever o que ocorre neste âmbito psicofísico, fonológico, ou seja, o que os ouvintes percebem auditivamente como diferentes. Utilizaremos, para este fim, as

³ Ladefoged and Johnson (2006, p.249) explicam que o stress é uma característica suprasegmental dos enunciados aplicada para sílabas inteiras (e não apenas em vogais ou consoantes). Esta sílaba, se for tônica, por exemplo, é pronunciada com uma quantidade maior de energia do que a sílaba átona e, por isso, costuma ser mais proeminente no fluxo da fala.

⁴ He suggests the use of the term 'prosody' to include all those changes which are often lumped together as accentual differences. They involve three independent factors: pitch (or tone), duration (or length, quantity), amplitude (or volume, stress, loudness) (tradução nossa).

⁵ Tone, force and quantity, which they say correspond to the three main attributes of sensation, voice-pitch, voice-loudness and subjective duration. The dimensions of fundamental frequency, intensity and time are their closest physical correlates (tradução nossa).

modalidades “assertiva neutra” e “interrogativa total” como base para verificar se curitibanos, ao ouvir curitibanos falando, seriam capazes de identificar seu próprio dialeto e, ainda, se curitibanos, ouvindo outros curitibanos e outras pessoas de distintos falares, seriam capazes de discriminar seu dialeto em contraste (ou não) com os demais.

Discorremos, então, a partir de agora, sobre as características e definições dos três parâmetros (frequência fundamental, duração e intensidade) relacionados às propriedades físicas e psicofísicas.

A frequência fundamental (doravante F0) é medida em *Hertz* (Hz) e corresponde ao número de vezes por segundo em que as pregas vocais completam um ciclo de vibração (Bolinger 1972, p.19). O correlato perceptual da F0, então, é chamado de *pitch*. Muitas vezes traduzido para o português como tom ou inflexão tonal, segundo Madureira (1999, p.55). Ladefoged e Johnson (2006) explicam que muitas línguas usam o *pitch*, por exemplo, para marcar as fronteiras das unidades sintáticas. “Em quase todas as línguas, a conclusão de uma unidade gramatical, como de um enunciado normal, é sinalizada por um *pitch* em queda”⁶ (p.254). Ou seja, uma sílaba lexicalmente acentuada, quando está no final de um enunciado afirmativo, costuma ter um *pitch* mais baixo se comparada a outras posições no mesmo enunciado.

Pike (1972 p.53) lembra que cada língua faz uso particular da variação de *pitch*, caracterizando o modo de falar, por exemplo, de cada cidade, estado e país. “Todos os falantes de uma língua usam sequências básicas de *pitch* de maneira similar e em circunstâncias similares”⁷ (p.53). Ainda sobre o *pitch*, segundo Crystal (1969), este é definido pela *American Standards Association*, como “o atributo, em termos do qual o som pode ser ordenado em uma escala de baixo para alto”⁸ (p.108).

⁶ In nearly all languages, the completion of a grammatical unit such as a normal sentence is signaled by a falling pitch (tradução nossa).

⁷ All speakers of the language use basic pitch sequences in similar ways under similar circumstances (tradução nossa).

⁸ That attribute of auditory sensation in terms of which sounds may be ordered on a scale from low to high (tradução nossa).

A duração se refere ao tempo gasto para produzir a fala, também definida como a diferença de tempo entre dois eventos (Barbosa 1999, p.23). Ela pode ser chamada ainda de comprimento (*length*) e quantidade (na percepção auditiva) e costuma ser medida em segundos (s) ou milissegundos (ms). Em línguas como o inglês, a duração costuma ser o segundo traço acústico mais importante (depois da F0), porque é responsável pelo nível acentual das palavras (Ladefoged e Johnson 2006). Os autores explicam também que é graças à duração que palavras em inglês como *week* (semana) [wik] e *weak* (fraco) [wi:k] são distinguidas, ou seja, ambas são ditas no mesmo tom, mas um dos aspectos que as diferencia é a duração alongada da vogal “i” na palavra “fraco”. Eles afirmam ainda que “uma maneira de descrever o ritmo (de uma língua) é considerar quanta variação de comprimento ocorre dentro de uma frase”⁹ (Ladefoged e Johnson, 2006, p.251).

