

Síntese por concatenação de variantes regionais: o falar do Porto

*Silvana Paiva¹, Lurdes de Castro Moutinho^{1,2}
e António J. S. Teixeira^{3,4}*

¹ Centro de Línguas e Culturas, Universidade de Aveiro

² Departamento de Línguas e Culturas, Universidade de Aveiro

³ Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA)

⁴ Departamento de Electrónica e Telecomunicações, Universidade de Aveiro

1. Objectivos

Com o advento do processamento computacional da Língua Portuguesa, onde se inclui a síntese de fala, torna-se necessário criar condições para que, também neste âmbito, as variantes sejam consideradas. Um dos pontos-chave na aceitação dos sistemas de voz, pela generalidade dos utilizadores, condição para a utilização generalizada desta tecnologia, prende-se com a necessidade dos sistemas estarem o mais próximo possível dos utilizadores, isto é, serem capazes de reproduzir as variantes próprias de um determinado espaço dialectal.

Concretamente, é nosso objectivo criar uma voz, de acordo com uma variante dialectal, a do Porto, para o sistema Festival, sistema aberto e com facilidade de adição de vozes para novas línguas, usando a síntese por concatenação de difones, uma das técnicas mais habituais na área da síntese.

Assim, o trabalho confronta alguns aspectos conhecidos da variedade normativa do Português Europeu (PE) com a variedade do Porto. Desta forma, foram seleccionados dois informantes: um, com realizações fonéticas características do falar do Porto; um outro que, não possuindo nenhuma marca própria de um qualquer falar, foi considerado um representante da variedade padrão.

Uma vez que este trabalho descreve e analisa fenómenos linguísticos que estão sujeitos a grande variação (dos quais os falantes podem ou não ter consciência), não pode, nem pretende ser, um estudo exaustivo da variedade dialectal do Porto. O estudo procura sobretudo salientar a importância que a variação linguística pode ter nos estudos linguísticos e nos estudos de síntese de voz.

2. Variantes e variação

Apesar da aparente homogeneidade da Língua Portuguesa desde há longas décadas que é reconhecida a existência de variação no PE, pelo que dialectólogos portugueses, com base nas suas investigações, constituíram o mapa dialectológico dos dialectos

continentais. Nesses mapas, são apresentadas as variantes encontradas em cada uma das áreas linguísticas por eles definida.

De acordo com os estudos dialectológicos que têm sido levados a cabo, podem ser identificados em Portugal Continental dois grupos de Dialectos, com base em diferenças fonéticas muito conhecidas. Cintra (1971(1983)) fala destes dialectos referindo-se ao “português setentrional” e “português centro-meridional” (op.cit., p.145)

Ainda que, em cada dialecto, o autor reconheça subdivisões, a verdade é que toda a zona, numa linha transversal, passando a Norte de Aveiro e abaixo de Castelo Branco pertenceria ao dialecto Setentrional, sendo o restante território continental pertencente ao dialecto Centro-meridional.

Paiva Boléo e Maria Helena Santos Silva (1959,1962), no *Mapa dos Dialectos e Falares de Portugal*, fazem já a distinção entre dialectos, falares e variedades. Boléo reconhece os dialectos mirandês, rioronês e guadranês, considerando todos os outros como falares. Leite de Vasconcelos (1901) refere o dialecto interamnese, o dialecto transmontano, o dialecto beirão e o dialecto meridional.

Apesar da existência de algumas diferenças quanto ao nome atribuído aos dialectos, ao longo do tempo, parece-nos, actualmente, existir um consenso na aceitação da classificação proposta por Lindley Cintra (1971(1983)).

Nos dialectos Setentrionais, destaca-se o falar do Porto, merecedor de uma atenção particular nos vários estudos dialectais.

2.1.1 Alguns aspectos caracterizadores do falar do Porto

Se tentarmos reunir o que ao longo dos estudos referidos tem sido dito acerca das características do falar do Porto, podemos destacar as seguintes:

- Em posição final, depois de ditongo ou consoante líquida [l] e [r], ocorre um fenómeno de paragoge, pela realização da vogal [ə].
- A vogal [o] sofre um fenómeno de ditongação, desde que ocorra em posição acentuada, independentemente do contexto fonético e da posição na palavra.
- Manutenção do ditongo [ow] em posição inicial, medial e final, independentemente da posição do acento.
- A vogal nasal [o~] sofre um fenómeno de ditongação, desde que ocorra em posição acentuada, em posição inicial e medial na palavra, independentemente do contexto fonético.
- As vogais [ɔ], [a] e [e] sofrem um fenómeno de ditongação, desde que ocorram em posição acentuada, antes das consoantes palatais [j], [ɹ] e [ʒ] e da semivogal [j].
- A consoante [v], independentemente do contexto fonético em que ocorre e da posição na palavra, sofre um processo de substituição ocorrendo a bilabial oclusiva [b].

