

A prosódia e o reconhecimento dialetal¹⁴

Daniel Oliveira Peres

Introdução

É possível, de maneira intuitiva, perceber diferenças entre as línguas. A partir de uma breve observação, nota-se que elas podem diferir de forma significativa quanto à sua matéria fônica, formação de palavras, construções de sentenças etc. O fato prosódico também faz parte dessas pistas que são buscadas pelos ouvintes na tarefa de reconhecimento de determinada língua. Os trabalhos de Atkinsons (1968), Bonte (1975), Maidment (1976) e Ohala e Gilbert (1978) corroboram a hipótese de que ouvintes são capazes de identificar sua própria língua por meio de informações prosódicas. Esta habilidade de buscar pistas prosódicas (ou segmentais) pode ser estendida também para as variações intralinguísticas. As características contidas nas variações de frequência fundamental, amplitude e duração desempenham o papel de sinalizadores de diversos tipos de variações linguísticas, sejam elas de caráter diatópico, diafásico ou diastrático. No âmbito da variação dialetal, alguns estudos foram feitos com o intuito de verificar se as variações prosódicas podem ser determinantes para o reconhecimento de uma certa variedade linguística, a saber, os trabalhos de Bush (1967) e Richardson (1973).

¹⁴ ENCONTRO DE PÓS-GRADUANDOS DA FFLCH/USP – EPOG, 5. São Paulo, 2010.

Pautado na hipótese de reconhecimento linguístico baseado nas informações prosódicas, esse trabalho visa a desenvolver um estudo perceptual em que se possa testar a habilidade dos falantes de português brasileiro na tarefa de reconhecimento de sua variedade linguística.

Objetivos

O presente trabalho tem como objetivo verificar a hipótese de que ouvintes são capazes de reconhecer sua variedade linguística por meio de suas características prosódicas. Em artigo, Ohala e Gilbert (1978) testaram a habilidade dos ouvintes em identificar línguas através apenas de suas pistas prosódicas: *“languages do differ in their prosody by seeing if listeners are successful in identifying the language used by speaker when they hear only the fundamental frequency, amplitude and certain timing characteristics of the original voice signal”*. A hipótese de Ohala e Gilbert partiu de estudos anteriores e da intuição de que línguas diferem não somente em sua parte segmental, mas também em suas variações prosódicas.

O presente trabalho diferirá, de certa forma, do objetivo do estudo comparativo entre línguas (ATKINSONS, 1968; BONTE, 1975; MAIDMENT, 1976; OHALA; GILBERT, 1981), pois o que estará em questão neste trabalho será o reconhecimento de variedades linguísticas regionais dentro de uma mesma língua.

Este estudo trabalhará com variedades do português brasileiro, sendo, dentre tantas variedades existentes, elencadas duas delas: variedade gaúcha (Pelotas) e paulistana (bairro da Mooca). A escolha dessas variedades está baseada no trabalho desenvolvido por Diegues (1960) sobre ocupação do território brasileiro, bem como suas características regionais.

Os ouvintes participarão de um teste perceptual elaborado de maneira a deixar para os participantes do experimento

somente as características prosódicas. Para a feitura do experimento será utilizado o script PURR (SONNTAG; e PORTELE, 1998), para Praat (BOERSMA e WENNINK, 1986). Em contrapartida, será aplicado um teste-controle que eliminará as variações prosódicas, restando somente a curva melódica monotônica, juntamente com os segmentos.

Entoação

Troubetzkoy (1964) já percebera a entoação como variação tonal nas frases, sendo uma constante oposição de direções tonais, mas tal característica aplicada somente a línguas que não possuem esse correlato para distinção lexical. Assim, Troubetzkoy postula duas direções tonais para a sua análise: *ascendente* - função de continuidade; *descendente* – função conclusiva. A prosódia estaria então subdividida em *entoação da frase, diferenças de registro na frase, acento frasal e pausa*.

Pike (1954; 1964) postulou que a entoação frasal da língua inglesa fosse analisada numa escala de 4 graus. Os trabalhos que o sucederam também propuseram uma análise entoacional a partir de níveis discretos; a saber, Maeda (1971), Pierrehumbert (1980;1990), Cruttenden (1986) e Ladd (1996).

