

Grapheme-Phone Transcription Algorithm for a Brazilian Portuguese TTS

Filipe Barbosa¹, Guilherme Pinto¹, Fernando Gil Resende¹,
Carlos Alexandre Gonçalves², Ruth Monserrat², and Maria Carlota Rosa²

¹ Escola Politécnica, Universidade Federal do Rio de Janeiro, Brazil
{filipe,guilherme,gil}@lps.ufrj.br

² Faculdade de Letras, Universidade Federal do Rio de Janeiro, Brazil
carlexandre@bol.com.br
monserruth@alternex.com.br
carlota@centroin.com.br

Abstract. This paper describes one of the aspects of an ongoing research to improve the synthetic speech quality for a Brazilian Portuguese text-to-speech synthesis system. This paper focuses on the grapheme-phone transcription algorithm. A complete set of rules for every grapheme is presented. Experimental results, obtained running the proposed algorithm through a text database, gave rise to 98,43% of accuracy rate.

1 Introduction

A text-to-speech (TTS) synthesizer is a system which should be able to read, with some degree of intelligibility and naturalness, any text in a given language, Brazilian Portuguese (BP) in the case of this paper. As simple as it seems, this definition poses a problem from the outset: which accent to choose amongst the different accents in BP. If the variety of Portuguese spoken in Rio de Janeiro was preferable to any other in Brazil in the beginning of the 20th century [1,2,3], its status has been changing drastically since the last decades of the century [4,5], and now it is considered too much nasalized, and sizzingly. So, in place of a no more prestigious accent, it was decided to provide the model with a neuter accent, understood as an accent which most of the Brazilian speakers could identify. According to Ramos [4], the accent in *Jornal Nacional*, the television news broadcast with the largest indexes of audience in Brazil, is felt in different degrees, by speakers of many parts in the country, as representative of their own dialects.

This work describes a complete transcription algorithm intended to support a BP TTS with a neuter accent. The proposed grapheme-phone transcription algorithm was tested using a randomly selected part of the CETEN-Folha text database [6]. The transcribed phones were checked, and 98,43% were correctly transcribed in a set of 7805 phones from 1364 words. In this paper, SAMPA [7] phonetic alphabet is used with a few extensions that we propose for BP. Official BP orthography is represented in italic fonts.

The remaining of this paper is organized as follows. Section 2 presents the transcription rules. Section 3 discusses experimental results, and Sect. 4 focuses on our conclusions and future work.

2 The Transcription Rules

This section presents the symbols used in the transcription algorithm, followed by the rules for every grapheme. Table 1 contains the symbolization adopted for the description of grapheme rules.

As some units needed for our BP transcription algorithms are not listed in the SAMPA phonetic units, we proposed the units listed in Table 2.

2.1 The Transcription Rules

The starting point for the transcription was the phone list proposed by Alcaim, Solewicz & Moraes [8], modified in Pinto, Barbosa & Resende Jr [9].

The rules for transcribing a given grapheme to the correspondent phone are shown in Tables 3–13. These tables are organized by grapheme blocks of rules. The left column has the grapheme sequences, the middle column contains the phone transcription, and the right column comes with one example for each rule.

Table 1. Table of symbols used to express the transcription rules

Symbol	Value	Symbol	Value
...	Any character (Punctuation or Grapheme)	C_v	Voiced consonants
<x>	A given grapheme x	C_uv	Unvoiced consonants
[y]	A given acoustic unit y	<C - x>	Except for <x>, any consonant
Hf	hifen	Pnt	Punctuation (, . ! ? ;)
Sp	Space between words	W_bgn	Beginning of word
V	Any vowel	<x><[y]>	When the transcription of the grapheme after <x> is [y]
V_s	Stressed Vowel	<x(STATE)>	STATE modifies x. Possibilities: V_us , V_s or W_bgn
V_us	Unstressed vowel	<x ₁ , x ₂ >	Graphemes x ₁ or x ₂
C	Any consonant	<x ₁ , x ₂ , (x ₃ , x ₄)>	<x ₁ x ₂ x ₃ > or <x ₁ x ₂ x ₄ >

