

## **Contributos para um modelo prosódico do Português: Estudo de um caso exemplar**

*Ana Margarida C. Vaz da Silva<sup>1</sup>, Lurdes de C. Moutinho<sup>1,2</sup>  
e António J. S. Teixeira<sup>3,4</sup>*

<sup>1</sup> Centro de Línguas e Culturas, Universidade de Aveiro

<sup>2</sup> Departamento de Línguas e Culturas, Universidade de Aveiro

<sup>3</sup> Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA)

<sup>4</sup> Departamento de Electrónica e Telecomunicações, Universidade de Aveiro

### **1. Algumas motivações**

A apresentação deste trabalho é fruto da investigação realizada para uma dissertação de Mestrado já defendida. Na sua base está um especial interesse relacionado com o processamento computacional da Língua Portuguesa e a sua inserção no Grupo das Ciências da Linguagem (Centro de Investigação de Línguas e Culturas). Este projecto pretende ser uma continuidade do Projecto Atlas Multimedia Prosodique de l'Espace Roman (AMPER) e prolongar alguns dos aspectos nele estudados. Tal pesquisa foi possível dada a já habitual cooperação entre o Laboratório de Fonética e o Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA) – da Universidade de Aveiro (UA).

### **2. Objectivos**

Ao desenvolver esta pesquisa pretendeu-se estudar características diferenciadoras de vários tipos de frase relativamente à entoação, dar continuidade e completar estudos desenvolvidos no Projecto Atlas Multimedia Prosodique de l'Espace Roman (AMPER), em curso, e contribuir para a construção de um modelo da entoação do Português, que possa vir a ser aplicável a sintetizadores de fala numa etapa futura.

### **3. Corpus**

Tal como antes foi referido, este trabalho surgiu como a possibilidade de dar continuidade à investigação que se tem vindo a desenvolver no âmbito do projecto AMPER. No entanto, pretendia-se, para além de nos debruçarmos sobre a mesma matéria, alargar o já desenvolvido. Daí que, na intenção de criar uma extensão do corpus do AMPER, se tenha partido de um corpus construído para o referido Projecto, mas alargando-se o estudo a outros tipos de frases.

A recolha para constituir o corpus foi feita no Laboratório de Fonética do Departamento de Línguas e Culturas da Universidade de Aveiro, numa sala não insonorizada. A gravação foi efectuada directamente para o disco de um computador portátil, usando o programa SFSWin, parte do SFS<sup>1</sup> (Speech Filing System) da autoria de Mark Huckvale do University College London. Foi utilizado o sistema CSL 4400 da Kay Elemetrics Corp. Para além do sinal acústico proveniente de um Microfone Shure SM48, foi gravado o sinal relativo às cordas vocais proveniente de um electroglotógrafo.

Durante a recolha, a produção da frase, com a estrutura e entoação pretendida, foi induzida pela utilização de estímulos visuais, que consistiam num conjunto de desenhos passíveis de induzir o informante à produção da frase pretendida, apresentando cada uma delas a mesma estrutura frásica: Sujeito+Verbo+Objecto, incluindo todas as estruturas acentuais – oxítone, paroxítone, proparoxítone – no Sujeito e no Objecto.

Foram solicitadas ao informante nove repetições das trinta e quatro frases que constituíam o corpus. Em cada repetição foi usada uma ordem diferente e aleatória das frases a produzir, para evitar o efeito de lista. Cada repetição das frases constituintes do corpus foi gravada num ficheiro individual em formato SFS, para posterior anotação.

O corpus foi gravado apenas por um informante, aluno da Universidade de Aveiro e proveniente desta mesma região, que produziu trezentas e trinta e uma frases, posteriormente analisadas.

Na construção deste corpus (Tabela 1) foram considerados os seguintes tipos de frase:

Declarativa: afirmativa (decl) e negativa (neg);

Interrogativa: global (int) e introduzida por pronome interrogativo (-que);

Imperativa: ordem (ord) e que subentende impaciência (int).

Tipo	Subtipo	Frase
Declarativa	Afirmativa simples	O Toneca toca no pássaro. O pássaro toca no Toneca.
	Negativa	O Toneca não toca no pássaro. O pássaro não toca no Toneca.
Interrogativa	Global (afirmativa)	O Toneca toca no pássaro? O pássaro toca no Toneca?
	Com pronomes	Quem toca no pássaro? Quem toca no Toneca?
Imperativas	Ordem	Toneca, toca no pássaro. Pássaro, toca no Toneca.
	Impaciência	Toneca, tocas no pássaro.[ou ♪♫⊗*□] Pássaro, tocas no Pateta.