A entoação pode ser encarada de duas maneiras. Nos estudos de prosódia podemos encontrar o termo ‘entoação’ para designar os fenômenos suprasegmentais com valores estritamente prosódicos. Neste viés, toda e qualquer manifestação prosódica como variações de F0, duração e amplitude estão arrolados no termo entoação. Numa abordagem mais estrita, a entoação seria a variação melódica, isto é, modulações do parâmetro f0 que ocorrem durante a fala (HIRST & Di CRISTO, 1998).

Vaissière (2004) faz uso do termo ‘entoação’ como variação melódica que acontece ao longo do discurso. Por outro lado, a autora cita o papel de outros parâmetros no auxílio das

variações prosódicas, incluindo nestes os de ordem fisiológica. Apesar de reconhecer que os estudos de entoação devem incluir o maior número de parâmetros possível, a sua categorização acerca das múltiplas funções da entoação está mormente condicionada às variações de F0.

Ferreira Netto (2006) trata a prosódia como uma série temporal, ou seja, um conjunto de observações sequenciadas e interdependentes, sendo o resultado da observação feita no momento $t+1$ condicionada ao momento t (FERREIRA NETTO, 2006). A análise está baseada na proposta de Xu e Wang (1997) de que alguns fatos prosódicos são decorrências de restrições *mecânico-fisiológicas* e outros configuram as intenções *expressivas* dos falantes. Para dar conta dessa proposta, faz-se a decomposição automática de F0 por meio do aplicativo Exprosodia®, dividindo-o em *componentes estruturadoras* e *semântico-funcionais*. Assim, as componentes ainda seriam subdivididas em finalização (F) e sustentação (S) - formando o ritmo tonal - foco/ênfase (E) e acento lexical (A), formando a face significativa da entoação frasal.

Percepção e produção

Ao trabalhar com um fator fonético e um fator linguístico. T'Hart, Collier & Cohen (1990) atestam que uma abordagem perceptual da entoação é capaz de trazer à tona o que é linguisticamente importante no tocante às características prosódicas. Para os autores, a percepção age como um selecionador de informações capaz de reter da cadeia sonora da fala o que é relevante. Dessa maneira, a abordagem perceptual não está focada numa visão psicoacústica da fala, tendo como expectativa um limiar de variações que possam ser não só percebidas, mas que elucidem quais propriedades do sinal acústico são importantes para determinadas questões e, num outro nível de interesse, quais mecanismos fisiológicos afetam diretamente a produção das variações prosódicas relevantes ('t

HART; COLLIER; COHEN, op.cit.). De maneira análoga, Patrícia Kuhl e seus colegas (KUHL; IVERSON, 1995; KUHL, 2000; KUHL *et al.*, 2001) propuseram o *perceptual magnet effect* (PME). A forma prototípica do falante, segundo ela, seria algum dos sons ouvidos em alguma etapa da vida do falante que, memorizada no todo, serve como pólo de atração de todos os estímulos que atinjam regiões próximas. Assim, nem toda a gama de variações é percebida, mas somente aquelas que desempenham papel significativo, neste caso, fruto de uma atração perceptual.

De outro lado, há vários outros estudos que trabalharam com a produção, analisando amostras de fala com intuito de buscar algumas pistas acerca das variações tonais e das características prosódicas como um todo. Vaissière enumera alguns tipos de abordagens da entoação em que, apesar de partirem de abordagens perceptuais, estabelecem modelos teóricos e representacionais diferentes para o estudo da produção. Uma abordagem trabalha com um sistema exclusivamente linear, em que a representação simbólica da entoação e as unidades prosódicas desempenham um papel principal (PIERREHUMBERT, 1980). Outra visão parte da hipótese de que o F0 é formado pela superposição de bases globais, componentes frasais e acento lexical (FUJISAKI; SUDO, 1971; OHMAN, 1967).