Table 2. List of SAMPA extension units proposed for BP

[fS]	<i>afta</i>	[bZ]	<i>abnegado</i>	[i_m]	<i>mbiá</i>
[kS]	<i>krenak</i>	[gZ]	<i>magnífico</i>	[m_i]	<i>amnésia</i>
[pS]	<i>piano</i>	[vZ]	<i>Ambey</i>	[i_n]	<i>ntogapide</i>

Table 3. Table of rules for graphemes <a, b>

Grapheme Sequence for <a> algorithm	Phone	Example
... <a><Pnt>...	[6]	<i>Ela gostou do doce.</i>
... <a n><Pnt>...	[6~]	<i>Ele pertence à Febraban.</i>
... <a m><Pnt>...	[6~w]	<i>É provável que cresçam.</i>
... <a(m,n)><C-h>...	[6~]	<i>antropofagia, amperímetro</i>
... <a(V_s)><m,n>...	[6~]	<i>cama, banho</i>
... <á>...	[a]	<i>antártica</i>
... <ã,â>...	[6~]	<i>amanhã, amago</i>
... <ã o>...	[6~w]	<i>avião</i>
... <a>...	[a]	<i>aracnofobia</i>
Grapheme Sequence for algorithm	Phone	Example
... <C - (r,l)>...	[bZ]	<i>abnegado</i>
... <Pnt>...	[bZ]	<i>Ele é o sub.</i>
... ...	[b]	<i>abacate</i>

Table 4. Table of rules for graphemes <c, d>

Grapheme Sequence for <c> algorithm	Phone	Example
... <c><C - (r,l)>...	[kS]	<i>icterícia</i>
... <c><Pnt>...	[kS]	<i>Aquele é o tic-tac.</i>
... <c><e,i>...	[s]	<i>aceitar; jacinto</i>
... <ç>...	[s]	<i>almoço</i>
... <c h>...	[S]	<i>acho</i>
... <c>...	[k]	<i>claro</i>
Grapheme Sequence for <d> algorithm	Phone	Example
... <d><i,[i]>...	[dZ]	<i>dia; Ela irá à tarde.</i>
... <d><C- r,l>...	[dZ]	<i>advogado</i>
... <d><Pnt>...	[dZ]	<i>Irão usar o raid.</i>
... <d>...	[d]	<i>dote</i>

Table 5. Table of rules for graphemes <e, f>

Grapheme Sequence for <e> algorithm	Phone	Example
... <e(m,n)><C - h>...	[e~]	<i>embora; entoação</i>
... <e,e,ê><m,n>...	[e~]	<i>tema, contém, têm</i>
... <ê>...	[e]	<i>bebê</i>
... <é>...	[E]	<i>época</i>
... <e><Pnt>...	[i]	<i>O índice é alto.</i>
... <e><s><Pnt>...	[i]	<i>Essas são as frases.</i>
... <e>...	[e]	<i>errado</i>
Grapheme Sequence for <f> algorithm	Phone	Example
... <f><C - (r,l)>...	[fS]	<i>afta</i>
... <f><Pnt>...	[fS]	<i>Fez um pif-paf.</i>
... <f>...	[f]	<i>faca</i>

Table 6. Table of rules for graphemes <g, h>

Grapheme Sequence for <g> algorithm	Phone	Example
... <g><C - (r,l)>...	[gZ]	<i>magnífico</i>
... <g><Pnt>...	[gZ]	<i>Foram os caças mig.</i>
... <g><e,i>...	[Z]	<i>gelo, giz</i>
... <g u><e,i>...	[g]	<i>quindaste, guerreiro</i>
... <g>...	[g]	<i>garoto</i>
Grapheme Sequence for <h> algorithm	Phone	Example
... <h>...	∅	<i>hoje</i>

Table 7. Table of rules for graphemes <i, j, k, l>

Grapheme Sequence for <i> algorithm	Phone	Example
... <i (m,n)><C-h>...	[i~]	<i>cacimba, sinto</i>
... <i (m,n)><Pnt>...	[i~]	<i>Fomos sim; É o professor Amin.</i>
... <i><m,n>...	[i~]	<i>imaginar, ninho</i>
... <V><i(V-us)>...	[j]	<i>sai</i>
... <i>...	[i]	<i>sai</i>
Grapheme Sequence for <j> algorithm	Phone	Example
... <j>...	[Z]	<i>jambo</i>
Grapheme Sequence for <k> algorithm	Phone	Example
... <k><Pnt>...	[kS]	<i>O que é krenak?</i>
... <k>...	[k]	<i>kulina</i>
Grapheme Sequence for <l> algorithm	Phone	Example
... <l><V>...	[l]	<i>ala</i>
... <lh>...	[L]	<i>alho</i>
... <l>...	[w]	<i>vogal</i>