Tabela 1 – Parte do corpus

<sup>1</sup> Este programa pode ser obtido, gratuitamente, no endereço <http://www.phon.ucl.ac.uk/resource/sfs/>

## 4. Análise

### 4.1 Anotação realizada

Após a recolha, foram identificados manualmente o início e o fim de cada uma das trinta e quatro frases de cada repetição. Com esta segmentação, e utilizando as facilidades do comando *wordchop* do SFS, foram criados ficheiros individuais para cada uma das frases e repetições. A cada um destes ficheiros individuais foi atribuído um nome, em que as três primeiras letras identificam o tipo de frase (exemplo: *dcl* para as declarativas afirmativas); a seguir dois números identificam o número de ordem da frase no corpus, no fim aparece o número referente à repetição após a letra “r”. Os ficheiros têm como extensão “.sfs”.

#### 4.1.1 Anotação manual dos fones

Para cada ficheiro, foi feita a segmentação e anotação manual de cada fone, usando o SFSWin, e o sinal acústico, complementado pelo espectrograma e sinal do electroglotógrafo.

Na anotação, foi utilizado o alfabeto fonético SAMPA<sup>2</sup> e seguidas as indicações acerca da transcrição fonética do CSLU (LANDER:1997). As oclusivas surdas foram divididas em duas partes: uma de silêncio, outra correspondente à “release” – por exemplo, para o [t] temos [t] e [tr].

#### 4.1.2 Anotações “automáticas”

Sendo nossa convicção que o material recolhido deveria ser anotado a outros níveis, como, por exemplo, a delimitação das sílabas, optámos por tentar automatizar, dentro do possível, a criação de todas as outras anotações. Saindo do âmbito deste trabalho uma descrição pormenorizada dos programas desenvolvidos, apenas se apresentam, de seguida, os procedimentos, e de forma resumida, numa perspectiva informativa e com intuito de documentar todo o processo.

Em termos gerais, para cada novo tipo ou grupo de tipos de anotação, foi desenvolvido um programa na linguagem SML (Speech Measurement Language) integrante do sistema SFS, e o respectivo ficheiro “batch” para o executar no ambiente DOS. Cada um desses programas faz uso do nome do ficheiro, nível de anotação manual – anteriormente descrito –, ou níveis criados automaticamente antes dele. Todo o processamento assume o resultado da anotação manual dos fones numa anotação com o nome “Eswin/AN(type=phones)”.

<sup>2</sup> O alfabeto SAMPA (*Speech Assessment Methods Phonetic Alphabet*) é um alfabeto fonético passível de leitura pelo computador sem necessidade de instalação de fonte própria, já que se serve dos caracteres disponíveis no teclado. Foi desenvolvido em finais dos anos 90 no âmbito de um projecto de um grupo de foneticistas e começou por ser aplicado às línguas da Comunidade Europeia. Tabelas e informações disponíveis em <http://www.phon.ucl.ac.uk/home/sampa/home.htm>.

Níveis de anotação automática: Tipo de frase; Estrutura acentual da frase; Sílabas; Palavra; Acento lexical; Vogal / consoante / ditongo; INTSINT /MOMEL.