A intensidade, medida em decibéis (dB), tem relação com a altura (física) da onda, ou seja com a amplitude do som e com a energia usada para produzir este som. Ela tem como correlato perceptual o volume (*loudness*).

2. O Amper-Por como base do experimento

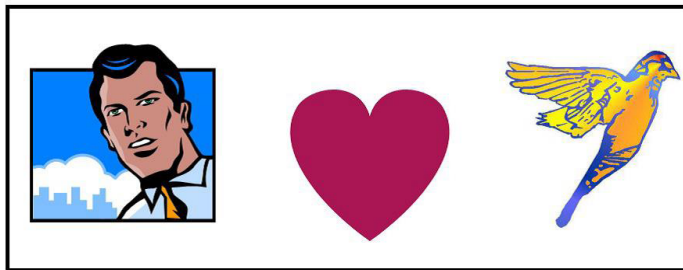
Esta pesquisa está vinculada ao projeto do Atlas Multimídia Prosódico do Espaço Românico do Português (Amper-Por). Trata-se de um projeto internacional que começou na França e na Itália, se expandiu para outros países europeus como Espanha e Portugal e chegou ao Brasil sob a coordenação da professora Lurdes de Castro Moutinho, da Universidade de Aveiro (Portugal). A metodologia é a mesma para todos os países, o que muda é a tradução do *corpus* conforme a língua onde ele é gravado. O *corpus* do Amper-Por adaptado para o português brasileiro é o que foi utilizado por nós. Por isso, primeiramente gravamos

⁹ One way of describing timing differences is to consider how much variation in length occurs within a sentence (tradução nossa).

curitibanos falando sentenças previstas pelo projeto e depois aplicamos testes de percepção a partir dos enunciados elaborados.

Esta primeira etapa, da gravação, foi feita da seguinte maneira. Os informantes visualizavam imagens no *Power Point* que, associadas, formavam frases a serem ditas. Antes de começar as gravações, os informantes participaram de um teste de familiarização. Este teste foi montado para que os informantes conhecessem os personagens (Renato, pássaro, bisavô), suas características (nadador, pateta, bêbado) e de que localidade poderiam ser (Mônaco, Veneza, Salvador). Também foi explicado a eles que a imagem do coração simbolizava o verbo gostar (que deve ser dito sempre na terceira pessoa do singular, no modo presente do indicativo, ou seja, “gosta”). Em seguida, os informantes visualizaram alguns *slides* montados com os personagens, suas características e o verbo para que pudessem treinar como deveriam dizer os enunciados. Na Figura 1, por exemplo, temos dois personagens (um homem e um pássaro) separados por um coração. O homem é designado Renato e o coração representa o verbo gostar. Sendo assim, temos a representação da frase “O Renato gosta do pássaro”.

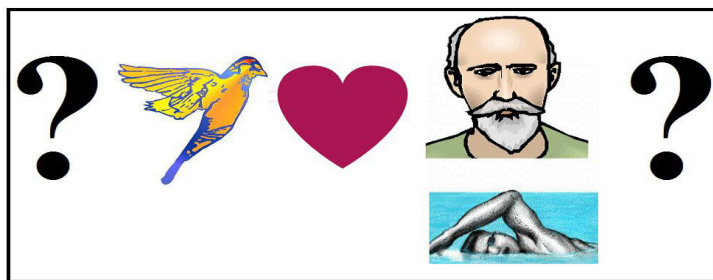
FIGURA 1: Slide do corpus do AMPER-POR para a frase “O Renato gosta do pássaro”



Para formular as interrogativas totais, a única coisa que muda é que o *slide* recebe um ponto de interrogação antes de todas as imagens e

um ponto de interrogação no final delas,¹⁰ como na Figura 2, em que o informante deve dizer “O pássaro gosta do bisavô nadador?”.

FIGURA 2: Slide do corpus AMPER-POR para o enunciado “O pássaro gosta do bisavô nadador?”



Assim que os informantes estivessem familiarizados com as figuras e os tipos de enunciados que deveriam proferir, começamos oficialmente a gravação das 66 frases que precisavam ser elaboradas após os informantes visualizarem cada um dos 66 slides no Power Point com as figuras que formavam os enunciados. Depois de gravar os enunciados, selecionamos alguns deles para montar os testes de percepção (veja mais a este respeito na seção 2.2 deste artigo).