Os estudos perceptuais e a variação linguística

Este trabalho está no limiar entre os estudos perceptuais e dialetológicos. O ponto de contato está na utilização de um instrumental perceptual para investigar fenômenos que são temas específicos dos estudos dialetológicos. Para Clopper e Pisoni (2004), o intuito dos estudos perceptuais estava mormente relacionado ao reconhecimento de categorias abstratas da língua, renegando a importância da variabilidade concreta do uso. Porém, em trabalhos recentes, os estudos perceptuais estão desempenhando papel importante nas

pesquisas de discriminação dialetal. O que está em questão nessa junção é a disparidade de pressupostos, já que para os pesquisadores que buscam uma regularidade, um modelo abstrato em que a percepção das categorias percebidas seja invariável, não importa a variação que essas categorias possam sofrer, sendo esta *“a source of noise [...] irrelevant to the underlying representations”* (CLOPPER; PISONI, 2004). Dessa forma, toda e qualquer variação acerca de uma categoria central torna-se dispensável.

Por outro lado, o objeto da sociolinguística está justamente nessa gama de variações que pode ou não ser significativa para identificar variedades regionais, sociais e estilísticas. Estreitando a abordagem para os interesses deste estudo, a percepção das variedades linguísticas regionais estaria relacionada à maneira como os ouvintes fazem uso da informação contida no sinal acústico, de modo que estas informações sejam relevantes para a identificação da variedade produzida pelo falante. Portanto, este trabalho pretende colocar uma abordagem a serviço da outra, em outras palavras, trazer para os estudos de percepção da fala, mais especificamente da prosódia, os fins encontrados nos estudos sociolinguísticos.

Sobre a escolha das variedades

O pressuposto para a feitura desse tipo de estudo é a variância inerente que as línguas naturais apresentam, a despeito de uma análise autônoma como sistema isolado e independente. Da extremidade saussureana da língua como sistema à visão sociolinguística - a qual tem como alicerces o contexto histórico-cultural, situação de uso e outras características trazidas pelos falantes no âmbito social - faz-se necessária uma breve reflexão sobre quais fatores são importantes para a configuração das variedades linguísticas. A definição sistêmica das línguas naturais colabora com o que há de uniforme e, portanto, caracterizador de uma determinada língua, assegurando uma

clara divisão entre o que é ou não parte de um determinado sistema. O ponto de vista sociolinguístico agrega ao que há de unificador a qualidade variável e mutatória das línguas.

Em seu trabalho, Diegues (1960), antes de traçar as regiões culturais brasileiras, afirma que é necessário pensar o Brasil como parte de um conjunto maior, de uma área cultural mais ampla, na qual estão inclusos Portugal e suas Províncias Ultramarinas. Essa área cultural luso-cristã, assim chamada, possui dois elementos caracterizadores: o elemento lusitano na aglutinação da população portuguesa e o cristianismo como base religiosa. Posto isso, é a integração desses elementos com a características autóctones que permite precisar traços comuns entres os vários países dessa grande área e, por outro lado, mantém a diversidade proveniente do ambiente natural e social de cada lugar.

A ocupação humana do território brasileiro não foi fruto de uma colonização bem definida, com metas e normas rígidas. Como consequência disso, nasce a possibilidade de pensar o Brasil como um conjunto de regiões culturais, pois cada uma delas é o reflexo das relações estabelecidas entre os povoadores, os povos nativos e a terra ocupada. Diegues afirma que é dessa diversidade que surge a unidade. É importante salientar que essa delimitação não é discreta, ou seja, as regiões culturais não são justapostas, mas são alteradas ao longo da ocupação.

Uma região cultural seria um conjunto ecológico de pessoas, agregadas pela similaridade das relações sociais, pela unidade das relações espaciais da população, somando a essas a estrutura econômica local. Do ponto de vista linguístico, Mufwene (2007) também defende que as características linguísticas são frutos da ecologia na qual elas foram geradas. Mas, é importante dizer que o foco de Mufwene é a distinção entre as mudanças linguísticas interna e externamente motivadas, propondo que tal distinção tem bases ideológicas,

mormente quando se trata do estabelecimento de um conceito sobre o que seriam as línguas crioulas. No caso de uma possível comparação entre o uso do termo 'ecológico' feito por Diegues e a abordagem de Mufwene, haveria uma intersecção no fato de que as características tanto culturais quanto linguísticas são frutos da ecologia interna de cada região.