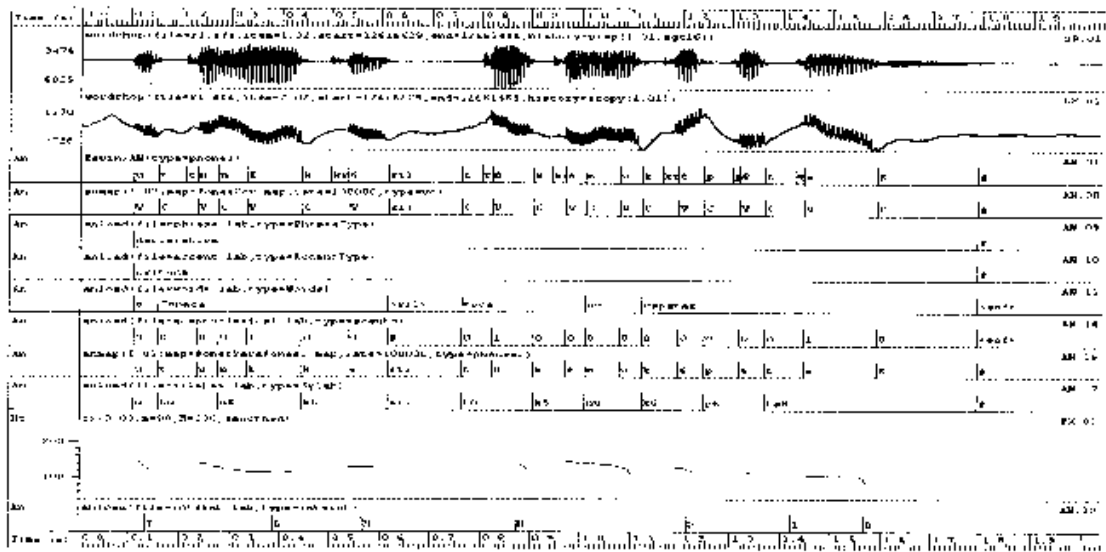


Figura 1 – Exemplo de uma frase devidamente anotada

#### 4.2 Parâmetros analisados

Depois de concluída a anotação, foram analisados os parâmetros da Frequência Fundamental (F0), Duração, Intensidade e Pausas, cujos resultados e conclusões a seguir se apresentam.

### 5. Resultados

#### 5.1 Frequência fundamental

##### 5.1.1 Frases declarativas

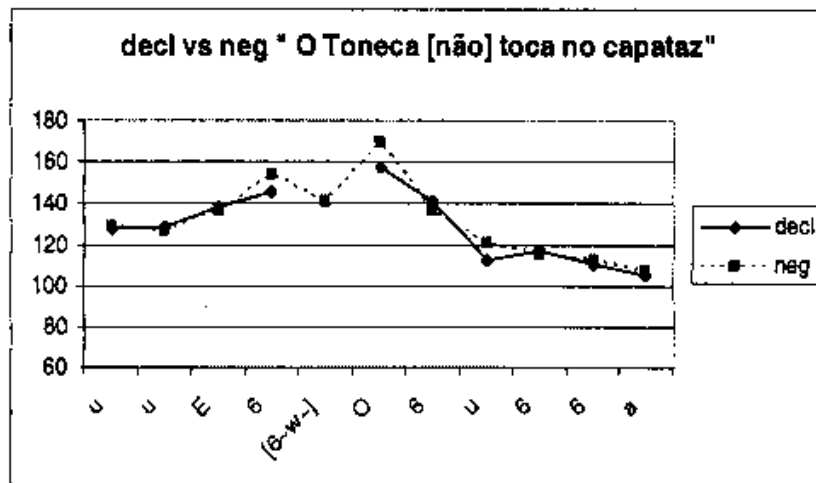


Figura 2 – Exemplo de análise comparativa de frases declarativas afirmativas e negativas.

À primeira vista, a distinção entre a declarativa afirmativa e a declarativa negativa não se faz pela diferença de movimentos da curva de F0, mas parece-nos, segundo os nossos dados, que o que distingue uma da outra é a estrutura sintáctica e a presença do advérbio de negação. Por isso, o advérbio de negação toma a posição de foco.

Há uma grande semelhança entre os dois padrões entoacionais comprovada pela quase sobreposição da representação das curvas melódicas obtidas em ambas as frases no início e final das mesmas.

### 5.1.2 Frases interrogativas

Globalmente a curva melódica destes dois tipos de frases assemelha-se de certa forma até à penúltima sílaba, onde se torna visível a dissemelhança. Assim, podemos verificar que há uma ligeira subida antes do sintagma verbal e em ambos os casos é este que detém o pico da frequência, após o qual a curva sonora desce até ao final, no caso da interrogativa-que, e até à penúltima vogal sofrendo uma ascensão na última vogal tónica, no caso da interrogativa – global.

