

Contributos para a caracterização das vogais nasais do Português

Liliana Ferreira⁽¹⁾, Rosa Lidia Coimbra^(2,3),
Lurdes Castro Moutinho^(2,3) e António Teixeira^(1,4)

⁽¹⁾Instituto de Engenharia Electrónica e Telecomunicações

⁽²⁾Centro de Línguas e Culturas, Universidade de Aveiro

⁽³⁾Departamento de Línguas e Culturas, Universidade de Aveiro

⁽⁴⁾Departamento de Electrónica e Telecomunicações, Universidade de Aveiro

1. Introdução

As vogais nasais são sons complexos, com uma natureza dinâmica, que apresentam várias dificuldades à sua caracterização completa. Em parte devido a estas dificuldades existem algumas áreas em que é necessária mais informação. Áreas como a variação regional, ou não, das características das vogais nasais do Português Europeu (PE), o comportamento da fonte glotal na produção destes sons, ou ainda a forma como as vogais nasais são produzidas e percebidas, são áreas pouco exploradas e que carecem de informação e estudo.

Em estudos anteriores foi utilizada a tecnologia *Electromagnetic Articulography* (EMMA) com o objectivo de obter mais informação tendo em vista as vogais nasais do PE. Foi desse modo obtido um corpus de aproximadamente uma hora com medição simultânea dos movimentos do véu palatino e orais (Teixeira e Vaz, 2001). No entanto esta técnica não é viável para um número elevado de oradores o que faz com que o sinal de voz, eventualmente complementado com sinais de fácil aquisição, tais como, o do Electroglotógrafo (EGG), seja a única informação passível de obtenção em estudos contemplando a variação regional. Deste modo, foi adquirido, no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”, tendo por parceiros o Centro de Línguas e Culturas da Universidade de Aveiro e o Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA), um novo *corpus* contemplando as várias regiões de Portugal Continental (Alentejo, Algarve, Beira Interior, Beira Litoral, Minho e Trás-os-Montes) contendo, quer o sinal de voz, quer o sinal do EGG (Teixeira *et al.* 2004), que é denominado no restante documento de *corpus* REGIONAL.

2. Objectivos

Este trabalho continua a análise relativa aos sons nasais do Português Europeu efectuada durante o projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português”. Pretende-se efectuar análises a todo o material de voz recolhido, o que

não foi possível durante a duração do projecto, e aproveitar, também, para estender os estudos a novas questões.

Com estas análises pretende-se o melhoramento do conhecimento sobre os sons nasais, em particular as vogais. O principal objectivo é a produção das várias vogais nasais do PE, usando o sintetizador articulatório da Universidade de Aveiro com qualidade, se possível, superior à resultante de trabalhos anteriores. No entanto, existem outras aplicações, tais como, a utilização do conhecimento adquirido para a melhoria do ensino da Língua Portuguesa e para materiais de terapia de fala, entre outros. Concretamente, pretende-se apresentar os estudos mais recentes e seus resultados. O primeiro tema abordado diz respeito à caracterização da fonte durante a produção das diversas vogais nasais. O segundo problema prende-se com a caracterização temporal das várias fases de uma vogal nasal quando produzidas entre oclusivas, e por último obtém-se informação sobre a configuração oral assumida em diferentes momentos, início e fim, da realização de uma vogal nasal nesse mesmo contexto.

Nestes estudos foram utilizados, quer o *corpus* REGIONAL, quer o *corpus* EMMA (Ferreira, 2005).

3. Características da Fonte

Para a continuidade do trabalho em síntese articulatória, em desenvolvimento na Universidade de Aveiro, torna-se imprescindível a obtenção de dados sobre a produção dos sons do PE, em particular, sobre a produção das diversas vogais nasais. O sintetizador articulatório necessita, para além da configuração dos articuladores, de informação sobre parâmetros relacionados com a excitação glotal, dos quais se destacam a frequência fundamental (F0) e o quociente de abertura (OQ). Análises focando este tipo de parâmetros tinham já sido efectuadas no âmbito do projecto “Fonética Aplicada ao Processamento de Fala: As Nasais do Português” (A. Teixeira *et al.* (2004b)). No entanto, o estudo foi realizado tendo por base informação relativa a, apenas, 6 dos 25 informantes.