Embora estas características se estendam a outros falares do distrito do Porto, elas são, sem sombra de dúvida, ainda hoje também características dos portuenses. Isto é, quando um habitante do Porto ou do seu distrito se desloca a outras zonas do país, sobretudo para sul, é, pelos seus traços linguísticos, imediatamente reconhecido como “falando à Porto” ou sendo do Porto.

2.1.2 O Falar do Porto: breve estudo sobre a situação actual

Uma vez enunciados os fenómenos fonéticos característicos desta variedade, surge a necessidade de testarmos a sua actualidade. De acordo com toda a informação recolhida nos estudos acima enunciados, tornou-se fundamental, antes de prosseguir com a criação da voz, fazer um estudo experimental sobre estes fenómenos.

Para aferir a actualidade e pertinência dos fenómenos fonéticos referidos por outros autores, decidimos realizar um estudo de caso. Para isso, procedemos a uma gravação informal, assim como à audição informal dos portuenses. Depois de procedermos à segmentação e análise das gravações efectuadas, constatámos que os fenómenos e traços diferenciadores referidos como característicos do falar do Porto ainda se mantêm, tendo mesmo sido encontradas outras variantes fonéticas, o que justifica, não só o aprofundamento do seu estudo, bem como a criação de uma voz com as características deste falar.

Por razões óbvias, os resultados detalhados da gravação acima referida, não cabem no âmbito deste artigo.

3. Construção da Voz

3.1 Definição do inventário de fones

O primeiro desafio na construção de uma voz consiste na definição dos fones necessários. As maiores dificuldades prendem-se, neste caso, com a existência ou não de vogais nasais adicionais, o [E~] e [a~] (Sampson, 1999). Optámos por incluir as duas na lista de fones.

3.2 Definição do corpus a gravar

Como para a síntese por concatenação de difones se torna necessária uma base de dados, contemplando todas as combinações de dois fones passíveis de utilização em palavras e/ou sequências de palavras da língua, é necessário definir um corpus a gravar. Muitas vezes, para simplificar a tarefa de criação do corpus, opta-se por sequências de fones que não constituem palavras. No nosso caso, optámos por utilizar palavras, de forma a tornar a tarefa de elocução pelo informante mais natural.

Começámos por produzir, automaticamente, todas as combinações de dois fones (mais o silêncio). De seguida, e por iterações sucessivas, foram procuradas palavras contendo as sequências. Depois de preenchidos os casos mais simples com palavras

comuns, seguiu-se a busca sistemática para certos tipos de sequências mais problemáticas. Numa fase mais avançada, foram consideradas sequências apenas existentes entre palavras e estrangeirismos. Foi feita uma pesquisa bastante minuciosa em dicionários de estrangeirismos (Machado, 1994 e Schmidt-Radefeldt, 1997) e na Linguateca. Pretendia-se, assim, dotar a voz da capacidade de pronunciar palavras estrangeiras, pelo menos as resultantes dos recentes movimentos migratórios, muito visíveis na região do Porto.

A elaboração da lista de palavras teve em conta: a substituição do [v] pelo [b], a utilização frequente de [@] no final de palavra, a consonantização e palatalização de vogais nasais em posição final (*sim*, [si~J@]) e a redução do ditongo nasal [ɔ~w~] à vogal nasal [o~], traços vulgarmente atribuídos a este dialecto. Também foram considerados os fenómenos de ditongação, sendo disso exemplo as palavras *toda* ['todɔ] e *pote* ['pOt@], que o portuense realiza ['twɔdɔ] e ['pwOt@], assim como *malha* ['maLɔ] e *tenho* [tɔJu], realizadas desta feita como ['majLɔ] e ['tɔjJu].

O nosso corpus é constituído por cerca de 1100 palavras, correspondentes ao mesmo número de difones, num total de 1599 difones teoricamente possíveis.