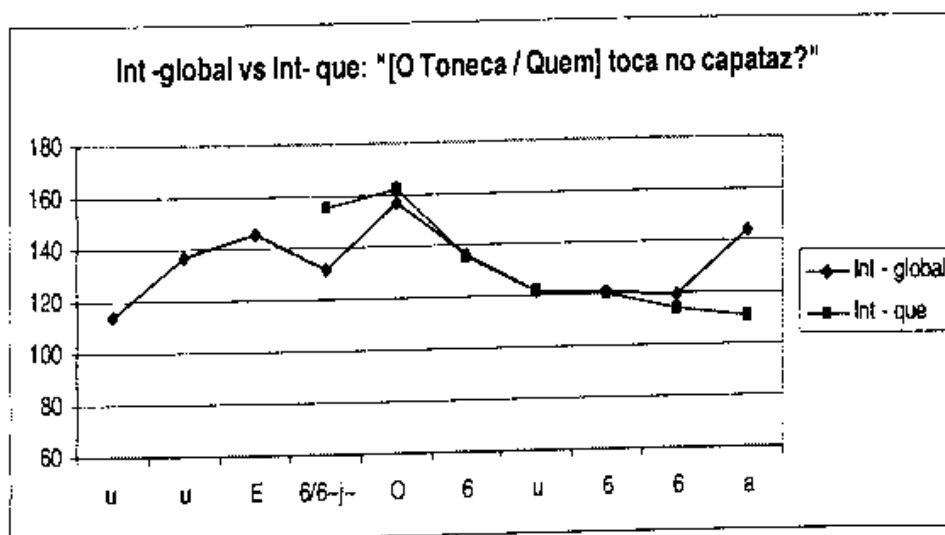


Figura 3 – Exemplo de análise comparativa de frases interrogativas global e introduzida por -que.

### 5.1.3 Frases imperativas

Tal como foi possível verificar noutros exemplos desta pesquisa e que agora não é possível demonstrar pela limitação espacial, as frases parecem assemelhar-se até ao sintagma verbal, onde acontece o pico de F0. Verifica-se, então, que, a partir do predicado, os dois tipos de frase têm um comportamento diferente, porque, enquanto que a imperativa-ordem desce a partir desse ponto, a imperativa – impaciência faz uma ligeira subida de frequência no início da primeira vogal da última palavra da frase, baixando um pouco de seguida, com uma subida de F0 bastante acentuada no final.

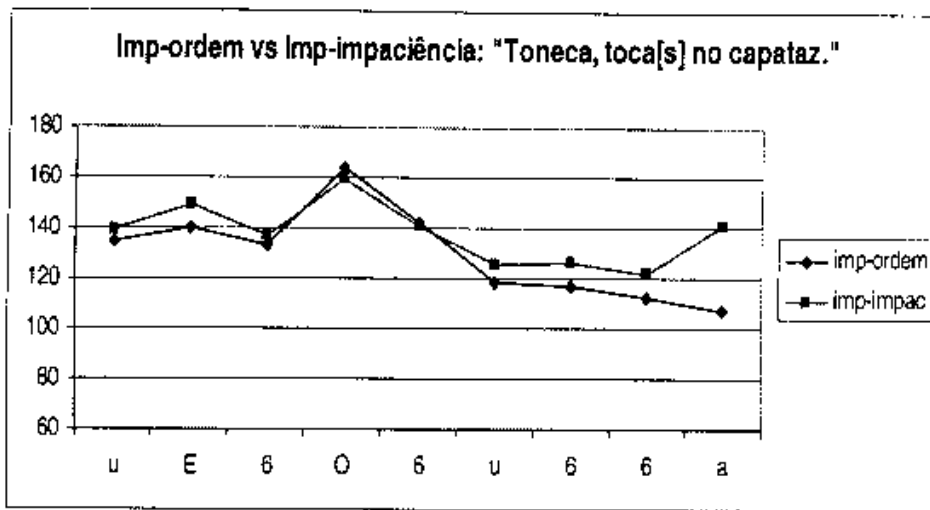


Figura 4 – Exemplo de análise comparativa de frases imperativas de ordem e que subentendem impaciência.