Actualmente, com todo o *corpus* segmentado e anotado torna-se relevante reanalisar estes parâmetros.

A obtenção dos parâmetros relacionados com a excitação foi realizada de uma forma automática sobre todos os ficheiros do *corpus* Regional. Para tal foi utilizado o sistema *Speech Filing System* (SFS), em particular a linguagem *Speech Measurement Language* (SML), que permite manipular os dados provenientes deste sistema.

De uma forma geral, a análise baseou-se na procura automática, para cada ficheiro do *corpus*, de todos os segmentos anotados como vogal (oral e nasal) e na obtenção dos parâmetros em estudo: F0, OQ e *jitter*. A informação sobre o falante, a região, o género do falante, a tabela, o item, etc., foi também guardada e posteriormente utilizada em análises estatísticas, realizadas com o auxílio do programa SPSS¹.

¹ <http://www.spss.com>

Para possibilitar o estudo de uma possível variação dos parâmetros, a F0 e o *jitter* foram calculados em pontos distintos da vogal. No caso da F0, o seu valor foi calculado em três pontos: próximo do início (concretamente a 10% da duração da vogal), no meio e próximo do final (a 90% da duração). O *jitter*, foi calculado recorrendo ao *Jitter Factor* (JF) o qual foi utilizado para obter valores de *jitter* em toda a duração da vogal e com uma janela de comprimento fixo no início, no meio e no fim.

3.1. Resultados

3.1.1. Frequência Fundamental

A variação do valor de F0 ao longo das vogais (início, meio e fim) em função da vogal é apresentado na Figura . Nesta figura, que ilustra os resultados obtidos para vogais nasais acentuadas, é possível verificar uma diminuição de F0 no final da realização das vogais.

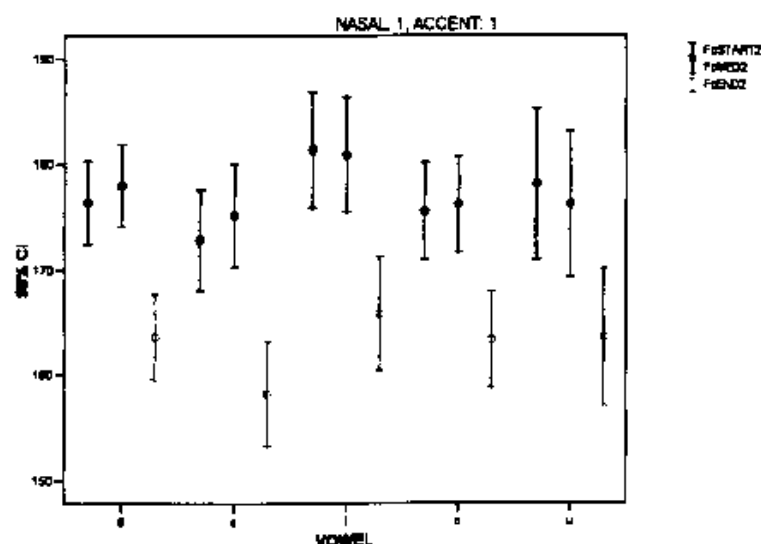


Figura 1: Intervalo de confiança para os valores de F0 em função da vogal, para nasais acentuadas.

Tendo-se verificado uma elevada variação nos valores deste parâmetro por falante efectuou-se um estudo com o objectivo de diminuir a influência deste factor. A Figura representa o diagrama de extremos e quartis para o valor de F0, obtido na parte medial da vogal (F0MED) e em função da vogal nasal, após uma normalização do seu valor pelo valor mediano de cada falante. A figura indica a existência de uma ligeira diferença no valor de F0MED para as nasais [i~] e [u~], sendo esta diferença especialmente notória no caso da vogal nasal [i~].