Table 8. Table of rules for graphemes <m, n>

Grapheme Sequence for <m> algorithm	Phone	Example
... <m><n>...	[m_i]	<i>amnésia</i>
... <m(W_bgn)>...	[i_m]	<i>mbiá</i>
... <e,é,ê,i><m><Sp><V>...	[J]	<i>Quem é?; Alguém usou; Eles têm amor; Tim está pronto.</i>
... <e,é,ê><m><Pnt>...	[j~]	<i>Que eles saquem; Está com alguém; Eles têm.</i>
... <m>...	[m]	<i>mameluco</i>
Grapheme Sequence for <n> algorithm	Phone	Example
... <n(W_bgn)><C>...	[i_n]	<i>ntogapide</i>
... <e><n><Pnt>...	[j~]	<i>Bonito abdômen.</i>
... <nh>...	[J]	<i>ganho</i>
... <n>...	[n]	<i>nata</i>

Table 9. Table of rules for graphemes <o, p>

Grapheme Sequence for <o> algorithm	Phone	Example
... <ó>...	[O]	<i>vovó</i>
... <ô>...	[o]	<i>vovô</i>
... <õ>...	[o~]	<i>coraçõ<u>es</u></i>
... <c(W_bgn)><o><Hf>...	[o]	<i>co-produçã<u>o</u></i>
... <o><Pnt>...	[u]	<i><u>O</u> músico é feliz.</i>
... <o(m,n)><C - h>...	[o~]	<i>ontem, compositor</i>
... <o (m,n)><Pnt>...	[o~]	<i>Comprar na Kibon; Est<u>á</u> no tom.</i>
... <o><m,n><V>...	[o~]	<i>soma, <u>sono</u></i>
... <o><s><Pnt>...	[u]	<i>carro<u>s</u></i>
... <o>...	[o]	<i>escopo</i>
Grapheme Sequence for <p> algorithm	Phone	Example
... <p><C - (r,l)>...	[pS]	<i>pneu</i>
... <p><i>...	[pS]	<i>piano</i>
... <p h>...	[f]	<i>esphera</i>
... <p>...	[p]	<i>pat<u>o</u></i>

Table 10. Table of rules for graphemes <q, r>

Grapheme Sequence for <q> algorithm	Phone	Example
... <q u><i,e>...	[kS]	<i>qu<u>ito</u>, qu<u>ente</u></i>
... <q>...	[k]	<i>qu<u>ando</u></i>
Grapheme Sequence for <r> algorithm	Phone	Example
... <r r>...	[R]	<i>carro<u>s</u></i>
... <r Sp r>...	[R]	<i>Pomar rodeado de fo<u>res</u>.</i>
... <r(W_bgn)>...	[R]	<i>A ru<u>a</u> foi interd<u>it</u>ada.</i>
... <r><V>...	[r]	<i>rat<u>oeira</u></i>
... <r><Sp><V,h>...	[r]	<i>Falta ac<u>ert</u>ar apenas uma; É p<u>re</u>ciso faltar<u>h</u>o<u>je</u>.</i>
... <r><Sp><C_v>...	[R]	<i>Injetar grã<u>o</u>s de arro<u>z</u>.</i>
... <r><C_v>...	[R]	<i>carg<u>a</u></i>
... <r>...	[X]	<i>Ela ir<u>á</u> se lasc<u>ar</u>.</i>

Within each grapheme block, individual rules are (a) disjunctively ordered, so that if a rule has been applied all the others are skipped; and (b) layered in the order they are checked, so that the last rule for every grapheme is applied whenever none of the other rules apply, i.e. it is the default rule.

