

Divisão Silábica Automática do Texto Escrito e Falado

Paulo Duarte Ferreira Gouveia
Dep. Informática, ESTiG,
Instituto Politécnico de Bragança.
Campus de Santa Apolónia,
5301-857 Bragança, Portugal.
pgouveia@ipb.pt

João Paulo Ramos Teixeira
Dep. Electrotecnia, ESTiG,
Instituto Politécnico de Bragança.
Campus de Santa Apolónia,
5301-857 Bragança, Portugal.
joaopt@ipb.pt

Diamantino Rui da Silva Freitas
CEFAT, Faculdade de Engenharia da Universidade do Porto.
Rua dos Bragas, 4050-123 Porto, Portugal
dfreitas@fe.up.pt

Abstract

This article presents an algorithm that allows to carry through the syllabic splitting automatically as a stage of the development of a more extensive work, that is the study of prosodic models for the European Portuguese, fit in the development of a synthesizer of speech. The algorithm of syllabic splitting is conceived for application in two distinct situations: in the first one it is applied to the text written and in second to the sequence of phonemes really produced in the locution of this exactly text. Each one of the applications has its peculiarities and difficulties, that are described, as well as the solutions adopted for its resolution. In the first case we obtains a tax of error of 0.06% and in the second case a tax of error of 0.89%. The algorithm is based on the consideration of syllables of types V, VC, VCC, CV, CVC, CCV and CCVC, being V a vowel or diphthong and C a consonant. We admit that this types of syllables covers all the existing syllables in Portuguese.

Resumo

Este artigo apresenta um algoritmo que permite realizar automaticamente a separação silábica como uma etapa do desenvolvimento de um trabalho mais extenso, que é o estudo de modelos prosódicos para o português europeu, enquadrado no desenvolvimento de um sintetizador de fala. O algoritmo de separação silábica foi concebido para aplicação em duas situações distintas: na primeira é aplicado ao texto escrito e na segunda à sequência de fonemas realmente produzidos na locução desse mesmo texto. Cada uma das aplicações está envolta nas suas peculiaridades e dificuldades, que são descritas, bem como as soluções adoptadas para a sua resolução. No primeiro caso consegue-se uma taxa de erro de 0,06% e no segundo caso uma taxa de erro de 0,89%. O algoritmo baseia-se na consideração de sílabas dos tipos V, VC, VCC, CV, CVC, CCV e CCVC, sendo V uma vogal ou ditongo e C uma consoante, que se admite cobrirem todas as sílabas existentes em Português.

1. Introdução

Em estudos realizados para outras línguas, como o espanhol, em Córdoba et al (1999), o inglês, francês, alemão, italiano, espanhol e holandês, em Fackerell et al (1999), constata-se que a sílaba é uma unidade relevante na determinação de parâmetros prosódicos tais como, as durações dos fonemas e as micro-variações da frequência fundamental na realização de síntese de fala a partir do texto. Estando os autores a realizar estudos prosódicos para o português europeu, com a

finalidade de implementar os modelos resultantes num sintetizador de fala, Teixeira et al (1998), torna-se aqui necessário um processo que automaticamente separe as sílabas, quer a partir de um texto escrito, quer a partir de uma sequência de fonemas identificados no sinal acústico resultante da gravação de textos lidos por um locutor.

Em alguns estudos realizados para o português, Catarino (2000), e para o espanhol, Benenati (2000), constatámos que, apesar de serem línguas diferentes, algumas regras de separação silábica são comuns. Verificámos também que as regras encontradas nas duas referências não constituem um conjunto consistente de regras passíveis de serem usadas num algoritmo que realize a divisão silábica automaticamente. Além disso, não são suficientes para resolver todos os casos. Nestas referências ocorrem também regras contraditórias como sejam, a separação dos dígrafos *rr* e *ss* para o caso da referência Catarino (2000) e não separação dos mesmos dígrafos na referência Benenati (2000). Por isso, a solução encontrada no presente artigo governa-se por um conjunto de regras distinto das referidas.

Apresentam-se de seguida as regras para a separação silábica, apontadas na referência Catarino (2000):

- Não se separam os ditongos e tritongos;
- Separam-se as vogais dos hiatos;
- Não se separam os dígrafos *ch*, *lh*, *nh*, *qu*, *gu*;
- Separam-se os dígrafos *rr*, *ss*, *sc*, *sç*, *xc*;
- Separam-se os encontros consonantais impuros;
- Separam-se as vogais idênticas e os grupos consonantais *cc* e *cç*;
- Nos prefixos terminados por consoante, esta liga-se à sílaba anterior se a palavra começa com consoante, e liga-se à sílaba seguinte se a palavra começa com vogal.