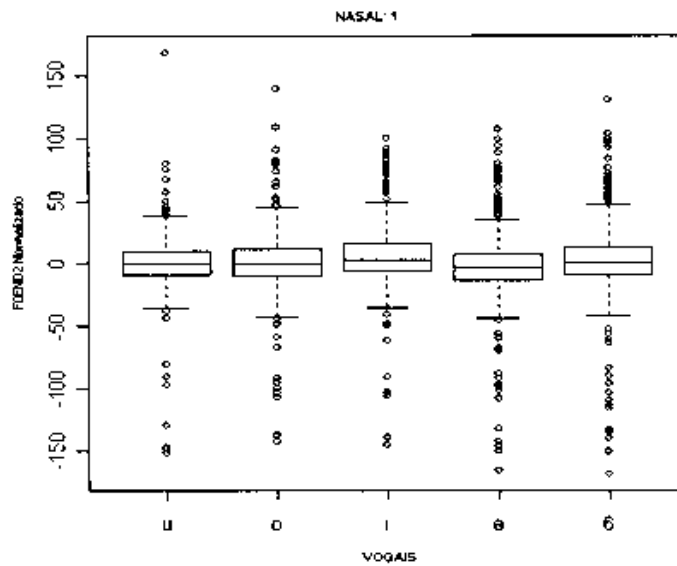


Figura 2: Diagrama de extremos e quartis dos valores de FOMED em função das vogais nasais após divisão pela mediana de FOMED.

3.1.2. Quociente de Abertura

A Figura 3 apresenta o diagrama de extremos e quartis dos valores de OQ em função da vogal e da nasalidade. É possível notar uma reduzida variação, quer para os valores medianos, quer para a dispersão dos valores. Esta análise é válida, quer para as vogais orais, quer para as correspondentes nasais.

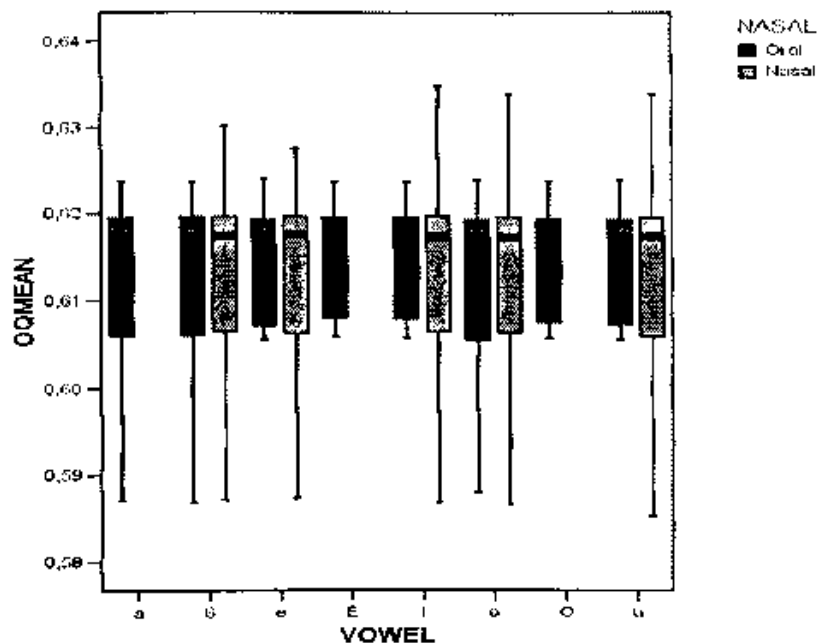


Figura 3: Diagrama de extremos e quartis dos valores do quociente de abertura em função da vogal e nasalidade.

3.1.3. Jitter

A Figura sumaria estatisticamente os valores de jitter obtidos para cada vogal, em função da nasalidade. Nesta primeira análise gráfica, verifica-se uma reduzida variação dos valores obtidos. No entanto, após uma análise mais detalhada dos valores médios correspondentes, apresentados na Tabela 1, evidencia-se a existência de um valor inferior do *jitter* quando medido a meio da vogal, relativamente ao início e fim, sendo o valor final o mais elevado.