## 5.2. Duração das vogais e das sílabas

### 5.2.1 Vogais

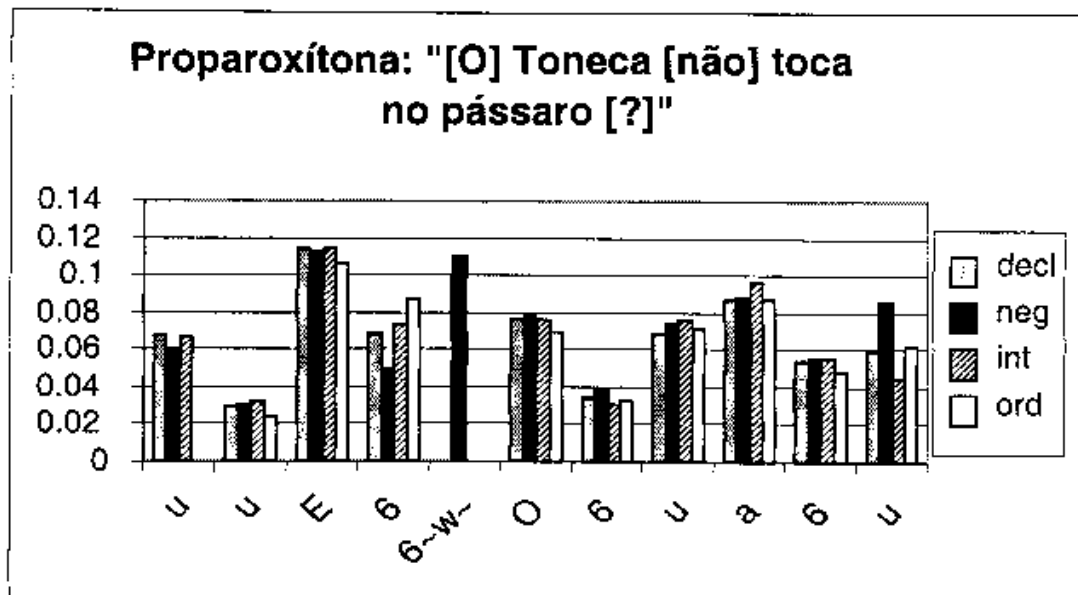


Figura 5 – Exemplo comparativo da duração das vogais em vários tipos de frase<sup>3</sup>

O gráfico acima, que apresentamos a título exemplificativo, permite-nos tecer as seguintes considerações:

<sup>3</sup> As frases tratadas são do tipo declarativo afirmativo (decl), declarativo negativo (neg), interrogativo-global (int) e imperativa-ordem (ord).

As vogais com menor duração são átonas e situam-se nas frases interrogativa – global e imperativa – ordem.

As vogais com maior duração são tónicas e detectam-se na palavra que corresponde ao sujeito ou vocativo (no caso das imperativas – globais) e na última palavra da frase que, pela sua posição, representa o acento frásico. Podemos notar a duração do advérbio de negação que, além de ser o foco da frase, é constituído por um ditongo o que, obviamente, prolonga a sua duração.

A duração é factor demarcador da qualidade da vogal como sendo tónica ou átona, o mesmo não se podendo afirmar, com base nos resultados por nós obtidos, no que diz respeito à distinção clara entre modalidades.

### 5.2.2 Efeito do acento nas vogais

Ao pretender verificar a importância do acento frásico e lexical na duração das vogais, analisámos o efeito destes factores no referido parâmetro aplicando, desta vez, o SPSS, usando ficheiros de dados contendo em cada linha o tipo de frase, o tipo de acento da frase, a posição do fone/sílaba na frase (contando a partir do início), informação acerca do acento, o fone ou sílaba, e o parâmetro em análise.

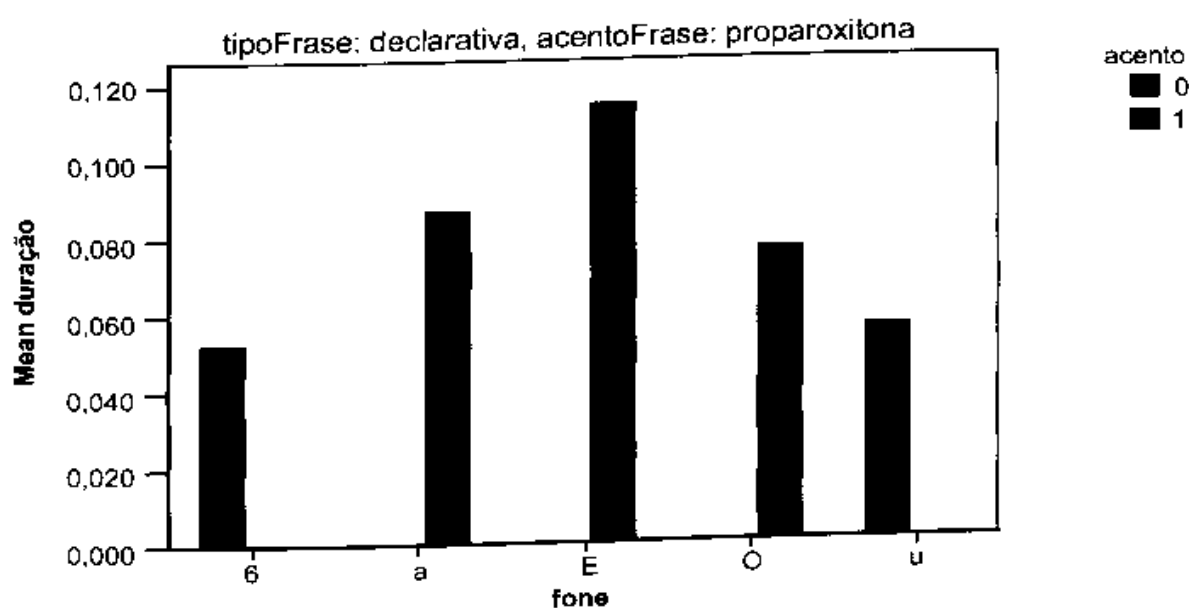


Figura 6 – Exemplo do efeito do acento na duração das vogais na frase [u tu'nEk6 'tOk6S nu 'pas6ru]

Como seria de esperar, verifica-se que as vogais acentuadas são as que apresentam maior duração. Neste caso, a mais longa é [E].