2.1 Participantes

Contamos com dois informantes curitibanos que gravaram as sentenças para os testes de percepção: um do sexo masculino e um do sexo feminino. Ambos nasceram e sempre viveram em Curitiba (capital paranaense) e se reconhecem como falantes deste dialeto e com grau de escolaridade ensino médio completo. Ele tinha, na época da gravação, 24 anos e, ela, 23 anos.

¹⁰ Nunes (2013) modificou os slides das interrogativas totais, acrescentando um ponto de interrogação no início de cada um deles, além do ponto de interrogação no final, porque a autora percebeu que os informantes, quando visualizavam slides sem este ponto de interrogação inicial, costumavam falar o enunciado como se fosse declarativo e, no final, quando se davam conta de que era uma pergunta, mudavam a entoação.

Em seguida contamos com mais 20 informantes, desta vez ouvintes, que julgaram as gravações dos dois primeiros informantes curitibanos em dois testes de percepção. Estes ouvintes eram curitibanos de nascença e sempre residiram na capital paranaense, sendo 10 homens e 10 mulheres, idade entre 25 e 62 anos (idade média 36 anos). Os testes foram aplicados no mês de dezembro de 2014. Damos preferência para ouvintes que não tinham nenhuma relação com a formação de Letras, a fim de evitar que houvesse algum conhecimento acadêmico envolvido com o teste, podendo, assim, distorcê-lo. Por isso, nossos 20 informantes têm perfil profissional variado: advogados, farmacêuticos, artesãos, jornalistas, funcionários públicos, médicos, nutricionistas e dentistas (todos com ensino superior completo).

2.2 Os testes de percepção

Esta pesquisa contou com um teste de identificação e outro de discriminação. Pela natureza destes dois testes, o de identificação costuma ser mais difícil aos ouvintes, pois estes precisam identificar o seu dialeto ao ouvir apenas uma sentença, sem nenhuma opção de comparação. Beddor & Gottfried (1995, apud Kluge 2004) dizem, sobre o teste de identificação, que “uma vantagem é que este tipo de teste tem uma quantidade limitada de alternativas para resposta deixando uma carga mais leve na memória dos ouvintes¹¹” (p. 222), porém, “a desvantagem deste teste é que ele envolve familiaridade com os dados”¹² (p.223).

Já o teste de discriminação, pela sua natureza, costuma ser mais fácil porque tem sempre um enunciado a ser comparado com outro e isso quer dizer que será possível ter o dialeto de Curitiba em contraste com outros dialetos ou o dialeto de Curitiba sem contraste com outros (porque as duas sentenças a serem ouvidas ou serão de Curitiba ou nenhuma será de Curitiba).

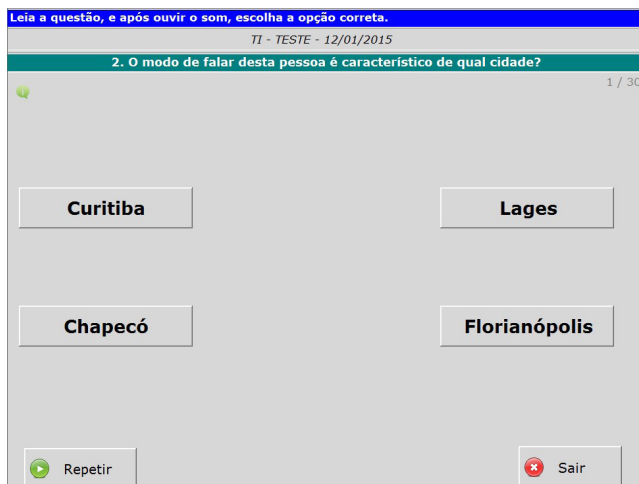
¹¹ An advantage of this kind of task is that the limited set of response alternatives causes “a lighter load on memory for listeners (tradução nossa).”

¹² A disadvantage of identification tasks is that labeling involves familiarity (tradução nossa).