As regras apresentadas na referência Benenati (2000) são as seguintes:

- Cada sílaba deve terminar em vogal sempre que possível.
- Duas vogais devem ser separadas a menos que uma delas seja semivogal.
- Em geral duas consoantes devem ser divididas. Os dígrafos *ch*, *ll* e *rr* são considerados como uma única consoante, não devendo ser divididos. Duplo *c* e duplo *n* devem ser divididos;
- As consoantes *r* e *l* não devem ser separadas da consoante precedente excepto se for um *s*;
- Na combinação de 3 ou 4 consoantes, a letra *s* deve pertencer à sílaba precedente.

Estes conjuntos de regras têm como perspectiva uma separação silábica do texto. Contudo a separação silábica que pretendemos realizar deve separar "sílabas fonéticas". Ou seja, pretende-se separar a palavra em sílabas da forma como se soletra, não fazendo assim sentido, por exemplo, dividir o dígrafo *rr* em sílabas diferentes porque juntos realizam apenas um fonema.

Pondo de parte as contradições encontradas atrás, conclui-se que, a questão da divisão silábica é mais ou menos objectiva, excepto nas situações em que duas vogais podem formar um hiato (duas vogais pertencendo a sílabas diferentes) ou um ditongo, e em alguns casos de encontros consonantais. Na possibilidade de duas vogais poderem constituir um ditongo crescente, não há problema em os considerar sempre como um hiato porque, sendo os ditongos crescentes instáveis, Cunha e Cintra (1997) e Bergström e Reis (1997), em pronúncias distintas podem resultar como hiato ou como ditongo. Quando a sequência de vogais indiciam um ditongo decrescente, normalmente é assim. Contudo, como veremos no relato dos resultados na secção 4, os erros encontrados devem-se exclusivamente a situações, bastante raras, de duas vogais serem erradamente interpretadas como um ditongo decrescente. No que concerne aos encontros consonantais em posição medial é, em algumas situações (*bc*, *bd*, *bj*, *bs*, *bt*, *cm*, *cn*, *ct*, *dj*, *dm*, *dq*, *cç*, *ds*, *dv*, *fn*, *ft*, *gd*, *gm*, *gn*, *mn*, *pç*, *pn*, *pt*, *tm* e *tn*), uma questão ambígua dos pontos de vista da linguística descritiva e da psicolinguística. Segundo os dados da linguística descritiva, dividem-se estes encontros, resultando uma divisão do tipo *rit-mo*, enquanto que do ponto de vista da psicolinguística, não se dividem, sendo a divisão do tipo *ri-tmo*. Segundo Cunha e Cintra (1997), ambas as divisões são possíveis numa pronúncia tensa. Ainda segundo o mesmo autor, caso este tipo de encontros consonantais ocorra no início de palavra, então são inseparáveis (ex: *psi-có-lo-go*). Ambas as situações foram implementadas em diferentes versões do algoritmo, contudo os

resultados finais, adiante apresentados, reportam-se à primeira destas duas possibilidades consideradas.

A divisão silábica é aplicada, como referido atrás, a duas situações um pouco distintas. Na primeira situação, pretendemos aplicar o algoritmo da divisão silábica ao texto escrito. Na segunda situação, para efeito de estudo prosódico, foi gravada, em sala insonorizada com boas condições acústicas, a fala resultante da leitura com pronúncia tensa dos textos escritos. Posteriormente, foi feita a identificação e marcação (etiquetagem) da sequência de fonemas, em código SAMPA, Wells (2000), usando para o efeito software apropriado e etiquetadores treinados. O resultado deste processo de etiquetagem é uma sequência de fonemas do género apresentado:

u~ !"por!t vi~!"taZ 6!ke"sew !"ti!tlu X 6"Raj6 "miu!d6 !"krES@ X 6 "OLZ "viS!tS X

Este Pequeno excerto da sequência de fonemas¹ resultantes da etiquetagem, corresponde ao texto "Um porto vintage aqueceu o título. Arraia miúda cresce a olhos vistos". Além dos códigos SAMPA usados, também aparecem os símbolos: X - marca uma pausa; " - marca a posição da sílaba tónica; ! - marca a oclusão das consoantes oclusivas (quando existe). Sempre que ocorre uma semivogal na sequência de fonemas, considera-se a existência de um ditongo, juntando-se a referida semivogal com a vogal vizinha. No caso da semivogal ser vizinha de duas vogais (antes e depois), considera-se que é a vogal anterior que forma com ela um ditongo, privilegiando-se assim, os ditongos decrescentes em detrimento dos crescentes, pelas razões atrás apontadas. Chama-se a atenção para o facto desta sequência de fonemas não ser a mesma coisa que a transcrição fonética. No caso da transcrição fonética existe sempre uma vogal em cada sílaba, mas neste caso, o orador, na leitura dos textos, omite muitas vogais átonas, resultando assim, sequências de várias consoantes pertencentes a sílabas diferentes que originalmente continham uma vogal.