## 5.2.3 Sílabas

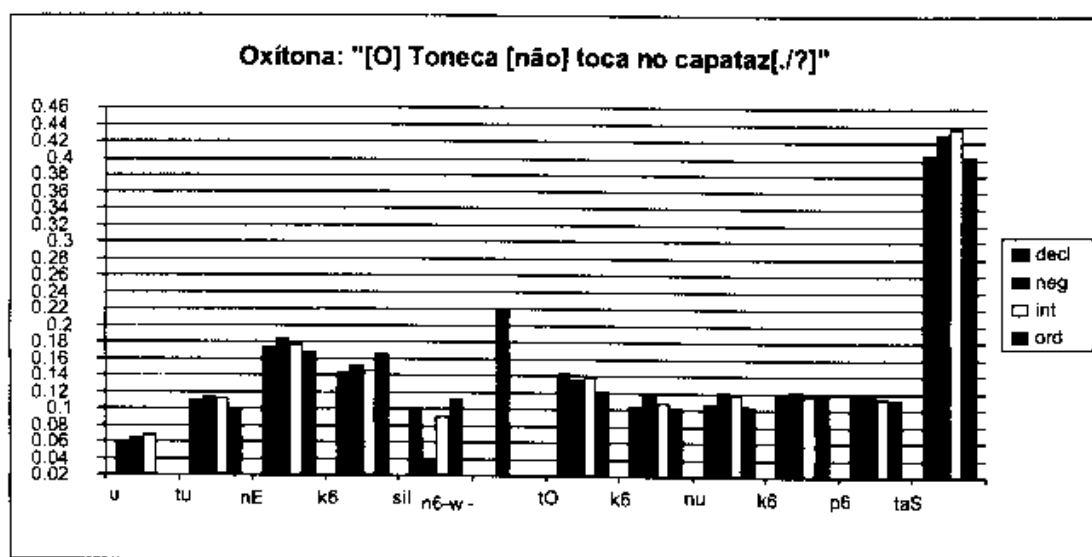


Figura 7 – Exemplo da duração das sílabas em quatro tipos de frase com acentuação oxítona

O gráfico da Figura 7 permite-nos constatar o seguinte: a sílaba tónica, que aparece antes da pausa, é a que tem maior duração.

Todas as frases apresentam silêncio: na frase declarativa tem duração média de 98ms, na negativa (declarativa) 37ms, a frase interrogativa 91ms e a imperativa (ordem) 113ms. É esta última que apresenta uma maior duração do silêncio, provavelmente a fim de demarcar o vocativo, função sintáctica distintiva neste tipo de frase.

O advérbio de negação, foco na frase negativa (declarativa), tem uma duração média de 210ms.

Parece-nos importante referir que, neste tipo de acentuação frásica, apenas a última sílaba é mais longa do que as outras. Não há qualquer distinção da penúltima sílaba ao contrário das outras frases analisadas antes. Como estas frases são oxítonas, a acentuação recai sobre a última vogal. Além disso, a última sílaba é composta por três fones [taS], ao contrário das outras sílabas finais analisadas que tinham apenas dois fones, o que logo à partida aumentará a sua duração, agravada pela presença da acentuação da frase. Deste modo, a sílaba final apresenta uma duração superior a qualquer outra: na frase declarativa esta tem duração média de 404ms, na negativa (declarativa) 429 ms, a frase interrogativa 427ms e a imperativa (ordem) 402ms.