A lista de palavras do nosso corpus foi constituída utilizando o alfabeto fonético SAMPA e as transcrições foram realizadas segundo a variedade normativa. Apenas foram transcritas de acordo com a variante do Porto as sequências que nos parecem ser específicas e pertinentes para o estudo.

3.3 Gravação e anotação

Após a construção da lista de palavras, passámos à gravação do corpus. Utilizámos o sistema EMU (<http://emu.sourceforge.net/>) para a gravação, visto ser um programa que nos facilitava as tarefas de gravação e anotação, tendo acesso, ao mesmo tempo, ao espectrograma e ao oscilograma.

Foram seleccionados dois informantes: um, com realizações fonéticas características do falar do Porto (Raul); um outro (Luís), representante da variedade padrão.

A gravação foi feita no laboratório de Fonética do Centro de Línguas e Culturas da Universidade de Aveiro.

A anotação foi feita inicialmente de forma automática e, posteriormente, sujeita a uma correcção manual, de forma a eliminar alguns erros da anotação automática. Considerámos pertinente a anotação não só do difone em estudo, mas de todos os fones de cada uma das sequências, para obtermos o máximo de informação possível.

Os detalhes referentes à criação da voz, bem como os comandos usados no processo, estão fora do âmbito deste artigo.

4. Teste perceptual para avaliação da qualidade da voz

Para a avaliação da inteligibilidade dos sinais produzidos, foi efectuado um teste para a avaliação das vozes. De entre os vários testes possíveis, foi escolhido o mais utilizado: o método de classificação de estímulos em cinco categorias, entre o mau e o excelente, que corresponde a uma pontuação de um a cinco e do qual resulta uma opinião média. Este teste é habitualmente designado por teste *Mean Opinion Score* (MOS).

4.1 Procedimento

O teste MOS foi realizado, recorrendo a um programa de computador (Teixeira, 2000), cuja interface com o utilizador permitia aos ouvintes atribuir a classificação utilizando o rato, permitindo, também, a audição do estímulo as vezes que o ouvinte desejasse. Só depois de classificar cada estímulo é que o computador realiza o estímulo seguinte.

A pergunta realizada aos ouvintes foi: “A voz é do Porto?”

Na classificação dos estímulos, os ouvintes utilizaram a seguinte escala: 1 – Não é de certeza do Porto; 2 – Não parece ser do Porto; 3 – Talvez sim, talvez não, não consigo decidir; 4 – Talvez seja do Porto; 5 – É com muita certeza do Porto.

Para permitir obter informação acerca da consistência das respostas de cada ouvinte e para se obter um número maior de avaliações, cada estímulo foi apresentado duas vezes para avaliação. A ordem de apresentação foi aleatória, no entanto, para cada ouvinte, a sequência era a mesma. O teste foi realizado individualmente, numa sala com ruído baixo a moderado, sendo os estímulos apresentados a um ouvinte de cada vez, com utilização de auscultadores.

4.2 Estímulos

Optámos por escolher pequenas frases que foram retiradas aleatoriamente do jornal *O Público*, nos dias 31 de Janeiro e 9 de Fevereiro do presente ano de 2005 que, depois de normalizado, isto é, com os números, as siglas e as abreviaturas tratados, constitui o nosso *corpus*. Devemos referir que, ao escolher as frases, houve a preocupação de seleccionar algumas, como é o caso da frase “o dever de votar”, onde os fenómenos característicos da variedade do Porto tivessem maior probabilidade de ocorrência. Estas deveriam facilitar ao ouvinte a sua decisão, permitindo-lhe reconhecer a voz como sendo do Porto e avaliando-a, por isso, com o máximo da pontuação.

Após a normalização, procedemos à transcrição automática das frases recolhidas, usando um sistema de conversão grafema-fone, baseado em regras, em desenvolvimento no Centro de Línguas e Culturas, Universidade de Aveiro (Oliveira *et al.*, 2004). Efectuámos manualmente a correcção/verificação das anotações geradas pelo sistema automático, de forma a colmatar as incongruências existentes.

De seguida, retirámos algumas pausas, uma vez que a nossa voz não tem implementados os módulos referentes à prosódia. Finalmente, procedemos à adaptação manual da transcrição para a pronúncia do Porto.

Cada uma das realizações obtidas para cada uma das frases foi sintetizada, utilizando ambas as vozes. Desta forma, foi possível obter quatro estímulos para cada frase: a voz standard com as regras standard, a voz standard com as regras da pronúncia do Porto, a voz do Porto com as regras standard e a voz do Porto com as regras da pronúncia do Porto.