Dessa forma, as regiões culturais seriam espaços territoriais delimitados por certas características que unificam as idéias, os sentimentos e o estilo de vida de um grupo populacional (DIEGUES, 1960: 7). Num breve parêntese, direcionando a discussão para uma abordagem sociolinguística, tais regiões também compartilhariam traços linguísticos particulares, pois o papel da interação social exerce grande influência na variação e mudança linguísticas. Labov (1972) considera que não há como tratar das diferenças linguísticas, sejam elas numa abordagem diacrônica - em caso de mudança -, ou sincrônicas - em caso de variação - sem levar em conta vida social da comunidade e as pressões sociais existentes nela.

Parece, então, que o pressuposto para a categorização tanto sociocultural como sociolinguística faz uso das relações e dos contatos externos e internos para compor uma região cultural, na concepção de Diegues, e uma comunidade de fala, na visão de Labov.

O método PURR e a curva melódica monotônica.

O método PURR (Prosody Unveiling through Restricted Representation) tem como principal motivação a possibilidade de se avaliar o componente prosódico de falas sintéticas sem que haja interferências de outros componentes da fala.

Segundo Sonntag e Portele (1998), as taxas de erro não podem ser avaliadas automaticamente pela inexistência de uma "prosódia correta". Mesmo com a existência de uma prosódia de referência, com a qual pudéssemos combinar um enunciado,

os desvios numéricos a partir dessa referência não fornecem qualquer informação sobre a distância real entre elas.

Outra observação dos autores é a relação não linear entre os parâmetros acústicos da prosódia e a sua correspondência auditiva. Por isso, grandes variações prosódicas apontadas por uma mudança brusca de F0 podem não significar uma alteração linguisticamente relevante. Por outro lado, pequenas alterações acústicas podem ser cruciais para a aceitação da percepção de uma realização prosódica. Pode-se dizer, pois, que não há padrões estabelecidos para a avaliação da prosódia. Além disso, há influência da qualidade segmental da fala na percepção da prosódia, tornando possível que os resultados sejam diferentes de acordo com o componente segmental empregado, tal como exposto em Peterson e Lehiste (1961). Assim, Sonntag e Portele (1998) propõem o uso de estímulos que deixem somente a informação da curva de F0, estrutura temporal e intensidade. Como este trabalho trata de variedades regionais e não de fala sintetizada, fez-se necessário um método que abordasse os segmentos de cada região analisada, isolando as variações de F0, mesmo que as demais informações prosódicas como a estrutura temporal e a intensidade permaneçam inalteradas. Acredita-se que com essa separação pode-se verificar qual a importância da prosódia e dos segmentos na discriminação das variedades regionais. A descrição dos experimentos realizados está no item seguinte.

Metodologia

No experimento com PURR, foram utilizadas amostras de fala adulta do sexo masculino divididas entre 6 informantes (distribuídas entre as 2 variedades escolhidas [SP e RS]). Ao todo, foram selecionados 6 trechos fluentes de cada sujeito (3 trechos curtos e 3 trechos longos). As frases foram obtidas por meio de gravações de fala espontânea. Os trechos escolhidos para o teste foram subdivididos em 3 trechos curtos (≤ 10

segundos) e 3 longos (entre 15 e 25 segundos). As gravações de São Paulo (Mooca) foram feitas com gravador digital ZOOM H2, com taxa de 44100 Hz. As gravações do Rio Grande do Sul (Pelotas) fazem parte do banco de dados VARX (UFPel), também com taxa de 44100 Hz.