Os dois testes foram elaborados e rodados no programa TP, um *software* gratuito (baixado do site <http://www.worken.com.br/tp>). Ele foi criado por Andréia Rauber, Anabela Rato, Denise Cristina Kluge, Giane dos Santos e Marcos Figueiredo com a finalidade de ajudar pesquisadores na elaboração e aplicação de testes perceptuais.

O primeiro teste então, de identificação dialetal, contou com 30 enunciados (10 sentenças repetidas três vezes) que foram avaliados pelos 20 ouvintes curitibanos. A partir da tela exposta na Figura 3, os ouvintes deveriam dizer se o enunciado que ouviam era de uma pessoa de Curitiba, de Florianópolis, de Lages ou de Chapecó.

FIGURA 3: Tela mostrando o programa TP para o teste de identificação dialetal

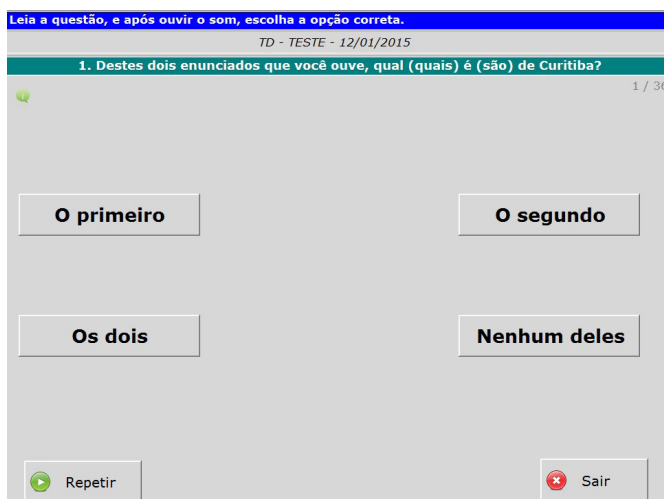


Existiam quatro enunciados (duas perguntas e duas afirmações) do dialeto de Curitiba e duas (uma pergunta e uma afirmação) de Chapecó, Lages e Florianópolis – estes últimos funcionaram como distratores e não nos interessam na análise dos resultados, ou seja, não levamos em conta se curitibanos não identificavam os dialetos de Santa Catarina, o que nos interessou foi ver se eles identificam os de Curitiba. As 10 sentenças

(todas de mulheres) foram repetidas três vezes (aleatoriamente pelo problema TP) e, após o participante ouvir cada uma das 30 sentenças, ele tinha a opção de ouvir mais uma vez (clcando no botão “repetir”) antes de clicar no botão de resposta. Ao clicar na resposta desejada, o participante ouvia a próxima sentença.

O segundo teste relacionado ao dialeto media se os ouvintes curitibanos reconheciam seu modo de falar quando este estava em contraste (ou não) com outra variedade através de um teste de discriminação. Conforme Figura 4, neste teste os ouvintes ouviam duas sentenças e deveriam apontar qual (ou quais) enunciado (s) era (m) de Curitiba: o primeiro, o segundo, nenhum deles ou os dois.

FIGURA 4: Tela mostrando o programa TP para o teste de discriminação dialetal



Neste teste de discriminação não era possível imaginar de onde eram as pessoas das gravações porque não aparecia, como possibilidade de escolha, a cidade de origem delas (ao contrário do teste de identificação). Apenas se sabia que, em algum momento, haveria a fala de Curitiba. Neste teste, diferentemente do primeiro, utilizamos sentenças de

informantes masculinos e femininos. Em cada estímulo, os ouvintes precisavam escutar dois enunciados e indicar qual ou quais era/eram de Curitiba. Montamos o segundo teste apenas com as declarativas (uma de um homem e uma de uma mulher) de Curitiba, de Lages e Chapecó (estes dois últimos também funcionaram como distratores).