4.3 Os ouvintes

Participaram nos testes 6 ouvintes, sendo três do sexo masculino e três do sexo feminino, todos falantes nativos do Português Europeu, naturais e residentes, na sua maioria, na zona litoral norte do país. As idades variavam entre os 22 e os 28 anos. Os níveis de escolaridade cobrem um intervalo situado entre o ensino secundário e a licenciatura, inclusive. Todos os ouvintes têm contacto com o Falar do Porto, apresentando, também eles, algumas características desse falar.

Para aferir a consistência das respostas de cada ouvinte, foi calculada a percentagem das decisões iguais entre as duas repetições da totalidade dos estímulos (coluna % igual). Como uma pequena diferença é muitas vezes insignificante, calculámos também as vezes que as respostas diferiam no máximo de 1 ponto.

Os maus resultados de um dos ouvintes na consistência entre as duas repetições levaram-nos a não o considerar para processamento.

4.4 Os resultados

Numa primeira análise, interessa-nos saber a média de cada voz, com e sem conversão grafema-fone, adequada à variante em estudo.

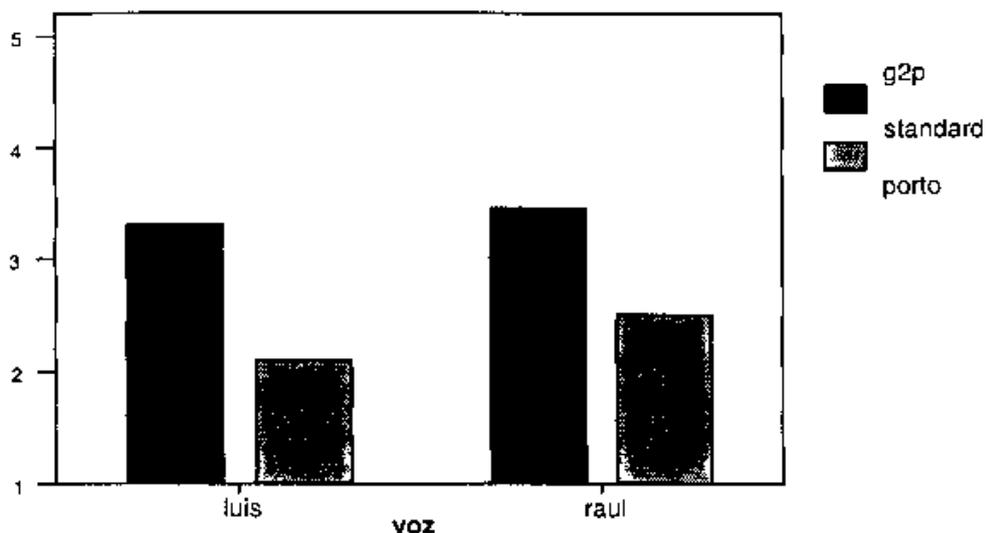


Figura 1 – Média de aceitação das vozes com e sem regras de conversão grafema-fone

Na Figura 1, encontramos do lado esquerdo a voz do Luís e do lado direito a voz do Raúl. As barras mais escuras representam as vozes onde foram aplicadas as regras passíveis de caracterizar a voz do Porto, sendo as outras respeitantes às vozes com as regras com características da variedade normativa.

Em termos médios, as vozes obtêm uma melhor classificação quando são aplicadas as regras do Falar do Porto. Ao aplicarmos as regras standard, os ouvintes preferem, apesar disso, a voz do Raúl, como podemos observar pelos resultados apresentados na figura.

Se analisarmos as vozes, tendo em conta a variação da respostas (dispersão), usando um intervalo de confiança a 95% para média, obtemos a Figura 2 abaixo apresentada.