Médias		
	Oral	Nasal
JFSTART	2,1200406	1,9419101
JFMED	1,6392915	1,4444812
JFEND	2,2557149	2,1352677
JFGLOBAL	2,10506703	1,85222505

Tabela 1: Valores médios obtidos para o Jitter Factor. À esquerda os valores para vogais orais, à direita os valores para vogais nasais.

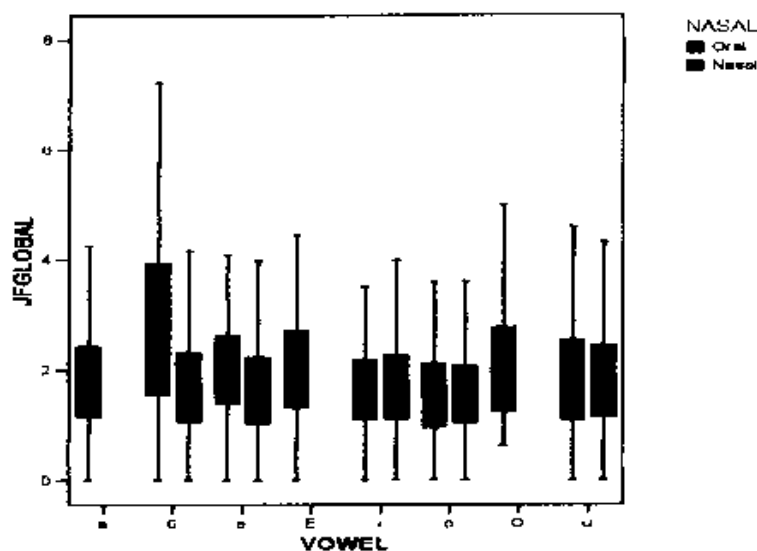


Figura 4: BoxPlot do valor do Jitter Factor para a totalidade da duração da vogal. São apresentados valores para cada vogal separados pela nasalidade.

4. Caracterização temporal

Um estudo da duração total das vogais nasais utilizando o *corpus* Regional, foi já

O novo estudo com o objectivo de explorar o *corpus* Regional, em combinação com o *corpus* EMMA, de modo a obter informação detalhada sobre as diferentes fases de uma vogal nasal quando produzida entre consoantes orais, preenche as secções seguintes. Estes sons nasais iniciam sempre com uma configuração onde domina uma configuração oral e termina com uma configuração nasal dominante (Teixeira e Vaz, 2001).

4.1. Utilizando as anotações dos sensores EMMA

Estamos interessados em obter, por exemplo, a duração da parte inicial da vogal em que o véu palatino se mantém fechado ou quase fechado e a duração do período de abertura e fecho. Concretamente foram estudadas as durações das seguintes fases:

- o tempo entre a abertura oral e início do movimento de abertura do velo – d_1
- a duração do intervalo existente entre o fecho dos lábios e o fecho do velo – d_2
- a duração do movimento de abertura do velo – d_0
- a duração do movimento de fecho do velo – d_c
- a duração do ciclo de abertura-fecho do velo – d_3

Para a obtenção das durações referidas acima continuaram-se as anotações em parte do *corpus* EMMA, nos contextos CVC, em quatro níveis: palavras, fonética, eventos do velo, eventos orais, utilizando o sistema EMU. Replicando o trabalho apresentado por Teixeira e Vaz (2001), agora com mais dados, foram medidas durações das várias fases de produção de uma vogal nasal com o auxílio da ferramenta SFS. Alguns exemplos de resultados são apresentados de seguida.

4.1.1. Resultados

A Tabela 2 apresenta os valores médios obtidos para as durações em análise. A última coluna contém o nível de significância obtido para o teste ANOVA, tendo como factor a vogal. Verifica-se que cada uma das durações apresentadas apresenta um valor para p não significativo, o que significa que as durações apresentadas não são dependentes das vogais.