Os estímulos eram compostos de nove pares de enunciados repetidos duas vezes (aleatoriamente pelo programa TP), totalizando 36 pares de gravações a serem ouvidos. Os ouvintes poderiam escutar os estímulos apenas mais uma vez, se desejassem, clicando na tecla “repetir” antes de clicar no botão de resposta. Os pares foram montados da seguinte maneira:

Curitiba (homem/mulher) x Curitiba (homem/mulher)
Curitiba (homem/mulher) x Chapecó (homem/mulher)
Chapecó (homem/mulher) x Curitiba (homem/mulher)
Chapecó (homem/mulher) x Chapecó (homem/mulher)
Curitiba (homem/mulher) x Lages (homem/mulher)
Lages (homem/mulher) x Curitiba (homem/mulher)
Lages (homem/mulher) x Lages (homem/mulher)
Lages (homem/mulher) x Chapecó (homem/mulher)
Chapecó (homem/mulher) x Lages (homem/mulher)

Decidimos não retirar do sinal acústico os segmentos produzidos nos enunciados seguindo o que defendem Vaissière e Boula de Mareüil (2004), de que muitas variações dialetais percebidas são frutos da combinação entre a entoação e a informação segmental e que, por isso mesmo, podem não ser reconhecidas quando apresentadas separadamente.

2.3 Tratamento estatístico para os dados

Para a análise dos dados foram consideradas 240 respostas (4 estímulos de Curitiba x 3 repetições x 20 ouvintes) no teste de identificação. Para

o teste de discriminação foram consideradas 880 respostas (22 estímulos x 2 repetições x 20 ouvintes). Não nos interessou medir se os ouvintes curitibanos erraram o reconhecimento do dialeto de Santa Catarina, porque estes funcionaram apenas como distratores. Por isso observamos somente se o dialeto de Curitiba foi discriminado. Nos detemos ainda a avaliar se o dialeto curitibano foi mais reconhecido quando havia contraste ou quando não havia contraste de dialetos.

Os dados – obtidos pelas respostas dos testes – passaram por tratamento estatístico, a partir do programa SPSS. Os resultados do teste de identificação dialetal foram avaliados pelo Teste de Wilcoxon para a comparação par em par: neste caso analisamos se o número de acertos na identificação do dialeto de Curitiba é significativo em relação ao número de erros na identificação deste dialeto ($p \leq 0.05$). Também nos propomos a observar se o número de acertos dos informantes masculinos é (ou não) significativamente maior (ou menor) que o de acertos femininos, por meio do Teste de Mann Whitney U para a comparação entre dois grupos independentes (neste caso o sexo feminino e masculino).

Os dados do teste de percepção para avaliar se curitibanos discriminavam o seu dialeto em comparação com outros três dialetos (Chapecó, Lages e Florianópolis) foram analisados pelo Teste de Friedman para a comparação entre seis instrumentos (quando a resposta certa era a primeira, a segunda, as duas e nenhuma delas ou quando as respostas mais acertadas foram quando havia contraste de dialeto durante a audição ou quando não havia este contraste). As diferenças foram localizadas por meio do Teste de Wilcoxon, como *post hoc* e, sempre que necessário, foi feita a Correção de Bonferroni (Martins 2011).¹³

¹³ A Correção de Bonferroni é utilizada para evitar que ocorram erros nos resultados significativos e consiste em dividir o valor do ponto de corte convencional relativo à significância ($p < 0,05$) pelo número de testes adicionais (MARTINS 2011).

3. Resultados

O teste de identificação dialetal, em que os ouvintes curitibanos escutavam um enunciado e deveriam dizer de onde era a pessoa da gravação (Curitiba, Florianópolis, Lages ou Chapecó), teve um grande percentual de reconhecimento do dialeto curitibano (lembrando que apenas nos interessava medir se curitibanos reconheciam o falar de Curitiba, não importando se errassem os dialetos de Santa Catarina). O valor total de acertos dos 20 ouvintes foi de 170 e o número de erros foi de 70 (de um total de 240 respostas que deveriam ser apontadas como de Curitiba). Este primeiro teste revelou que curitibanos são capazes de reconhecer seu dialeto ao ouvir sentenças isoladamente, já que o percentual de acertos foi de 71%.

O Teste de Wilcoxon mostrou que a diferença entre acertos e erros é significativa ($Z=-3,421$, $p=0,001$). Também aplicamos o Teste de Mann Whitney para medir se a média de acertos, conforme o sexo do ouvinte, era significativamente diferente: as mulheres ouvintes reconheceram, em média, 9,95 enunciados (de 12) e os homens 11,05 (de 12). O teste revelou que a diferença de acertos entre ouvintes masculinos e femininos não é significativa ($U=44.500$, $p=0,674$), indicando que homens e mulheres identificam o dialeto de Curitiba da mesma forma.