Ainda na fase da etiquetagem, o etiquetador marca a sílaba tónica com um código apropriado, imediatamente antes do primeiro fonema da sílaba. Também nesta fase são colocadas marcas de separação de palavra. Contudo, surgem, nesta matéria, situações complicadas de resolver quando acontece a contracção da última vogal de uma palavra com a primeira vogal da palavra seguinte, resultando uma única vogal que pertencerá apenas a uma das palavras pela referida marcação.

É a esta sequência de fonemas resultante da marcação, que é aplicado o algoritmo aqui desenvolvido. Nesta situação surgem casos distintos da primeira, devidos precisamente ao facto da sequência de fonemas falados não ser coincidente com a transcrição fonética do texto. Como já referido, os casos em que as vogais são simplesmente suprimidas durante a fala, surgem situações de difícil decisão, pois juntam-se várias consoantes que pertencem a sílabas diferentes, (ex: *futebol* -> *fibOl*). É muito frequente o fonema @ ser omitido em posição final (ex: *sete* -> *sEt*). Acontece também com frequência que a forma como são pronunciadas determinadas palavras conduz a uma divisão silábica diferente da do próprio texto por este ser lido com sílabas "legais", sendo por conseguinte difícil de conhecer automaticamente a palavra original e a respectiva divisão correcta (ex: *para* pronunciada *pr6* é considerada como uma só sílaba).

A marca do primeiro fonema da sílaba tónica é usada como ajuda na separação silábica, já que indica, sem margem para dúvidas, o início de uma sílaba.

O artigo está organizado da seguinte forma: na secção 2 descrevem-se as regras consideradas para a divisão silábica do texto escrito e a respectiva implementação. Na secção 3 descreve-se as regras consideradas para a divisão silábica do texto falado e a respectiva implementação. Na secção 4 apresentam-se os resultados obtidos para os dois métodos e uma análise das situações de erro. Finalmente, na secção 5 elaboram-se as principais conclusões.

¹ Na verdade são usados fones e não fonemas, já que se faz a separação entre a oclusão e a explosão das consoantes oclusivas. Contudo, para efeitos da divisão silábica este aspecto é irrelevante já que os símbolos que marcam a oclusão são considerados como pertencendo à consoante seguinte.

2. Divisão Silábica do Texto Escrito

Na secção 2.1, começamos por referir as regras que consideramos genericamente aplicáveis à divisão silábica do texto, ressaltando algum pré-processamento para identificação dos ditongos e sua substituição por um determinado código, e para conversão dos dígrafos que correspondam a vogais nasais ou consoantes fonéticas (ex: *am, rr, ss, ch, lh, nh*), por códigos que os identifiquem. Na secção 2.2 apresentamos o algoritmo que implementa as referidas regras.