	[ɛ~]	[e~]	[i~]	[o~]	[u~]	p
d_1	90,7	99,6	64,5	70,1	39,3	0,064 (ns)
d_2	112,9	105,3	72,9	91,0	114,2	0,000 (***)
d_0	181,4	173,1	158,4	176,5	163,1	0,261(ns)
d_c	125,8	128,9	106,3	117,5	128,8	0,223(ns)
d_3	307,2	302,0	264,7	293,9	291,9	0,056(ns)

Tabcla 2: Durações das Fases (ms).

4.2. Com base apenas no sinal acústico: estudo exploratório

Existindo interesse em obter dados similares para um número alargado de pessoas, o que é inviável com a tecnologia EMMA, foi efectuado um estudo exploratório relacionado com a obtenção de informação, sobre as fases referidas anteriormente, com base apenas no sinal acústico.

Para tal foi utilizado um método de segmentação do sinal existente na ferramenta SFS que se baseia na determinação de zonas com características espectrais estáveis. Esta segmentação foi efectuada para 3 e 4 zonas e posteriormente efectuou-se um estudo da correlação entre os limites dos segmentos obtidos e os eventos anotados no sensor do velo e nos sensores que controlam a passagem oral (sensores dos lábios e sensores da língua, em particular o colocado na parte posterior).

4.2.1. Resultados

A elevada correlação existente entre o primeiro segmento obtido por segmentação automática (SEG1) e o primeiro evento anotado no sensor do velo (NASALOPEN), ilustrada na Figura 5, bem como o elevado índice de ajuste entre estes dois parâmetros, permite concluir que a segmentação automática é um assunto que vale a pena ser explorado, pelo menos para obtenção indirecta de eventos relacionados com o velo.

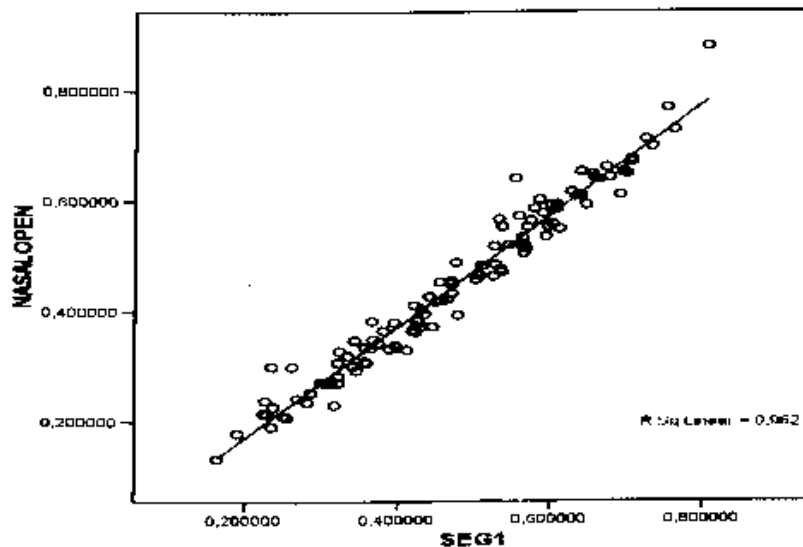


Figura 5: *Scatter Plot* entre o evento de abertura do velo e o primeiro segmento automático.

5. Características relacionadas com o tracto

O objectivo deste tipo de análise passa pela obtenção de informação relativa à configuração dos articuladores, em especial da língua, durante as diferentes fases de uma vogal nasal.

Numa primeira fase foi feito um estudo utilizando as duas primeiras formantes na fase inicial, por forma a obter informação acerca da altura e zona de articulação (Teixeira *et al.* 2003).

Deixando para mais tarde a análise da zona de transição, foi efectuado um estudo com o objectivo de investigar a relação entre as características da fase final de um vogal nasal e a consoante seguinte. Não é claro nos resultados apresentados por outros investigadores (Lacerda e Head, 1966) até que ponto essas características finais são essencialmente dependentes da coarticulação com a consoante seguinte ou se dependem apenas das características intrínsecas das vogais nasais.