O alto índice de identificação do dialeto de Curitiba (71%) também foi alcançado em estudos de outros dialetos brasileiros, descritos na seção introdutória deste artigo. Peres (2011) teve uma média de acertos de 75% na identificação dialetal de São Paulo (por paulistas), de Pelotas (RS) (por pelotenses) e de Senador Pompeu (CE) (por senadorenses), quando os segmentos foram mantidos nos testes. O índice de acertos na identificação dos três dialetos estudados por Peres, porém, caiu para uma média de 66,3% quando os enunciados ouvidos estavam apenas com a informação tonal (“sem as palavras”).

O teste de Constantini (2014) também manteve um alto índice de identificação dialetal, 77,9% do falar do Rio de Janeiro foi reconhecido

por cariocas. Não citaremos os outros dialetos pesquisados pela autora por causa da falha na aplicação do teste, conforme descrevemos na introdução deste artigo.

Ainda, nos estudos de Nunes (2015) o índice de acertos para o dialeto de Florianópolis (apenas com informação tonal) foi de 56,8% e de 58% para a fala de Aracaju. Além disso, quando a informação segmental foi mantida nos testes de Nunes, os florianopolitanos aumentaram a média de acertos de seu dialeto para 84%.

Estas pesquisas, acrescidas à nossa, demonstram que perceptualmente os dialetos podem ser identificados com êxito pelos seus falantes.

O teste perceptual de discriminação dialetal exigia dos ouvintes curitibanos reconhecer o falar de Curitiba em contraste (ou não) a outros falares (sem identificar de onde eram os outros dialetos). Então, eles ouviam duas sentenças em cada estímulo e deveriam dizer qual (quais) era (m) de Curitiba: a primeira, a segunda, as duas ou nenhuma. Lembrando que havia a possibilidade de o falar de Curitiba aparecer nas duas sentenças ou em nenhuma delas (sem contraste), ou ainda, poderia aparecer na primeira gravação ou na segunda (com contraste).

Os resultados mostraram que os ouvintes curitibanos discriminaram com maior êxito seu dialeto quando este não estava em contraste com outro dialeto e quando as duas gravações eram de falantes curitibanos. O percentual de acertos foi de 72,5%, ou seja, eles acertaram 58 de 80.

Quando havia contraste, ou seja, quando ouviam o dialeto de Curitiba em comparação com a gravação de um outro dialeto antes ou depois, acertaram 38%, o que quer dizer que das 320 respostas que deveriam apontar que o dialeto de Curitiba estava por primeiro ou segundo, eles acertaram 123.

Quando as duas gravações não eram de Curitiba (ou sejam, eram de Santa Catarina), o número de acertos caiu para 28% (90 respostas apontaram corretamente que se tratava de fala catarinense e que ali não havia a variedade de Curitiba, de um total de 320). Importante

ressaltar que não nos interessou observar se curitibanos seriam capazes de reconhecer a fala de outros dialetos, porém, citamos o número de acertos deste último critério do teste – quando não havia falas de Curitiba nas duas gravações – porque os ouvintes apontaram em alguns casos, erroneamente, que havia fala de Curitiba onde era fala catarinense.

O Teste de Wilcoxon, de par em par, mostrou que não é significativa a diferença entre os acertos quando (i) o falar de Curitiba vinha por primeiro (média de 2,95 acertos de 8) ou quando (ii) o dialeto curitibano aparecia por segundo (média de 3,2 acertos de 8) ($Z=-1,387$, $p=0,166$), indicando que a posição do falar de Curitiba no estímulo não teve efeito nos resultados. Porém, o mesmo teste apontou diferença significativa no par de comparação de acertos quando (i) as duas gravações eram de Curitiba ou quando (ii) nenhuma era de Curitiba ($Z=-2,123$, $p=0,034$), revelando que, quando não havia contraste, os participantes tiveram mais acertos quando as duas sentenças eram faladas por curitibanos.