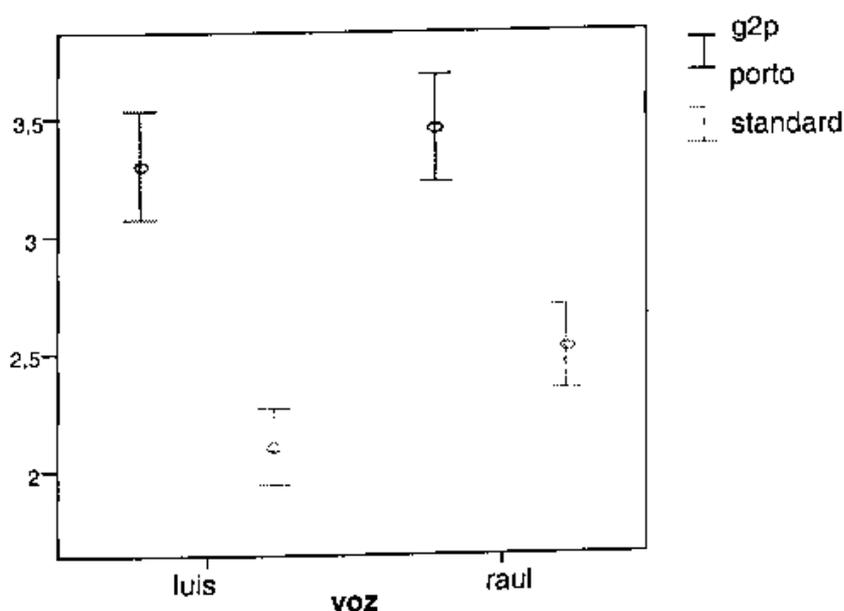


Figura 2 – Média de aceitação das vozes com e sem regras de conversão grafema-fone com um intervalo de confiança de 95%

A diferença entre a voz do Raúl e a do Luís, com as regras de conversão standard, é significativa. Os seus intervalos de confiança não se sobrepõem. Também é significativa a diferença entre as vozes com e sem a utilização das regras de conversão grafema-fone do Porto.

4.4.1 Resultados para as várias frases

Numa segunda análise, exemplificamos e apresentamos algumas frases, cujos resultados são pertinentes para este estudo. Estes resultados representam a avaliação média atribuída por todos os ouvintes.

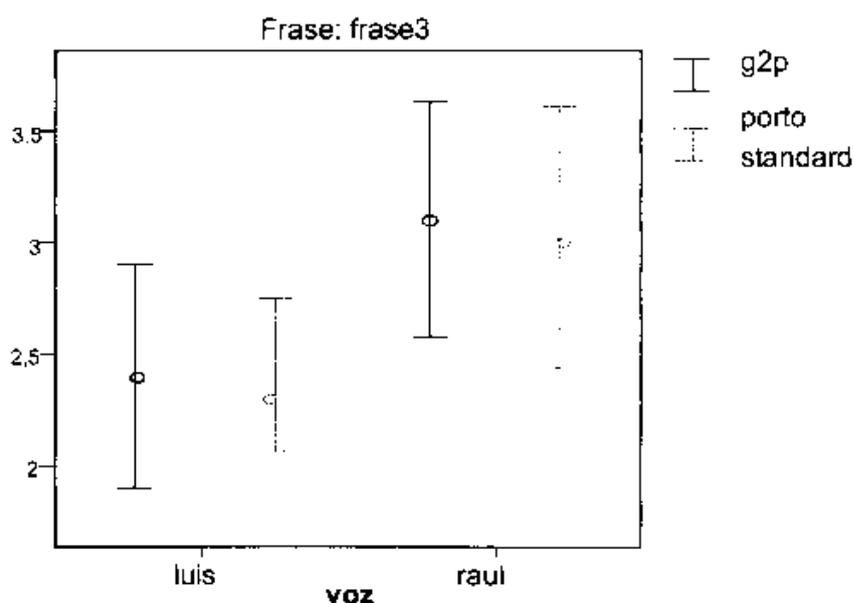


Figura 3 – Resultados da aceitação das vozes para a frase 3

A frase 3, na Fig. 3, “Quase metade das famílias portuguesas habita em casas frias”, mostra que quando aplicadas as regras de conversão grafema-fone relativas à voz do Porto, melhora a qualidade das duas vozes. Manifesta-se uma preferência pela voz do Raúl, mesmo aplicando as regras do Porto à voz do Luís. Assim, a voz do Raúl sem aplicação de regras do Porto, continua a ser manifestamente preferida. A voz parece ser determinante para distinguir as produções com pronúncia standard, das produções com pronúncia da variante do Porto. Para efeitos de síntese, parece ser importante que a voz gravada contenha já algumas características da variante que se pretende sintetizar.