2.1. Regras para a Divisão Silábica

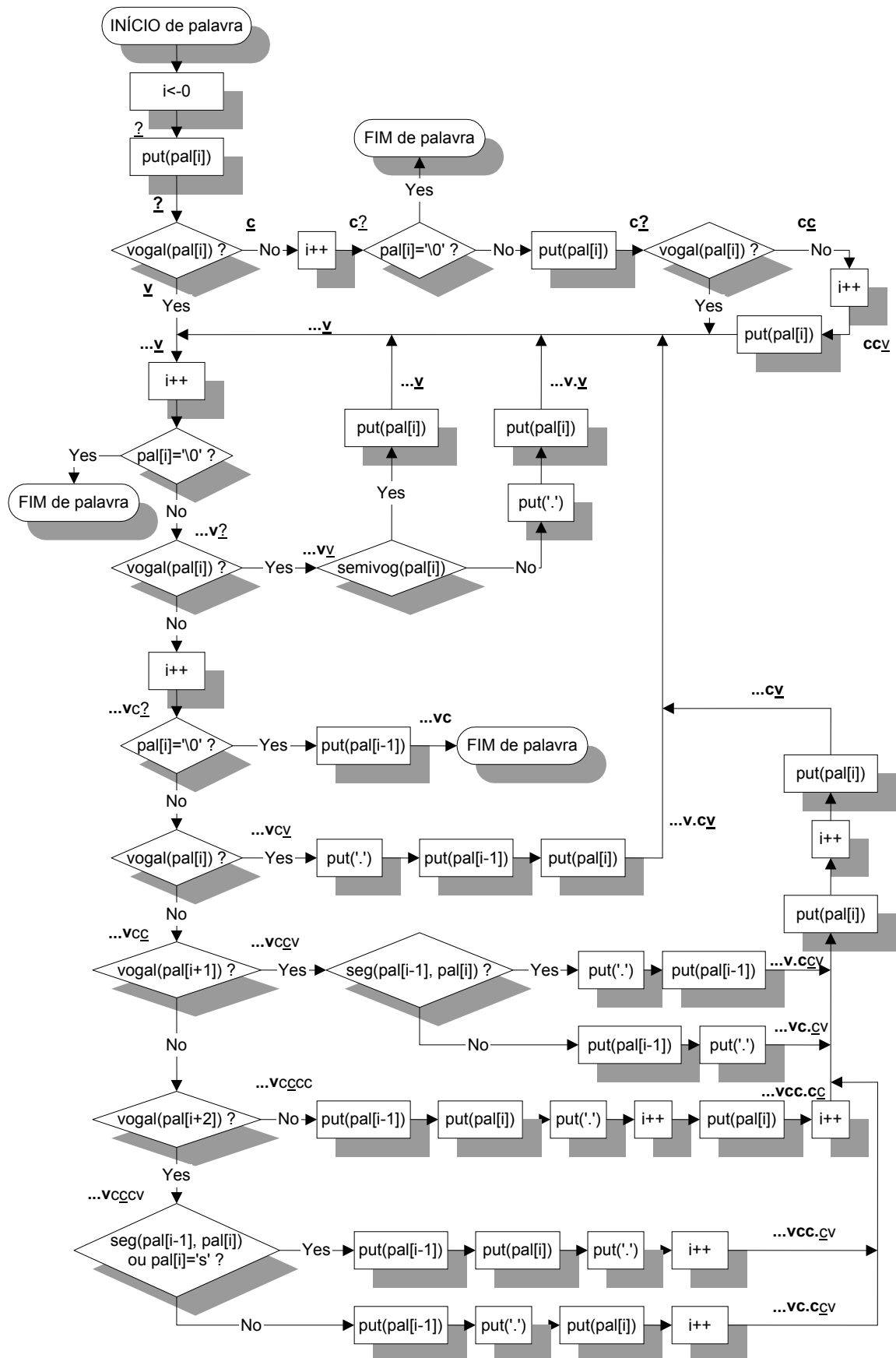
As regras consideradas para efeito da divisão silábica baseiam-se no pressuposto de que as sílabas só podem ser dos seguintes tipos: **V**, **VC**, **VCC**, **CV**, **CVC**, **CCV** e **CCVC**, em que **C** representa uma consoante fonética e **V** uma vogal ou ditongo. Admitindo apenas a existência deste tipo de sílabas, o processo de separação silábica fica, em grande parte, definido. Apenas nos casos em que estas regras não são suficientes, é que se recorre a um conjunto de regras complementares. A insuficiência destas regras constata-se apenas em dois tipos de casos: no primeiro quando ocorrem duas consoantes entre duas vogais (...**VCCV**...) e no segundo quando ocorrem três consoantes entre duas vogais (...**VCCCV**...). O primeiro caso é resolvido a partir da regra de que a divisão silábica nunca pode ocorrer entre uma consoante e uma vogal (**C-V**); então, as duas consoantes começam uma nova sílaba se constituírem um par de consoantes inseparáveis, isto é, a primeira delas pertencer ao conjunto {*b, p, d, t, g, k, v e f*} (*k* corresponde a uma das letras *k, c* ou *q*), e a segunda pertencer ao conjunto {*l, e r*} (ex: *a-tlas*); caso contrário, a divisão silábica ocorrerá necessariamente entre as duas consoantes (ex: *al-tas*). O segundo caso, (...**VCCCV**...), é resolvido da seguinte forma: como três consoantes seguidas nunca podem pertencer a uma mesma sílaba, a divisão ocorrerá entre a segunda e terceira consoantes se as primeiras duas forem inseparáveis ou então se a segunda for a letra *s*, uma vez que quando a letra *s* precede outra consoante as duas pertencerão a sílabas diferentes (ex: *obs-tar*); caso contrário, a divisão silábica ocorrerá necessariamente entre as primeiras duas consoantes (ex: *ul-tra*).

