

(R) I Stop Being Afraid of This Problem (1/4) [10 Points]

One common task in developing technologies for human languages is grapheme-to-phoneme conversion (or G2P).¹ In G2P, you convert words written in orthography (the practical writing systems that people use day-to-day) to a standardized phonetic transcription that can be used for recognizing and synthesizing speech, among other things. In this problem, you will help develop G2P for Carib (*Karinja*), a Cariban language spoken by about 7,400 Carib people in Venezuela, Guyana, Suriname, French Guiana, and Brazil. Below are some words in Carib with their pronunciations and meanings (a guide to the phonetic symbols is provided on the next page). Here are some words to get you started:

Word (orthography)	Phonetic transcription	Meaning
<i>aikuma</i>	/ˈaih.kʲu.ma/	to make something juicy
<i>yrama</i>	/ˈɾa.ma/	I am turned
<i>asaperary</i>	/a.ˈsaʔ.pe.ˈra.ɾa/	your cup
<i>ōkaikō</i>	/ˈoŋ.gai.giō/	combs
<i>taweipore</i>	/ta.ˈweih.pʲo.ɾe/	well-lit
<i>saraisarai</i>	/sa.ˈɾai.ʃa.ɾai/	the sound of raking
<i>kynapoaton</i>	/ki.ˈnaʔ.po.ˈjaʔ.ton/	they feel it
<i>anipynary</i>	/a.ˈniʔ.pʲi.ˈna.ɾa/	your love (i.e., the object of your love)
<i>tainawerùke</i>	/ˈtai.nʲa.ˈwe.ruɦ.ke/	having a skin fungus on the hand
<i>asewenàpota</i>	/a.ˈse.we.ˈnah.po.da/	one after another
<i>sapera</i>	/sa.ˈbe.ra/	cup
<i>apo</i>	/ˈaʔ.po/	feel
<i>erèny</i>	/e.ˈreʔ.ni/	nervous motion
<i>ytonoroipory</i>	/ˈto.no.ˈɾoiɦ.pʲo.ɾo/	my <i>Matayba</i> tree
<i>sampura</i>	/ˈsam.bu.ɾa/	drum
<i>sukurusaniry</i>	/su.ˈgu.ɾu.ˈsa.ni.ri/	of the candy
<i>Sipanijorory</i>	/ʃi.ˈbʲa.ni.ˈjʲo.ɾo.ɾo/	of the Spaniard
<i>yjenarĩkepy</i>	/ˈje.na.ˈɾiŋ.gie.bi/	I stop being afraid
<i>sikirima</i>	/ʃi.ˈgʲi.riʔ.mʲa/	to divide into pieces

(Table is continued on the next page.)

1. G2P also showed up in problem (L), Stopping for a Spell. However, solving problem (L) will give you no advantage in solving this problem, and vice versa.



(R) I Stop Being Afraid of This Problem (2/4)

(Table continued from the previous page.)

<i>pěputu</i>	/ˈpɛm.bu.du/	dung beetle
<i>aperěperěka</i>	/a.ˈbe.rɛm.ˈbe.rɛŋ.ga/	you are made to flap
<i>enaro</i>	/e.ˈna.ɾo/	governor
<i>ytuwarõ</i>	/ˈtu.wa.ɾõ/	I am forgotten
<i>pirai</i>	/ˈpi.rai/	piranha
<i>uwěposapariky</i>	/u.ˈwɛm.bo.ˈsaɾ.pa.ˈɾiɾ.ki/	puffiness of the belly
<i>mòwusa</i>	/ˈmoɾ.wu.sa/	to sharpen
<i>jarã</i>	/ˈja.ɾã/	fence

As you can see, some consonants in the phonetic transcription of Carib are different from those of English. To help you understand how these consonants are pronounced, here is some information on their location in the mouth and manner of articulation (ɾ is pronounced farther back in the mouth than r). “Voiced” and “voiceless” indicate that the vocal cords vibrate and don’t vibrate, respectively, during pronunciation. Note also that ^j after a consonant represents palatalization, or softer articulation, and ˈ before a syllable indicates that it is stressed (i.e., pronounced more emphatically, like the first syllable of the geographical feature “desert”, or the second syllable of the food “dessert”). The boundary between syllables is marked with a period.

		Labial (lips)	Coronal (tip of tongue against roof of mouth)	Palatal (middle of tongue against hard palate)	Velar (back of tongue against soft palate)	Glottal (vocal cords)
Nasal	Voiced	m	n		ŋ	
Stop	Voiceless	p	t		k	ʔ
	Voiced	b	d		g	
Fricative	Voiceless		s	ʃ		h
Liquid	Voiced		ɾ, ɽ	j	w	



(R) I Stop Being Afraid of This Problem (3/4)

R1. Fill in the following table to provide rules for the pronunciation of orthographic *y*. Note that for each answer you provide in the “Phoneme” column (i.e., answers **(c)**, **(e)**, **(g)**, and **(h)**), you should answer with a single phoneme (i.e., a single character used in the phonetic transcriptions). Multiple correct answers are possible.

Grapheme	Environment		Phoneme
if <i>y</i>	(a)	it is silent	—
but if <i>y</i>	(b)	it is pronounced as	(c)
but if <i>y</i>	(d)	it is pronounced as	(e)
but if <i>y</i>	(f)	it is pronounced as	(g)
otherwise	—	<i>y</i> is pronounced as	(h)

R2. Fill in the following table to provide rules for the pronunciation of orthographic *p*, *t*, and *k*. For this question, you should ignore palatalization (none of your answers in the “Phonemes” column should include the ^j symbol). Note that for each entry you fill in the “Phonemes” column (i.e., **(c)**, **(e)**, and **(f)**), you should provide three answers of one or more phonemes each, and you should order your answers respectively, separated by commas (i.e., with the pronunciation of *p* first, then *t*, then *k*). Multiple correct answers are possible.

Grapheme	Environment		Phonemes
if <i>p</i> , <i>t</i> , or <i>k</i>	(a)	they are pronounced as	h.p, h.t, h.k respectively
but if <i>p</i> , <i>t</i> , or <i>k</i>	(b)	they are pronounced as	(c) respectively
but if <i>p</i> , <i>t</i> , or <i>k</i>	(d)	they are pronounced as	(e) respectively
otherwise	—	they are pronounced as	(f) respectively



(R) I Stop Being Afraid of This Problem (4/4)

R3. Fill in the blanks in the following table, using the G2P rules you developed based on the data (including, but not limited to, the rules from **R1** and **R2**).

Word (orthography)	Phonetic transcription	Meaning
<i>makopamy</i>	(a)	to grow dark
<i>aitopòma</i>	(b)	homeless
<i>kerikeri</i>	(c)	a species of bird
<i>parimy</i>	(d)	son in law of
<i>kurijara</i>	(e)	boat
<i>ykurijarary</i>	(f)	my boat
<i>tykupimy</i>	(g)	what needs to be bathed

R4. Explain the G2P rules you developed for Carib based on the data. You do not need to repeat your rules from **R1** and **R2**.