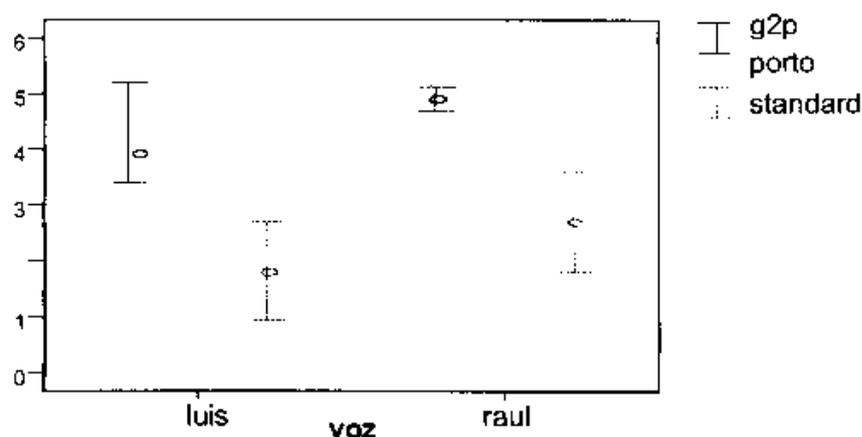


Figura 4 – Resultados da aceitação das vozes para a frase 5

Remetendo para a frase 5 (Fig. 4), “futebol clube do porto foi hoje agraciado pelo governo com a medalha de mérito turístico”, demonstra que a junção dos dois factores, a voz e as regras de conversão grafema-fone da variante do Porto, permite aos ouvintes classificar, com poucas dúvidas, a frase. Apesar da voz do Luís, com a aplicação das regras de conversão grafema-fone relativas à Porto, ter uma boa classificação, ao

observarmos o gráfico é notória a importância da classificação da voz do Raúl com as regras do Porto, resultando uma média próxima de 5, obtendo, assim, a pontuação máxima. Isto é: “é com muita certeza do Porto”, principalmente num intervalo de confiança de reduzida amplitude.

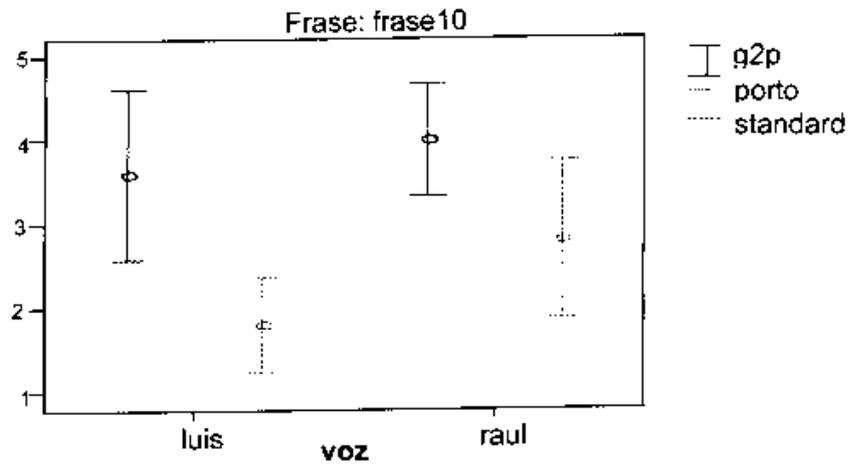


Figura 5 – Resultados da aceitação das vozes para a frase 10

Na frase 10, “bloco diz que a igreja não deve tomar posições políticas”, confirma-se a preferência pela voz do Raúl.

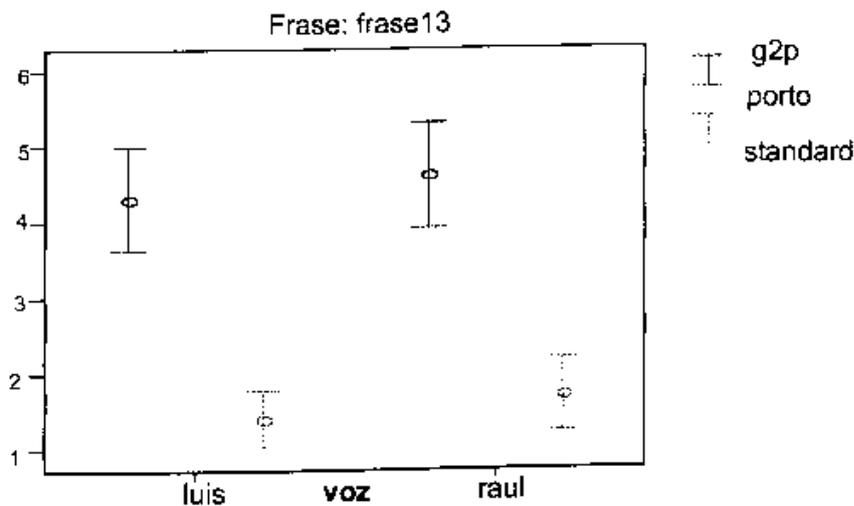


Figura 6 – Resultados da aceitação das vozes para a frase 13

Uma última frase, a 13, “o dever de votar”, realça a importância das regras e sua inclusão na criação das vozes. A diferença nas classificações é, de facto, notória, principalmente quando se trata de frases de tamanho reduzido.

