Pronounceability for a generated character sequence

Pseudocode

Cathryn Ploehn

Method within the that returns an integer score for level of pronounceability of a word depending on how many different matching syllables

Uses the onset – nucleus – coda syllable classification scheme used in English phonology.

Pull letter groups that match the following:

Regex: [possible onset consonants][possible nucleus letters][possible coda letter sequences]

Possible onset consonants (following the rules at <http://en.wikipedia.org/wiki/English_phonology#Onset>):

P

B

T

D

K

G

F

Ph

V

Th

S

Z

M

N

L

R

W

H

Y

Sh

Ch

J

Pl

Bl

Kl

Gl

Pr

Br

Tr

Dr

Kr

Gr

Tw

Dw

Gw

Kw

Pw

Sp

Sk

St

Sm

Sn

Sph

Sth

Spl

Scl

Spr

Str

Scr

Squ

Sm

Sphr

Wr

Gn

Xy

ps

Possible nucleus letters:

I

E

A

O

U

Oo

Ea

Ir

Y

Oy

ee

ea

ou

o

ow

Possible coda letters:

P

B

T

D

K

G

F

Ph

V

Th

S

Z

M

N

L

R

Y

Sh

Ch

Lp

Lb

Lt

Lch

Lge

Lk

Rp

Rb

Rt

Rd

Rch

Rge

Rk

Rgue

Lf

Lve

Lth

Lse

Lsh

Rf

Rve

Rth

Rce

Rs

Rsh

Lm

Ln

Rm

Rn

Rl

Mp

Nt

Nd

Ch

Nge

Nk

Mph

Mth

Nth

Nce

Nze

Gth

Ft

Sp

St

Sk

Fth

Pt

Ct

Pth

Pse

Ghth

Tz

Dth

X

Lpt

Lfth

Ltz

Lst

Lct

Lx

Mth

Rpt

Rpse

Rtz

Rst

Rct

Mpt

Mpse

Ndth

Nct

Nx

Gth

Xth

xt

ng

For each letter group that matches the expression, add one point to variable “score”

Return score