### 5.3 Intensidade

#### 5.3.1 Intensidade em diferentes tipos de frase

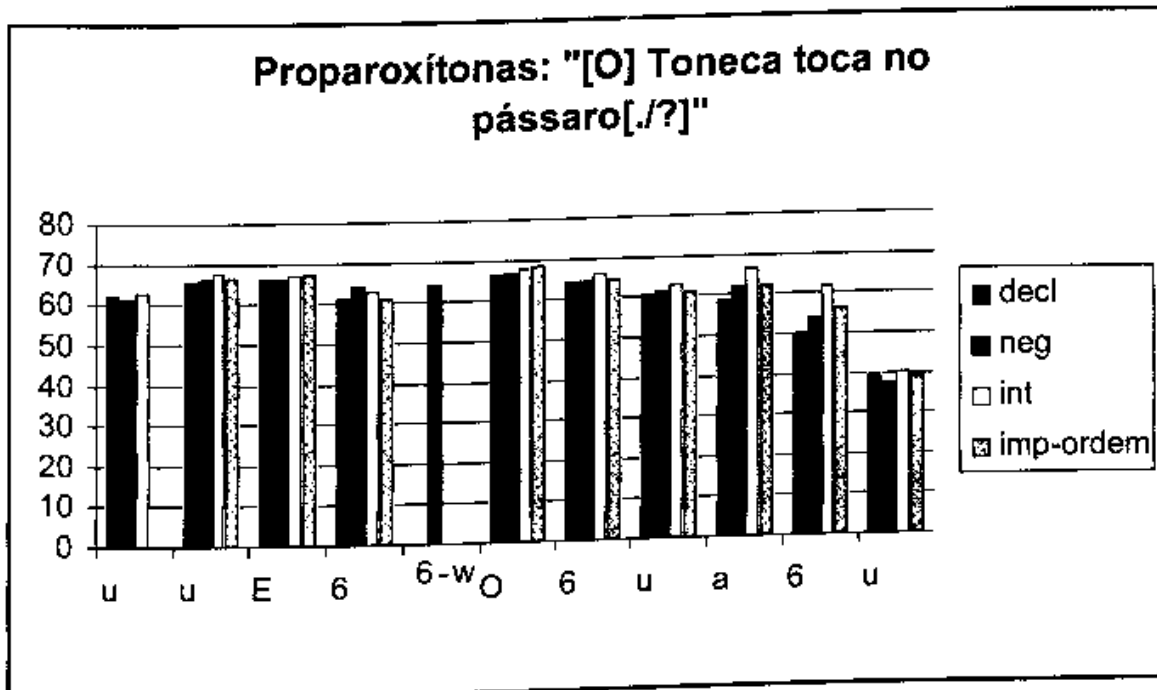


Figura 8 – Exemplo da análise comparativa da energia das vogais em quatro tipos de frase

No exemplo acima apresentado, tratamos frases do tipo declarativo afirmativo (decl), negativo (neg), tipo interrogativo global (int) e frases do tipo imperativa de ordem (ord).

Através deste gráfico podemos verificar que a frase declarativa e negativa, aparentemente, têm uma média bastante igual. Não podemos esquecer que ambas são declarativas. Tanto a frase interrogativa como a frase imperativa-ordem têm geralmente mais intensidade do que a declarativa e a negativa. Apenas uma única vez a frase imperativa tem menos intensidade do que a negativa, provavelmente porque a intensidade da vogal desta aumenta antes do advérbio de negação e foco da frase. Como dissemos as vogais mais intensas aparecem como pertencendo tanto à imperativa como à interrogativa. No entanto, notamos que nas últimas quatro vogais a maior intensidade é sempre da interrogativa.

#### 5.3.2 Efeito da posição da vogal na frase nos valores de energia.

Para verificar a influência da posição da vogal dentro da frase nos valores da energia das vogais, analisámos o efeito deste factor no referido parâmetro, utilizando para isso o programa estatístico SPSS. Este programa havia já sido utilizado, quando procedemos à análise da influência do acento frásico e lexical na duração das vogais.