For a given text to be transcribed, the algorithm of mapping a given grapheme into its respective acoustic unit follows the order of appearance of each grapheme. For every grapheme of the sentence, the correspondent algorithm is called, concatenating into the transcribed sentence the generated acoustic unit. It is worth noting that a rule can skip the next grapheme to be analyzed, such as in the fourth rule of grapheme <a> algorithm, where both graphemes <a><n> or <a><m> are associated with the phone [6~]. For the grapheme <h>, listed in

Table 11. Table of rules for graphemes <s, t>

Grapheme Sequence for <s> algorithm	Phone	Example
... <V><s><V>...	[z]	<i>asa</i>
... <s><C_v>...	[z]	<i>transgredir</i>
... <s s>...	[s]	<i>assar</i>
... <s c><e,i>...	[s]	<i>cre<u>s</u>cer, cre<u>s</u>cido</i>
... <s ç>...	[s]	<i>cre<u>s</u>çam</i>
... <s h>...	[S]	<i>shi<u>s</u>tu</i>
... <s Sp j>...	[Z]	<i>E<u>s</u> jogaram bola.</i>
... <s><Sp><V,h,C_v>...	[z]	<i>O<u>s</u> aros são cromados; O<u>s</u> helicópteros são aerodinâmicos; Ela<u>s</u> gostam disso.</i>
... <tr(a,â)n><s><V>...	[z]	<i>transa<u>s</u>ção, trâns<u>s</u>ito</i>
... <ob><s><(é,e)qui>...	[z]	<i>obs<u>s</u>équio, obs<u>s</u>equioso</i>
... <s>...	[s]	<i>E<u>s</u> ficaram com o prêmio .</i>
Grapheme Sequence for <t> algorithm	Phone	Example
... <t><C - (r,l)>...	[tS]	<i>algor<u>t</u>mo</i>
... <t h><Pnt>...	[tS]	<i>R<u>t</u>h</i>
... <t h>...	[t]	<i>Ar<u>t</u>hur</i>
... <t><i,[i]>...	[tS]	<i>t<u>t</u>ia, met<u>t</u>e</i>
... <t><Pnt>...	[tS]	<i>A<u>t</u>aque se<u>t</u> foi difícil.</i>
... <t>...	[t]	<i>t<u>t</u>ato</i>

Table 12. Table of rules for graphemes <u, v, w>

Grapheme Sequence for <u> algorithm	Phone	Example
... <ü>...	[w]	<i>ling<u>ü</u>ística</i>
... <u(m,n)><C - h>...	[u~]	<i>ab<u>u</u>dante, ret<u>u</u>mbante</i>
... <u (m,n)><Pnt>...	[u~]	<i>Vam<u>u</u> até Kun<u>u</u>n; Ele come <u>u</u>at<u>u</u>m</i>
... <u><m,n>...	[u~]	<i>u<u>u</u>ma, u<u>u</u>ha.</i>
... <u>...	[u]	<i>U<u>u</u>birajara</i>
Grapheme Sequence for <v> algorithm	Phone	Example
... <v><Pnt>...	[vZ]	<i>Am<u>v</u>ev</i>
... <v>...	[v]	<i>vo<u>v</u>ando</i>
Grapheme Sequence for <w> algorithm	Phone	Example
... <w>...	[w]	<i>u<u>w</u>att</i>

Table 6, no phonetic representation is used, as this grapheme is not pronounced in Portuguese.

Also, it is important to highlight that a given grapheme can be mapped into more than one acoustic unit, as can be seen in Table 13. In the fifth rule of this table, the grapheme <x> is mapped into two acoustic units, [kS][s], as well as in the sixth rule, if the word that contains <x> belongs to the exception list of Table 14.