4.4.2 Influência da diferença entre a pronúncia do Porto e standard

Numa terceira análise, comparámos os resultados, tendo em conta a ocorrência ou não de fenómenos fonéticos característicos do Falar do Porto.

Para cada frase, foram determinados os fenómenos modificadores de pronúncia entre o standard e o Porto. De seguida, foram contadas quantas inserções, alterações e elisões ocorreram, utilizando-se a soma desses valores como indicativo do grau de mudanças.

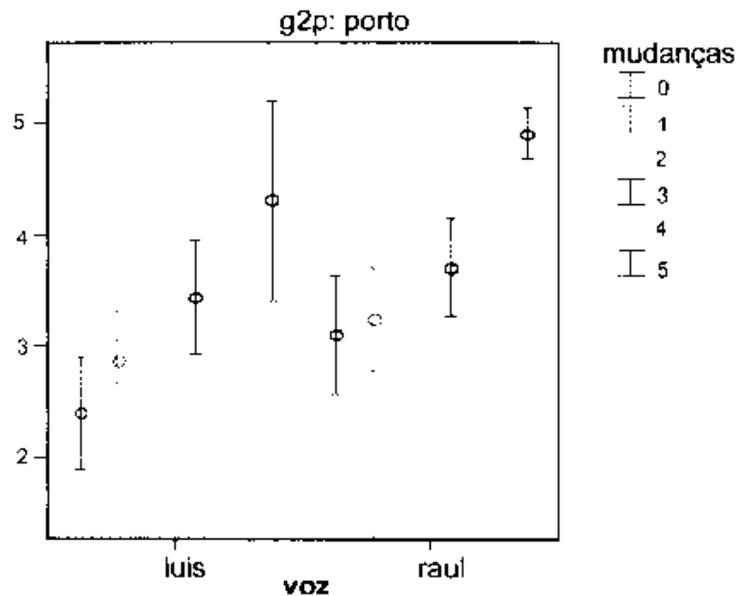


Figura 7 – Avaliação das vozes em função do número de mudanças com as regras de conversão grafema-fone do Porto

Analisando a figura, tendo em conta a realização das frases, aplicando as regras de conversão grafema-fone relativas à voz do Porto, verificamos que as vozes apresentam comportamentos diferentes.

É notório que a classificação obtida pela voz do Luís está intrinsecamente ligada ao aumento das alterações/fenómenos característicos do Porto. Ou seja: quanto mais regras aplicarmos na voz do Luís, mais elevada é a sua classificação. Este aumento é progressivo como podemos constatar na figura.

A voz do Raúl é mais insensível à inserção dos fenómenos, apresentando-se de forma geral com um valor médio próximo de 3. Apenas quando lhe adicionamos o número máximo de regras, e só neste caso, é que o seu valor difere, obtendo uma pontuação muito próxima do 5, a pontuação máxima.

As regras são também um factor relevante para a aceitação/rejeição da voz como sendo do Porto.