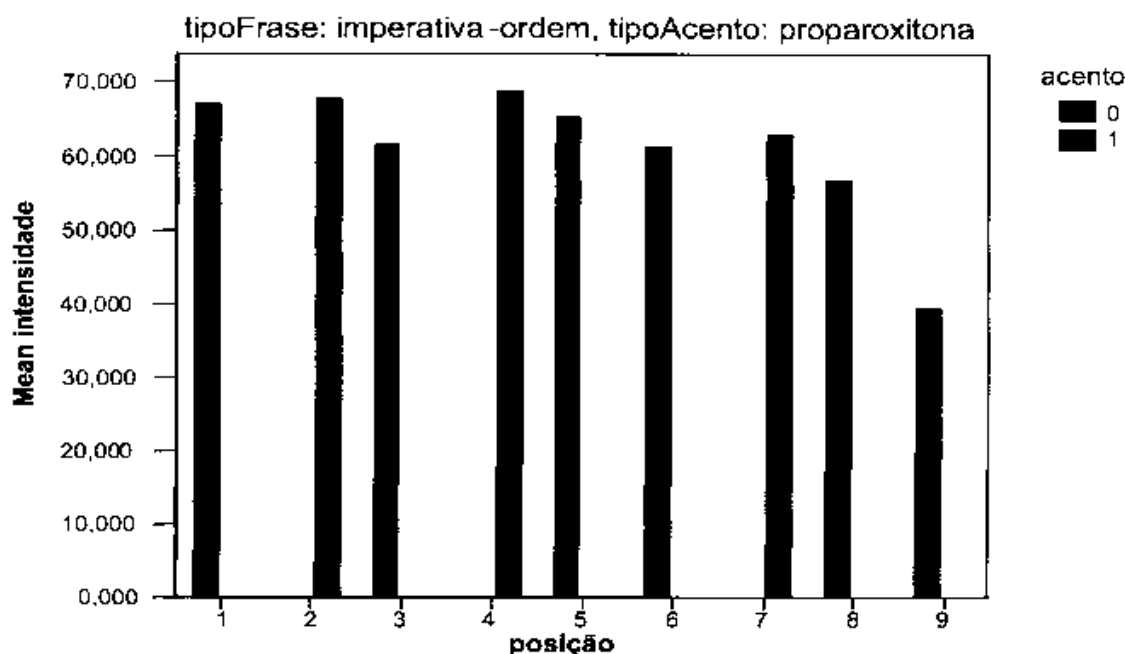


Figura 9 - Exemplo da análise comparativa a frases do tipo imperativa ordem

Pela análise da Figura 9, é possível verificar pequenas oscilações de intensidade ao longo da frase. As vogais tónicas têm valores mais altos do que as vogais átonas que a ladeiam. Notamos, porém, que a partir da posição 4, que corresponde ao predicado, há um decréscimo de valores que voltam a subir na última vogal tónica, detentora do acento de frase, descendo nas seguintes, até ao final da frase.

#### 5.4 INTSINT

Para além da análise que temos vindo a descrever, foi também efectuada uma análise INTSINT, não só com o objectivo de complementar os gráficos de Excel e SPSS, mas também de testar a sua aplicabilidade à Língua Portuguesa. O resultado foi, sem dúvida, animador, pois esta análise automática foi capaz de extrair os movimentos essenciais da trajectória da frequência fundamental para os diversos tipos de frase.

No âmbito deste artigo, e por razões facilmente entendíveis, não nos é possível apresentar exemplos para todos os tipos de frase. Escolhemos, por isso, apenas para exemplificar, uma das frases declarativas afirmativas à qual o modelo de análise foi aplicado. Na figura 10, o resultado da anotação automática INTSINT encontra-se na última linha (AN.19).



Relativamente à intensidade:

- A intensidade revela-se um factor pouco importante, sendo apenas usada para marcar o acento lexical;
- Só associada ao movimento de F0, pode ter algum valor distintivo.

A experiência de utilização do INTSINT, pelos resultados obtidos, mostra que esta análise automática foi capaz de extrair os movimentos essenciais da trajectória de F0 para os diversos tipos de frase, sendo, por isso, uma metodologia a considerar em trabalhos futuros.

### Referências Bibliográficas

- CRUZ-Ferreira, M. (1998) Intonation in European Portuguese, in: HIRST, D.; DI CRISTO (Coord.): *Intonation Systems: a survey of twenty languages*, Cambridge University Press, pp. 167-178.
- CUNHA, C.; Cintra, L. (1994) *Nova Gramática do Português Contemporâneo*, Edições João Sá da Costa, Lisboa, 1994.
- FALÉ, Isabel (1998) Duração das vogais tónicas e fronteiras prosódicas: uma análise em estruturas coordenadas, *Actas do XIII Encontro da Associação Portuguesa de Linguística*, Vol. 1, Edições Colibri, Lisboa, pp. 255-269.
- GOSCINNY, R.; UDERZO, A.: *Astérix legionário*, Meribérica/Liber Editora, Lda., Lisboa, 2002, pp. 21.
- HIRST, Daniel (1998) Automatic analysis of Prosody for multi-lingual speech corpora, Paper presented at COST 258 Workshop: Naturalness of Synthetic Speech, Porto.
- MARTINS, Maria Raquel Delgado (1993) Estruturação temporal de Fala: análise acústica e reconhecimento perceptivo, *Actas do Encontro de Processamento da Língua Portuguesa falada e escrita*, Lisboa.