Levando em consideração as seis condições (acerto Curitiba primeiro, acerto Curitiba segundo, acertos para os dois de Curitiba, acerto para nenhum de Curitiba, acerto quando o dialeto de Curitiba estava em contraste com outro, acerto quando o falar de Curitiba não estava em contraste) o Teste de Friedman apontou que a diferença entre o valor destes acertos é significativa ($X^2(5, N=20) = 66,654$, $p=0,000$). Como o Teste de Wilcoxon mostrou que não há diferença entre os valores de acertos para Curitiba por primeiro ou segundo, juntamos estes dados. Então, rodamos o Teste de Wilcoxon – como *post hoc* de Friedman – para ver onde estava a diferença entre as três condições (acertos quando nenhum falar era de Curitiba, quando os dois eram de Curitiba ou quando um deles – em contraste com outro – era de Curitiba). Por se tratarem de três condições, fizemos a Correção de Bonferroni¹⁴. Este teste reportou que o número de acertos é significativamente maior quando curitibanos ouviam o seu dialeto em contraste com outro. Porém, analisando o percentual de acertos com relação ao total de enunciados para cada caso

¹⁴ O valor de significância 0,05 dividido por três: $p = 0,017$.

(quando os dois eram de Curitiba, quando nenhum era de Curitiba ou quando um deles era de Curitiba) podemos afirmar que ainda assim o reconhecimento maior é quando não há contraste de dialeto e quando o falar de Curitiba aparece nas duas gravações.

Os resultados do nosso teste de discriminação não serão confrontados com outros estudos porque não encontramos trabalhos desta categoria que pudessem ser comparados ao nosso.

Considerações finais

A pesquisa descrita neste artigo tende a indicar que testes perceptuais podem ser um bom auxílio para estudos prosódicos a respeito de dialetos do português brasileiro. No âmbito perceptual, os ouvintes de Curitiba identificaram com êxito (71% do total) seu dialeto quando este estava em sentenças isoladas, classificando-as corretamente como pertencentes ao dialeto curitibano.

No teste de discriminação dialetal, o reconhecimento também foi alto quando as duas sentenças ouvidas eram de Curitiba (72,5%), porém, o restante do teste teve o número de acertos reduzido consideravelmente. Quando os dois falares eram catarinenses, erroneamente os ouvintes apontaram em grande parte do teste que um dos falares (quando não os dois) eram de Curitiba, por isso acertaram apenas 28% dos estímulos, dizendo que eles não eram da variedade da capital paranaense. Quando uma gravação de Curitiba vinha em contraste com outra catarinense, o número de acertos no reconhecimento do falar curitibano foi de 38%.

Diante destes dados, acreditamos que o teste de discriminação se mostrou mais difícil aos ouvintes do que o teste de identificação. Temos algumas hipóteses para estes resultados, porém elas precisam ser cientificamente melhor investigadas. Uma delas é a natureza do teste de discriminação. Como os ouvintes precisam escutar duas gravações para dizer qual (ou quais) é (são) de Curitiba, isso talvez dificulte

a memorização de qual foi a gravação ouvida por primeiro. Ou seja, quando os ouvintes estão na segunda gravação, talvez já não se recordem muito bem de como era a primeira.

Outra questão é que os ouvintes apontaram a dificuldade de distinguir o dialeto de Curitiba (PR) com o de Chapecó (SC), por isso o maior número de erros esteve quando os dois dialetos eram comparados e os ouvintes diziam que ambos eram do falar da capital paranaense. Isso nos leva a questionar o quanto seriam parecidos estes dois dialetos e de que maneira (se é que isso é possível) eles poderiam ser diferenciados prosodicamente.

Referências

AUBERGÉ, Veronique. **A gestalt morphology of prosody directed by functions:** The example of a step-by-step model developed at ICP. In *Proceedings of Speech Prosody*, Aix-en-Provence, p. 151-155, 2002.

BAGNO, Marcos. **Nada na língua é por acaso: por uma pedagogia da variação linguística.** São Paulo: Parábola Editorial, 2007.

BARBOSA, Plínio A. **Revelar a estrutura rítmica de uma língua construindo máquinas falantes:** pela integração de ciência e tecnologia de fala. In: SCARPA, Ester M. (Org.). **Estudos de Prosódia.** Campinas, SP: Editora da Unicamp, p. 21-52, 1999.

