IPA HANGMAN

Reed Blaylock
reed.blaylock@gmail.com

Sarah Harper
skharper@usc.edu

University of Southern California

WORDS

There are 36 words with every English sound (at least twice)

Selection method:

1. Initial pool: 4,269 words in both the CMU dictionary and COCA's 5,000 most frequent words
2. Trimmed pool 1: 359 words from the initial pool that contained at least one of 6 rare sounds ("rare" = less than 100 tokens: /a͡ʊ/, /ð/, /ɔ͡ɪ/, /θ/, /ʊ/, and /ʒ/)
3. Trimmed pool 2: 182 words from step 2 that contained at least one of a new set of rare sounds ("rare" = less than 20 tokens)
4. All 4 remaining instances of /t͡ʃ/ were preserved. 32 other words were pseudo-randomly selected until the full set of 36 had at least 2 tokens of every sound

SCORING

The word bonus is the Levenshtein distance between each word's orthographic and IPA form

Levenshtein distance: +1 for every sound that must be changed, added, or removed to get from the orthographic form to the IPA form

Orthography:	" n	o	i	s	e "	
IPA:	/ n	ɔ͡ɪ		z	/	
Levenshtein:	0	+1	+1	+1	+1	= 4

IPA HANGMAN RESOURCES

Blaylock , R. and Harper, S. (2018). IPA Hangman. http://reedblaylock.com/ipahangman.html.

MacKichan, P. (2004). /fəˈniːmɪk/ Hangman http://www.e-lang.co.uk/mackichan/call/phonhang/phonhang.html.

SammyJackson. (2016). IPA Hangman! https://scratch.mit.edu/projects/110936363/.

REFERENCES

Weide, R. L. (1998). The CMU pronouncing dictionary. URL: http://www.speech.cs.cmu.edu/cgibin/cmudict.

Davies, M. (2008). The corpus of contemporary American English. BYE, Brigham Young University.

Levenshtein, Vladimir I. (February 1966). "Binary codes capable of correcting deletions, insertions, and reversals". Soviet Physics Doklady. 10 (8): 707–710.