Table 13. Table of rules for graphemes <x, y, z>

Grapheme Sequence for <x> algorithm	Phone	Example
... <(W_bgn)(e,ine)><x><V, C_v>...	[z]	<i>ex̄ecrar,</i> <i>in̄existência</i>
... <(W_bgn)(e,ê,ine)><x><C_uv>...	[s]	<i>ex̄cêntrico,</i> <i>êxtase,</i> <i>in̄expressivo</i>
...<(W_bgn)(e)><x><Hf><V, C_v>...	[z]	<i>ex̄-amigo</i>
...<(W_bgn)(e)><x><Hf><C_uv>...	[s]	<i>ex̄-presidente</i>
... <x><Pnt>...	[kS][s]	<i>Ele era o Març̄.</i>
... <V - e><x><V>...	Phone exception	Exceptions in Table 14
... <x>...	[S]	<i>faz̄ina</i>
Grapheme Sequence for <y> algorithm	Phone	Example
... <y><C>...	[i]	<i>Yguaçu</i>
... <y>...	[j]	<i>Yanomami</i>
Grapheme Sequence for <z> algorithm	Phone	Example
... <z><Sp><C_uv>...	[s]	<i>Ferraz̄ furou o ferro.</i>
... <z><Sp><V,C_v,h>...	[z]	<i>Faz̄ anos que não o vejo;</i> <i>Ferraz̄ gosta de ferro;</i> <i>Faz̄ horas que o vejo.</i>
... <z><Pnt>...	[s]	<i>Isso não se faz̄.</i>
... <z>...	[z]	<i>zebra</i>

Table 14. Table of exceptions for grapheme <x>

Phone of the rule 8 for <x>	List of words associated
[kS][s]	<i>oxítono, oxidar, complexo, reflexo, anexar, oxigênio, oxíuro, oxalato, úxer, uxoricida, axila, axiologia, íxia, taxi, sintaxe ixofagia, ixomielite, ixolite, ixômero, ixora, ixoscopia, ox-acético</i>

3 Experimental Results

The proposed rules were implemented and tested using part of the text from the CETEN-Folha database [6]. The phones originated by the algorithm were checked, and 98,43% of them were correctly transcribed. A resume of the errors can be seen in Table 15.

Table 15. Table of errors on mapping the graphemes

Error type	Occurrences	Occurrence(%)
[e] or [E] misplaced	22	0.28%
[o] or [O] misplaced	18	0.23%
Incorrect foreign word phones	31	0.40%
Diphthongs	35	0.44%
Incorrect acronym phones	17	0.22%

As can be seen from Table 15, the errors come from diphthongs, foreign words, acronyms and confusion between [O] or [o], and [E] or [e]. Rules to handle these cases are subject of ongoing research.

4 Conclusions

This paper presents rules for generating a sequence of phones from a grapheme sequence to be applied in a BP TTS system. The proposed rules were tested using part of the CETEN-Folha text database giving rise to 98,43% of correctly transcribed phones. Present research concentrates on proposing rules to deal with foreign words, names, diphthongs, and decision for the phones [O] or [o], and [E] or [e]. Rules to discriminate the different levels of nasality and stress for some acoustic units, such as [6] and [6~], for example, are also subject of future work.

References

1. Cunha, Celso.: *Língua portuguesa e realidade nacional*. 2^a. ed. atualiz. Rio de Janeiro: Tempo Brasileiro, 1970.
2. Anais do Primeiro Congresso Brasileiro de Língua Falada no Teatro, 1958.
3. Anais do Primeiro Congresso da Língua Nacional Cantada, 1938.
4. Ramos, Jânia M.: *Avaliação de dialetos brasileiros: o sotaque*. In: Revista de Estudos da Linguagem. Belo Horizonte: UFMG. 6 (5):103–125. jan.–jun. 1997.
5. Almeida, M.J.A.: *Étude sur les attitudes linguistiques au Brésil* Université de Montreal, 1979.
6. *Corpus de Extractos de Textos Eletrônicos NILC/Folha de São Paulo (CETEN-Folha)*. <http://acdc.linguatca.pt/cetenfolha/>
7. *Speech Assessment Methods Phonetic Alphabet (SAMPA)*. <http://www.phon.ucl.ac.uk/home/sampa>
8. Alcain, Solewicz, and Moraes: *Frequência de ocorrência dos fones e listas de frases foneticamente balanceadas no Português falado no Rio de Janeiro*. Revista da Sociedade Brasileira de Telecomunicações. Revista da Sociedade Brasileira de Telecomunicações. 7(1): 23–41.
9. Pinto, G.O., F. Barbosa, and F.G. Resende Jr. A.: *Brazilian Portuguese TTS based on HMMs*. 2002. In: Proceedings of International Telecommunications Symposium. Natal, Rio Grande do Norte.