5.1. Método

De modo a estudar estas características das vogais nasais, foram obtidas as características espectrais no final das vogais nasais e compararam-se com as características espectrais das consoantes nasais e de outras vogais nasais (a mesma vogal e outras). Para calcular a semelhança espectral utilizam-se diversas medidas que traduzem uma distância espectral. Neste estudo obtiveram-se medidas baseadas em distâncias euclidianas aplicadas a vectores com coeficientes cepstrais em escala Mel (dMFCC), medidas com base no espectro de potência como é o caso das distâncias de Itakura (dITA), Itakura-Saito (dITAS) e Itakura-Saiko simétrica (dCOSH). A implementação foi baseada nos pacotes existentes para o MatLab *sfsmatlab*² e *voicebox*³. Foi ainda implementada uma abordagem baseada em *Perceptual Linear Prediction* (PLP) usando o pacote *rastamat* (dPLP).

5.2. Resultados

Os resultados obtidos apresentam claras contradições entre si. Estas contradições começam pelo facto de as medidas consideradas não atribuírem sempre à distância entre vogais iguais o menor valor. As medidas dPLP e a dCOSH diferenciam-se, no entanto, pelo facto de terem atribuído mais vezes valores correctos (100% para a dPLP e 80% para a dCOSH) a esta distância. Deste modo, apenas se pode analisar aquilo que estas distâncias parecem indicar. Existe uma tendência clara para que a dPLP classifique como menor a distância existente entre vogais iguais ($V_i V_i$), logo seguida da distância entre uma vogal e as outras vogais nasais ($V_i V_j$). Parece, assim, indicar que os espectros do segmento final das vogais nasais se assemelham entre si, corroborando a teoria que defende que as características finais das vogais nasais são intrínsecas. A dCOSH, por outro lado, inverte os resultados anteriores, classificando a distância entre diferentes vogais nasais como superior à distância entre vogais nasais e consoantes nasais. Esta contradição patente nos resultados impede a obtenção de qualquer conclusão. Sugere-se a utilização de outro tipo de medidas de distância espectral para o esclarecimento deste assunto.

² <http://www.phon.ucl.ac.uk/resource/sfs>

³ <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox>

6. Conclusões

Este trabalho explora assuntos pouco analisados em estudos anteriores. Os estudos apresentados têm como objectivo contribuir para uma melhor caracterização das vogais nasais do PE. Existem mesmo assuntos que pretendem esclarecer controvérsias existentes na área, como é o caso do estudo efectuado com o objectivo de analisar a relação das características da fase final das nasais com as consoantes nasais e outras vogais. No entanto, este estudo, bem como o efectuado com o objectivo de obter informação sobre as diferentes fases das vogais nasais de uma forma automática, devem considerar-se exploratórios e os resultados obtidos indicam a necessidade de um maior aprofundamento e a aplicação de técnicas adequadas.

7. Referências

- Ferreira, L. (2005) *Contributos para a caracterização das vogais nasais do Português*. Dissertação de Mestrado, Departamento de Electrónica e Telecomunicações, Universidade de Aveiro (aprovada a 30 de Março de 2005).
- Lacerda, A. e Head, B.F. (1966) Análise dos sons nasais e dos sons nasalizados do português. *Revista do Laboratório de Fonética Experimental de Coimbra*, VI pp. 5-70
- Teixeira, A. e Vaz, F. (2001) European Portuguese Nasal Vowels: An EMMA Study. *EuroSpeech*. Scandinavia.
- Teixeira, A., Castro Moutinho, L. e Coimbra, R. L. (2003) Production, Acoustic and Perceptual Studies on European Portuguese Nasal Vowels Height. *ICPhS*.
- Teixeira, A., Ferreira, L., Moutinho, L., Coimbra, R.L. e Lisboa, R. (2004). An Acoustic Corpus Contemplating Regional Variation for Studies of European Portuguese Nasals. *Proc. 2004 International Conference Language Resources and Evaluation*. Lisboa, Portugal.
- Teixeira, A., Vaz, F., Moutinho, L., Coimbra, R. e Lisboa, R. (2004b) Para a melhoria da síntese articulatória das vogais nasais do português europeu: Estudo da duração e características relacionadas com a fonte glotal. *Revista de Estudos da Linguagem (The Journal of Language Studies)*.