

Spoken Language Processing (CSCI 780/3813)

Prof. Rosenberg

Fall 2011

Homework 1 – Representing Speech Sounds

Question 1a – 5 pts. Write your name in IPA.

Question 1b – 5 pts. Write your name in ARPABET.

Question 2 – 10 pts. Transcribe the following words in IPA

Dark

Suit

Greasy

Wash

Water

Question 3 – 10 pts. Write words that start with phones with the following features.

Voiced bilabial nasal
Voiced velar stop
Unvoiced velar stop
Unvoiced glottal fricative
Voiced dental fricative
Alveolar lateral approximant (liquid)
High back rounded vowel
High front unrounded vowel
Low front unrounded vowel
Schwa

Question 4 – 10 pts. Write the English words that are pronounced as follows in IPA.

[grɪn]
['ɔrɪndʒ]
[spɪtʃ]
['saɪəns]
['skɛdʒul]

Question 5 – 10 pts. Write the English words that are pronounced as follows in ARPABET.

[t ay p ih ng]

[p ah z ax l]

[dh ow]

[s ay k ih k]

[f l aw ax r]

Question 6 – 40 pts. Write a command line program that reads a text file containing IPA Unicode, and produces a text file with the same content in ARPABET.

Masters students – Extend this program to include a command line option to convert a file containing ARPABET and generates a Unicode file containing IPA material. That is, reverse the conversion done in the first part of the assignment.

You may write this program in java, c + + or python. Your program must compile and run on venus.cs.qc.cuny.edu.

Here is a resource for representing IPA in Unicode.

[http://en.wikipedia.org/wiki/IPA_Extensions_\(Unicode_block\)](http://en.wikipedia.org/wiki/IPA_Extensions_(Unicode_block))

50% of the points are for correctness of the conversion, 30% for compilation, 20% for documentation and style.

Documentation requirements: 1) a readme file that describes how to compile and run your program (on venus). 2) each method in your program must have a comment describing

what it does, what its input and outputs are. 3) Variables and method names must have intuitive and understandable names. 4) avoid extra blank lines (e.g. at the start or end of blocks) in your code unless structuring related lines of code together. 5) be consistent with naming styles (e.g. if you use camel casing for method names, like `readTextFile(...)`, use this for every method) 6) be consistent with spacing, e.g. “`if (a == b)`” or “`if (a = = b)`”. “`a = b + c`” (no spaces) or “`a = b + c`” (all spaces), but NOT “`a = b + c`” (spaces around equals, but not plus)

Question 7 – 10 pts. Identify your team for the term project. Each team should have three members. If possible, if you are a graduate student, try to make your team include only graduate students, and vice versa. If there are concerns about the size of the team or the makeup of the members, let me know, and I will try to work with you to construct a team.