Quando no texto, surgem duas ou mais vogais seguidas, é necessário verificar se formam um ditongo decrescente ou um hiato². Para a detecção dos ditongos decrescentes, procuram-se ocorrências de uma vogal seguida de uma semivogal. Considera-se uma semivogal fonética se se tratar de um *i* ou *u* e não lhe suceder um *r* ou um *l* como última letra da palavra ou como primeira de duas ou mais consoantes (ex: *cai, cai-ro, ca-ir* e *ca-irmos*). Não é considerada semivogal quando precedida por uma vogal igual (ex: *ni-ilismo*), ou então, se lhe suceder a vogal *u* (ex: *ca-íu*) ou uma consoante indicadora de nasalidade (ex: *a-in-da*). Para além dos casos referidos, considera-se ainda como semivogal a letra *o* precedida pela letra *a* (ex: *ao*).

2.2. Algoritmo

A implementação foi realizada em linguagem C, apresentando-se adiante um fluxograma ilustrativo da parte responsável pela divisão silábica de uma palavra. A palavra a processar encontra-se na *string* designada por *pal*, sendo representada como em linguagem C: o carácter '\0' é usado como finalizador de palavra, e ao primeiro carácter da *string* está associado o índice 0. Para além da variável *pal*, faz-se uso da variável *i*: índice que referencia a letra da palavra a ser processada. Do diagrama fazem ainda parte algumas invocações a funções: *vogal(x)* e *semivog(x)*, cujo objectivo é verificar se o carácter *x* é uma vogal fonética e semivogal, respectivamente; *put(x)* que tem como função enviar para a *string* de saída o carácter *x*; e a função *seg(x,y)* que permite verificar se *x* e *y* formam um par de consoantes inseparáveis (que se devem manter na mesma sílaba). Por fim, para uma melhor compreensão do diagrama, refira-se que é usado o carácter '.' como separador de sílabas (exemplo: tendo como *string* de entrada a palavra "fluxograma", obtém-se como resultado a *string* "flu.xo.gra.ma").

² Os ditongos crescentes podem ser interpretados como hiatos, como já se referiu na introdução.



LEGENDA:

v: vogal
c: consoante

... : qualquer sequência de letras
.: separador de sílabas
?: letra ainda não determinada

negrito: letra já gravada na string de saída
sublinhado: letra referenciada pelo índice *i*

3. Divisão Silábica do Texto Falado

Na secção 3.1 apresentamos as regras especificamente usadas para a divisão silábica do texto falado. São justificadas as razões de algumas modificações que foi necessário introduzir no conjunto de tipos de sílabas considerados neste caso e a forma como são tratados os ditongos.

Na secção 3.2 apresentamos o algoritmo que implementa as referidas regras.

3.1. Regras para a Divisão Silábica

A questão da distinção entre ditongos e hiatos aparece simplificada já que neste caso as semivogais que dão origem aos ditongos estão identificadas.

O maior problema ocorre devido à supressão de muitas vogais, dando origem a sequências de consoantes pertencentes a sílabas diferentes, assim complicando, de sobremaneira, a correcta identificação de sílabas de acordo com o texto lido. Este fenómeno leva a considerar mais dois tipos abstractos de sílabas, além das referidas em 2.1, que são **C** e **CC**. A ocorrência destas sílabas deve-se precisamente à supressão das vogais em sílabas dos tipos **CV**, **CVC** e **CCV**. Para o caso da sílaba tipo **CCVC** é muito raro ocorrer a supressão da vogal (ex: *p@nEtr6S*), pelo que se optou por não se considerarem legais sílabas com três consoantes juntas, de forma a evitar alguns erros que daí resultariam.

A divisão silábica do texto falado baseia-se também no pressuposto de considerar apenas sílabas dos tipos: **V**, **VC**, **VCC**, **CV**, **CVC**, **CCV**, **CCVC**, **C** e **CC**. No entanto, devido à dificuldade adicional introduzida pela supressão de algumas vogais, foi necessário elaborar um conjunto adicional de regras a aplicar em situações específicas.