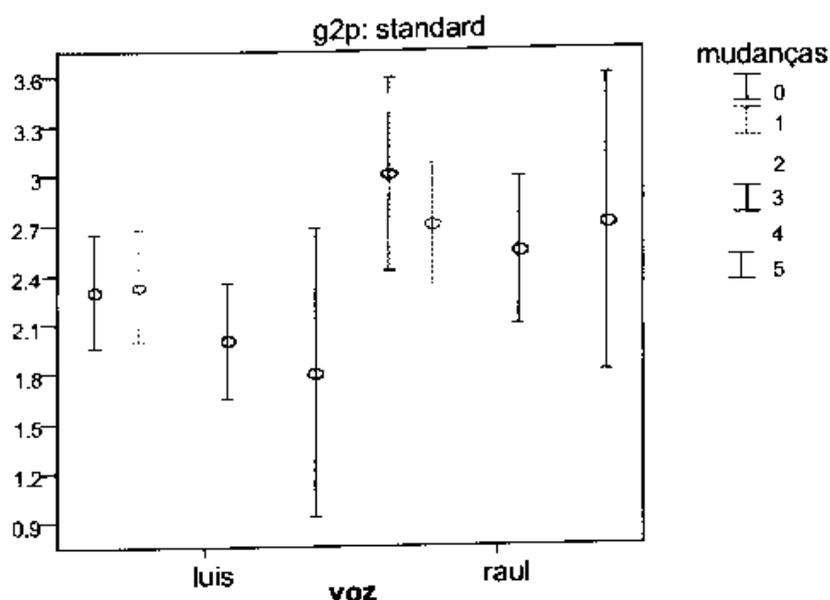


Figura 8 – Avaliação das vozes em função do número de mudanças com as regras de conversão grafema-fone standard

Fazendo uma análise similar, na figura 8, agora com a aplicação das regras de conversão grafema-fone da voz standard, concluímos que, sem as regras características de um falar, a voz propriamente dita torna-se muito mais importante para a avaliação das realizações produzidas.

Quanto mais diferenças existem para a voz do Porto, pior é a classificação da voz do Luís com pronúncia standard. A sua pior média encontra-se quando a frase apresenta o maior número de fenómenos do Porto.

A voz do Raúl, com pronúncia standard, mantém-se com valores superiores. Esta apresenta-se mais imune às regras, apresentando novamente uma média alta.

5. Comentários finais

Motivados pela variação linguística do Português Europeu e da inexistência de vozes sintéticas referentes a essas mesmas variantes, trabalhamos na criação de uma voz sintética do falar do Porto, que denominamos de Voz do Porto.

Surgida a necessidade de termos outra voz, característica da variedade normativa, que nos permitisse avaliar a voz do Porto em comparação com a standard, gravámos o mesmo material, mas desta vez com um informante com características da variedade padrão, repetindo todo o processo da criação da voz.

O resultado principal deste trabalho é a voz criada. Esta voz poderá ser usada como ferramenta em testes de percepção, para testar as diferentes produções na realização de palavras.

Outro dado importante é a obtenção da confirmação, com base nos resultados obtidos, de que a criação de vozes para variantes, como o Falar do Porto, passa, não só pela criação e aplicação de regras de pronúncia adequadas, mas também pela criação de vozes com falantes dessa variante.

Para além disso, constatámos que, quando a variante em estudo se afasta da realização da considerada standard, a inclusão das regras de pronúncia da variante em estudo permite obter bons resultados. No entanto, em muitas situações, principalmente quando a variante se aproxima da realização considerada standard é necessária a utilização de uma voz da variante que pretendemos criar para a obtenção de bons resultados.

Este trabalho tem constituído um desafio e uma forma interessantíssima de aplicação e ponto de partida para investigação na área da Linguística.

Referências Bibliográficas

- BOLEÃO, M. P. e Santos Silva, M. H. (1961) O Mapa dos dialectos e falares de Portugal Continental. *Boletim de Filologia*, 20.
- CINTRA, L. F. L. (1983) *Estudos de Dialectologia Portuguesa*, Sá da Costa.
- LEITE de Vasconcellos (1901) *Esquisse d'une Dialectologie Portugaise*. Thèse de Doctorat, Université de Paris.
- MACHADO, J. Pedro (1994) *Estrangeirismos na Língua Portuguesa*. Coleção Linguística, Editorial Notícias.
- MOUTINHO, L., (1988) *Analyse Sociolinguistique du Parler de Porto: Etudes phonétiques et phonologiques*. Thèse Doctorat, Université Strassbourg.
- MOUTINHO, L. (2001) *Falar do Porto com todos os bês*. Campo das Letras, Porto.
- OLIVEIRA, C. et al. (2004) Um novo sistema de conversão grafema-fone para o PE baseado em transdutores. *Actas do II Congresso Internacional de Fonética e Fonologia*, Universidade Federal do Maranhão (no prelo).
- SAMPSON, R. (1999) *Nasal Vowel Evolution in Romance*. Oxford: Oxford University Press.
- SCHMIDT-Radefeldt, J. (1997) *Dicionário dos anglicanismos e germanismos na Língua Portuguesa*. Frankfurt: TFM.
- TEIXEIRA, A. (2000) *Síntese Articulatoria das Vogais Nasais do Português Europeu*. Tese de Doutoramento, Universidade de Aveiro.