No experimento com PURR, houve um breve treino em que os ouvintes familiarizaram-se com o que é entonação. Para a fase de treinamento, apresentou-se uma frase de cada variedade, apresentada em sua versão não manipulada e manipulada com PURR, com o intuito fazer o ouvinte entender o procedimento. Após o treinamento, 5 ouvintes de São Paulo e 5 do Rio Grande do Sul ouviram uma sequência de 27 trechos curtos e outra com 27 trechos longos. No experimento com curva melódica monotônica, as variações de frequência foram “monotonizadas”, utilizando-se a função *To Manipulation*, do Praat. Com isso, as frases apresentadas aos ouvintes tiveram um ‘platô’ tonal sem que houvesse nenhum tipo de variação de F0. Este procedimento seguiu o tom médio, obtido por meio do aplicativo Exprosodia®, de cada frase manipulada. Para a feitura deste experimento, foi gerada uma senoide no menu *sound*, do Praat, obedecendo os tons médios obtidos pelo Exprosodia. As camadas tonais dos trechos foram substituídas pelas senoides, por meio das funções *extract pitch tier* e *replace pitch tier*.

Resultados

No experimento com PURR, os ouvintes de ambas variedades tiveram, em dados brutos, um reconhecimento satisfatório de suas variedades. Nas tabelas abaixo temos os resultados para os trechos curtos:

Trechos Curtos Pelotas (RS)

	acertos	erros
Valores brutos	57	33
Porcentagem	63%	37%

Trechos Curtos São Paulo (SP)

	acertos	erros
Valores brutos	51	39
Porcentagem	57%	43%

Tabelas 1 e 2: Resultados de discriminação de trechos curtos com PURR.

No que se refere ao número de acertos e sua predominância no teste de percepção dos trechos curtos, aplicando-se o teste qui-quadrado, observou-se que no grupo de Pelotas a diferença foi significativa ($p < 0,05$), enquanto que o mesmo não foi observado no grupo de São Paulo ($p > 0,05$). É importante salientar que um dos ouvintes paulistanos teve um desempenho muito baixo na discriminação. Aplicando-se novamente o teste, subtraindo os julgamentos desse ouvinte, tivemos $p < 0,05$ para os trechos curtos (SP). Pode-se, pois, atribuir o primeiro resultado dos paulistanos a esse ouvinte, o que não espelha o resultado dos demais. Para os trechos longos, temos os seguintes resultados:

Trechos Longos Pelotas (RS)

	acertos	erros
Valores brutos	53	37
Porcentagem	59%	41%

Trechos Longos São Paulo (SP)

	acertos	erros
Valores brutos	59	31
Porcentagem	65,50%	34,50%

Tabelas 3 e 4: Resultados de discriminação de trechos longos.

Com os trechos longos, observou-se que a diferença entre erros e acertos no grupo de Pelotas não foi significativa ($p > 0,05$), enquanto que foi no grupo paulista ($p < 0,05$). Para efeito de comparação, foi feito um teste F para saber se as diferenças de discriminação entre trechos curtos e longos foi significativa, isso entre cada variedade. Os resultados apontaram para uma diferença não significativa ($p > 0,05$), logo a duração de cada estímulo parece não influenciar no resultado. No experimento com curva melódica monotônica, os resultados foram os seguintes:

CMM – RS

	acertos	erros
Valores brutos	44	16
Porcentagem	73%	27%

CMM - SP

	Acertos	erros
Valores brutos	45	15
Porcentagem	75%	25%

Tabelas 5 e 6: Resultados de discriminação de trechos com curva melódica monotônica.

Aplicando-se o teste qui-quadrado, observou-se uma diferença significativa entre erros e acertos em ambas as variedades (valor- $p < 0,05$). Tal resultado demonstra que as escolhas não foram aleatórias.

Considerações finais

Apesar de se tratar de um breve estudo, pode-se ter uma ideia aproximada do quanto os ouvintes são capazes de discriminar sua própria variedade linguística. De fato, a hipótese de Ohala e Gilbert (1978) é válida no âmbito intralinguístico, no tocante à população analisada neste trabalho. Necessita-se ainda de uma maior abrangência nos resultados, a qual obter-se-á com um aumento da população, envolvendo também os ouvintes cearenses. O próximo passo, além do aumento quantitativo da análise, será saber como os ouvintes se reconhecem nos estímulos com PURR. Quais serão, dentro dos estímulos envolvidos, os pontos de ancoragem perceptual que os ouvintes se prendem para aceitarem ou rejeitarem uma variedade? Para tal, ainda será empreendida uma análise de produção e o seu cotejo com os resultados perceptuais.