- As consoantes {**l**, **r**, **S**, **z** e **Z**}, quando seguidas de outra consoante, ligam-se sempre à sílaba anterior (ex: *sal-tu*). Neste grupo inclui-se a consoante **Z** devido à vocalização da consoante não vocalizada **S** em contexto vocalizado (ex: *meZ mu*).
- As consoantes {**S**, **z** e **Z**}, em final de palavra, associam-se à sílaba anterior³.
- Quando ocorre uma sequência de uma vogal seguida por duas consoantes, se as duas consoantes constituírem um dos casos seguintes {*bk, bd, bZ, bs, bt, km, kn, kt, ks, dZ, dm, dk, ds, dv, fn, ft, gd, gm, gn, mn, ps, pn, pt, tm* e *tn*}, então separam-se as consoantes, criando uma sílaba do tipo **VC** (ex: *ap-tu*). Esta mesma sequência em início de palavra nunca se separa (ex: *pnew*), Cunha e Cintra (1997).

3.2. Algoritmo

Antes de aplicação do algoritmo para cada palavra, são retirados os símbolos de oclusão das consoantes oclusivas (!) e as semivogais, ficando a referência das suas posições. Depois de aplicado o algoritmo, os referidos símbolos e semivogais são re-introduzidos nas posições originais da palavra e ligados ao fonema associado – no caso do símbolo de oclusão, associado à consoante oclusiva, e no caso das semivogais, associadas à vogal que com ele formam um ditongo de acordo com as regras atrás descritas.

O algoritmo apresentado é responsável por identificar onde termina uma sílaba, após o qual, volta ao início para repetir o processo até terminar a palavra.

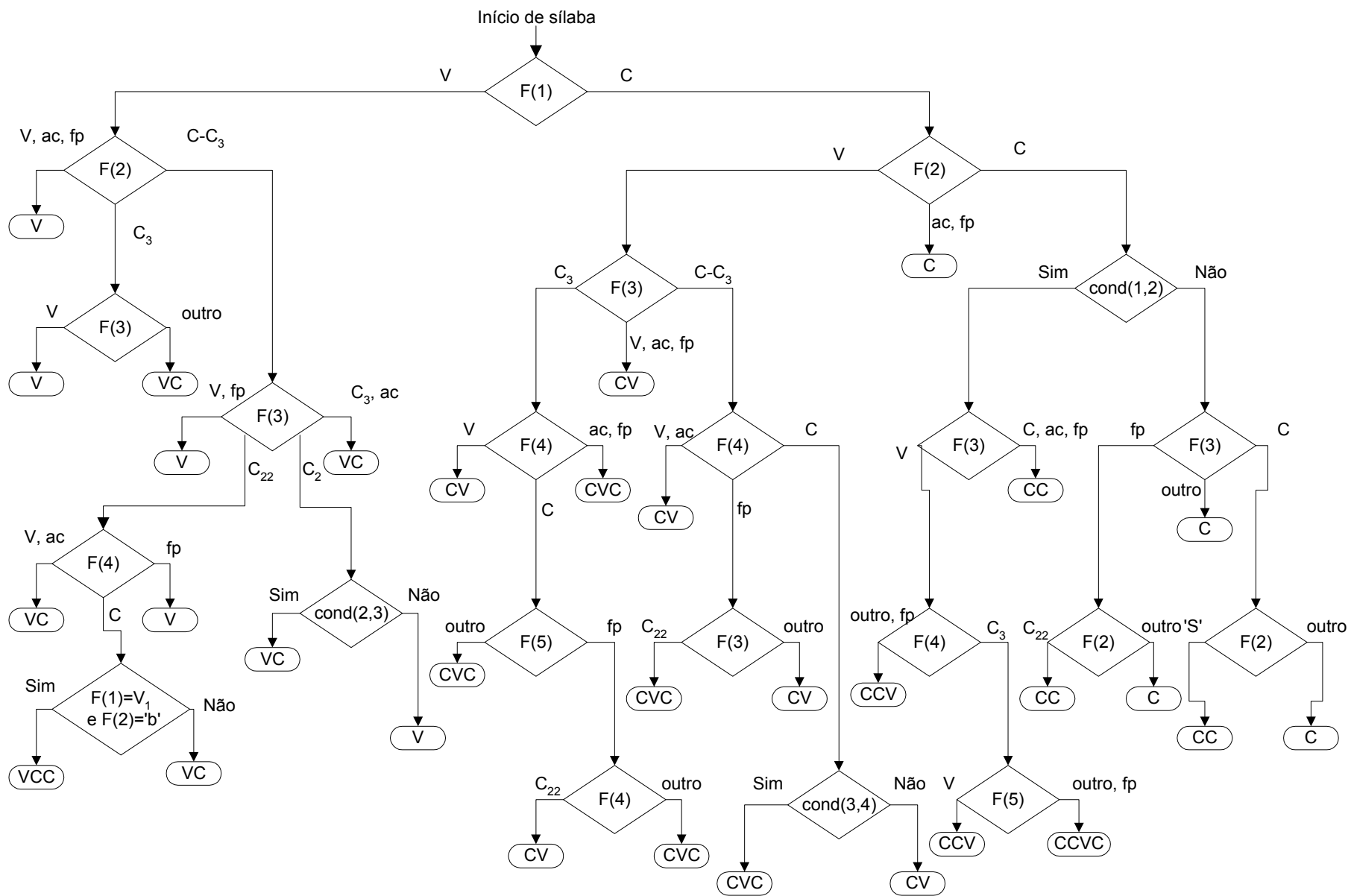
Um ciclo de execução do algoritmo termina numa decisão do tipo de sílaba detectada, **V**, **CV**, **CVC**, etc.