BELINE, Ronald. **A variação linguística.** In: FIORIN, J.L (Org.). **Introdução à linguística: objetos teóricos.** São Paulo: Editora Contexto, p. 120-140, 2011.

BOLINGER, Dwight. **Around the edge of language: intonation.** In: BOLINGER, Dwight (Ed.). *Intonation.* Inglaterra: Great Britain, p. 19-29, 1972.

_____. **Intonation and its Parts**. London: Edward Arnold Publishers, 1985.

CASTILHO, Ataliba T. de; **Gramática do português brasileiro**. São Paulo: Editora Contexto, 2010.

CONSTANTINI, Ana Carolina. **Caracterização prosódica de sujeitos de diferentes variedades falas no português brasileiro em diferentes relações sinal-ruído**. 115 fls. Tese de doutorado – Instituto de Estudos da linguagem, Universidade Estadual de Campinas, Campinas, 2014.

CRYSTAL, David. **Prosodic Systems and intonation in English**. Londres: Cambridge University Press, 1969.

FÓNAGY, Ivan. **Des fonctions de l'intonation: essay de synthèse**. In: Flambeau, Tokyo, n. 29, p. 1-20, 2003.

JAKOBSON, R; HALLE, M. **Fundamentals of Language**. The Hague, Mouton, 1956.

KLUGE, Denise C. **Perception and production of English syllable-final nasals by brazilian learners**. 175 f. Dissertação de Mestrado – Programa de Pós-graduação em Letras-Inglês e Literatura Correspondente, Universidade Federal de Santa Catarina, Florianópolis, 2004.

LADD, Robert D. **Intonational phonology**. 2ª ed. Nova Iorque: Cambridge University Press, 2008.

LADEFOGED, Peter; JOHNSON, Keith. **A course in phonetics**. Boston (USA): Wadsworth, Cengage Learning. 6ª ed, 2006.

LEHISTE, I. **Suprasegmentals**. MIT, Cambridge: Massachusetts, 1970.

MADUREIRA, Sandra. **Entoação e síntese de fala: modelos e parâmetros**. In SCARPA, Ester M. (Org.), Estudos de Prosódia. Campinas, SP: Editora da Unicamp, p. 53- 68, 1999.

MARTINS, C. **Manual de análise de dados quantitativos com recurso ao IBM SPSS**. Braga: Psiquilibríos Edições, 2011.

MORAES, J.A.; RILLIARD, A. **Illocution, attitudes and prosody: a multimodal analysis**. In: Raso, T., et al. (Org.), Spoken Corpora and Linguistic Studies, Amsterdam: John Benjamins Publisher, p. 233-270, 2014.

MORAES, João. **Recherches sur l'Intonation Modale du Portugais Brésilien Parlé à Rio de Janeiro**. Thèse de Doctorat de Troisième Cycle. Université de la Sorbonne Nouvelle, Paris III, 1984.

NUNES, Vanessa. **Características entoacionais de sentenças interrogativas totais em falares catarinenses e sergipanos**. 119 f. Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina. Qualificação de doutorado, Florianópolis, 2013.

_____. **A prosódia de Sentenças Interrogativas Totais nos falares Catarinenses e Sergipanos**. 323 fls. Tese de doutorado – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2015.

PERES, Daniel Oliveira. **O papel da prosódia na identificação das variedades regionais do português brasileiro**. 152 fls. Dissertação de mestrado – Faculdade de Filosofia, Letras e Ciências Humanas, Programa de Pós-Graduação em Filologia e Língua Portuguesa, Universidade de São Paulo, São Paulo, 2011.

PIKE, Kenneth L. **General characteristics of intonation**. In: BOLINGER, Dwight (Ed.). *Intonation*. Inglaterra: Great Britain, p. 53-82, 1972.

PROJETO AMPER-POR. Disponível em: <<http://pfonetica.web.ua.pt/AMPER-POR-POR.htm>>.

VAISSIÈRE, J.; BOULA DE MAREÜIL, P. **Identifying a language or na accent**: from segments to prosody. In: *Modelling for the identification of languages (MIDL) Workshop*. Paris, p. 1-6, 2004.

Recebido em 29/03/2015 e Aceito em 09/06/2015.