A função F(x), lê o fonema com índice x após o início do algoritmo. Este fonema pode ser: **V** – vogal; **C** – consoante; **V₁** – uma das Vogais [*a*, **6**, **O**, *o*]; **C₁** – uma das consoantes [*b*, *p*, *d*, *t*, *g*, *k*, *v*, *f*]; **C₂** – uma das Consoantes [*l*, *r*]; **C₂₂** – uma das Consoantes [**S**, **z**, **Z**]; **C₃** – uma das Consoantes [*l*, *r*, **S**, **z**, **Z**]; **C-C₃** – uma Consoante que não pertença a **C₃**; **ac** – o fonema lido contém a marca de acento da sílaba tónica; **fp** – fim de palavra (a palavra termina no fonema x-1).

³ Esta regra leva à ocorrência de erros em palavras cuja sílaba final, da transcrição fonética original, seja uma destas consoantes seguida de uma vogal, mas em que no texto falado foi suprimida essa vogal (ex: palavra original – "reage" transcrição fonética – "*RiaZ@*" texto falado – "*Ri aZ*").

A função $\text{cond}(a,b)$, retorna o valor lógico 1 (sim) se $[(F(a)=C1 \text{ e } F(b)=C2)$ ou (sílabas em início de palavra) e $(F(a)F(b))$ são uma das sequências: ***bk, bd, bZ, bs, bt, km, kn, kt, ks, dZ, dm, dk, , ds, dv, fn, ft, gd, gm, gn, mn, ps, pn, pt, tm e tn***], senão retorna o valor lógico 0 (não).

Como exemplo apresenta-se a decisão tomada pelo algoritmo para dividir a palavra ***kaz6***: F(1) é ***k***, como é uma consoante segue no ramo da direita e lê F(2) que é o ***a***, como é vogal vem para o ramo da esquerda e lê F(3) que é ***z***, como se trata de uma consoante C_3 , vem para o ramo da esquerda e lê F(4) que é ***6***, tratando-se de uma vogal a decisão foi tomada considerando uma sílaba do tipo ***CV*** separando então a primeira sílaba ***ka***. Num novo ciclo do algoritmo lê F(1) que agora é o ***z***, como é uma consoante, segue pelo ramo da direita e lê F(2) que é ***6***, sendo uma vogal segue pelo ramo da esquerda e lê F(3), como F(3) não existe, trata-se do fim de palavra (fp), pelo que a sílaba é do tipo ***CV***. Assim, terminou a palavra, resultando a divisão ***ka-z6***.



4. Testes e Análise dos Resultados

A aplicação dos algoritmos desenvolvidos e respectivas regras conduz a resultados diferentes para os dois casos referidos (aplicação ao texto e aplicação à sequência de fonemas), como seria de esperar, devido à omissão de vogais no segundo caso. Contudo, para os dois casos a taxa de erro na divisão silábica é muito baixa.

A taxa de erro foi calculada com o número de erros cometidos na divisão por número de sílabas divididas, em percentagem.

A metodologia seguida baseou-se no princípio da independência entre os conjuntos de desenvolvimento e de teste. Com efeito, para o aperfeiçoamento e verificação das regras, utilizou-se um primeiro conjunto de textos com separação silábica manual. Subsequentemente, realizou-se o teste dos algoritmos com novos textos extraídos de publicações periódicas.

No caso da divisão silábica do texto, o algoritmo foi testado com um conjunto de palavras não repetidas, com duas ou mais letras, retiradas de cinco artigos de um periódico diário. Por comparação com o resultado duma divisão manual previamente efectuada sobre o mesmo texto constatou-se que o algoritmo apenas errou em duas palavras num total de 1164. Assim, em 3387 sílabas divididas ocorreram apenas 2 erros, correspondendo a uma taxa de 0.06% de erros por sílaba. Embora tenham ocorrido em palavras distintas ("cai-re-mos" e "reu-ni-ão"), representam um mesmo tipo de erro: hiato interpretado como ditongo decrescente. É a situação em que aparece uma vogal seguida de um *i* ou de um *u* que, excepcionalmente, não se comportam como semivogais (a regra usada trata a generalidade desses casos como ditongos decrescentes). Os erros não são de solução imediata uma vez que se encontram outras palavras do género que se comportam de maneira diferente na divisão silábica (exemplo: Cai-ro e reu-má-ti-co). Esperamos eliminá-los num futuro próximo por introdução de considerações sobre o contexto silábico.

No caso da divisão silábica da sequência de fonemas, o algoritmo foi aplicado à sequência de fonemas etiquetados, de um sinal de fala de cerca de cinco minutos de duração, correspondente à leitura de dois artigos, igualmente de um periódico diário. A divisão foi aplicada apenas a palavras, não se tendo considerado como tal os monossílabos – estes são, por definição, uma só sílaba. Num total de 1569 sílabas divididas ocorreram 14 erros, correspondendo a uma taxa de 0.89% de erros por sílaba. Os erros ocorreram em sete palavras diferentes, algumas repetidas. Apresentam-se a seguir as divisões silábicas das palavras em que ocorreram os erros, com as sílabas erradamente separadas a sublinhado (as aspas são a marca que identifica o primeiro fonema da sílaba tónica): (futebol) ft-“bol, (evidentemente) iv-de~-t-“me~-t, (ministério) nniS-“tE-riw, (irresponsabilidade) iRS-po~-s6-bli-“da, (industrial) i~d-S-tri-“al, (acusação) 6k-z6-“s6~w, (demonstração) dmo~S-tr6-“s6~w. Todas as situações de erro resultam em sílabas em que foi omitida a vogal, tendo a consoante dessas sílabas ficado associadas às sílabas vizinhas. Não foram considerados erros nas situações em que, por omissão de uma vogal, na leitura do texto, resulta uma "sílabas fonética legal" (ex: palavra *para* lida *pr6*).

5. Conclusões

Os resultados obtidos vêm de encontro às expectativas criadas no início do trabalho. Infelizmente, não podemos comparar os resultados com outras implementações já que não encontramos publicados outros trabalhos que automaticamente realizem a divisão silábica.

Para as aplicações que se pretendem dar aos algoritmos, estudo prosódico num caso, e parte do bloco prosódico de um sintetizador no outro, pensamos que os objectivos foram suficientemente atingidos.

Agradecimentos

O presente trabalho foi realizado no âmbito de um projecto com a finalidade de desenvolver um Conversor-Texto-Fala de elevada naturalidade, para o qual têm contribuído entidades e projectos a que os autores gostariam de deixar um agradecimento, nomeadamente à Acção COST 258 “NATURALNESS OF SYNTHETIC SPEECH” e ao projecto Antígona. Agradecemos ainda à ESTiG de Bragança as facilidades concedidas para a prossecução dos trabalhos nesta área.

Finalmente, uma palavra de apreço e agradecimento ao Dr. João Veloso da FLUP, pela sua empenhada contribuição em termos de esclarecimento de algumas dúvidas na área de linguística.

Referências

- Benenati C. (2000). *Separacion en Silabas*, <http://www.lclark.edu/~benenati/silabacento/silabas.html>.
- Bergström M. and Reis N. (1997). *Prontuário Ortográfico e Guia da Língua Portuguesa*, Editorial Notícias.
- Catarino D. (2000). *Separação Silábica*, <http://www.option-line.com/members/dilson/Silabas.htm>.
- Córdoba R., Vallejo J., Montero J., Arriola J., López M. and Pardo J. M. (1999). *Automatic Modeling of Duration in a Spanish Text-To-Speech System Using Neural Networks*, in EUROSPEECH'99 vol. 4.
- Cunha C. and Cintra L. (1997). *Nova Gramática do Português Contemporâneo*, Edições João Sá da Costa.
- Fackerell J., Vereecken H., Martens J. P. and Coile B. V. (1999). *Multilingual Prosody Modelling Using Cascades of Regression Trees and Neural Networks*, in EUROSPEECH'99 vol. 4.
- Teixeira J., Freitas D., Gouveia P., Olaszy G. and Németh G. (1998). *Multivox – Conversor Texto Fala Para Português*, in PROPOR'98, Porto Alegre, Brasil.
- Wells J. (2000). *SAMPA computer readable phonetic alphabet*, <http://www.phon.ucl.ac.uk/home/sampa/home.htm>.