



Phoneme-Grapheme Mapping Project

Erica Varga

College Park Scholars – Science & Global Change Program
Neurobiology and Physiology
evargal@umd.edu

College Park Scholars Academic Showcase, May 1, 2020



Introduction

My research deals with *phonemes*- the smallest unit of sound in words, and *graphemes*- the series of letters that represent phonemes in a word, and how to map them to each other. Different series of letters (graphemes) can have different corresponding sounds (phonemes).

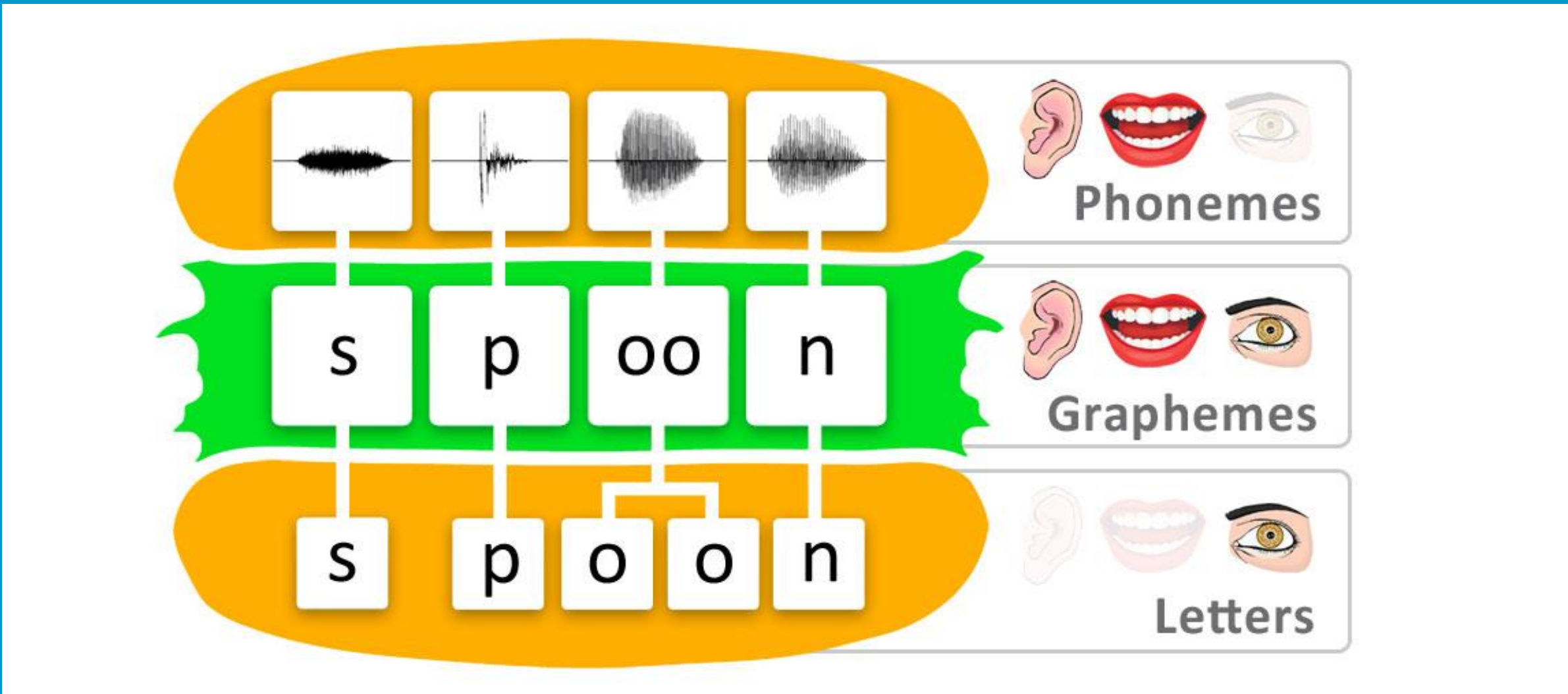


Image from readingdoctor.com.au

Site Information:

Laboratory for the Neurodevelopment of Reading and Language
3304 Benjamin Building 3942 Campus Dr.
College Park, MD 20742

Jeremy Purcell

Mission: to study how the brain learns to read and the cognitive and neural bases of reading and language ability and disability

Goal of my research: to determine whether pronunciation/ spelling consistency measures derived from an academic study done by Hannah (see ‘Materials’) are superior to other, preexisting measures.

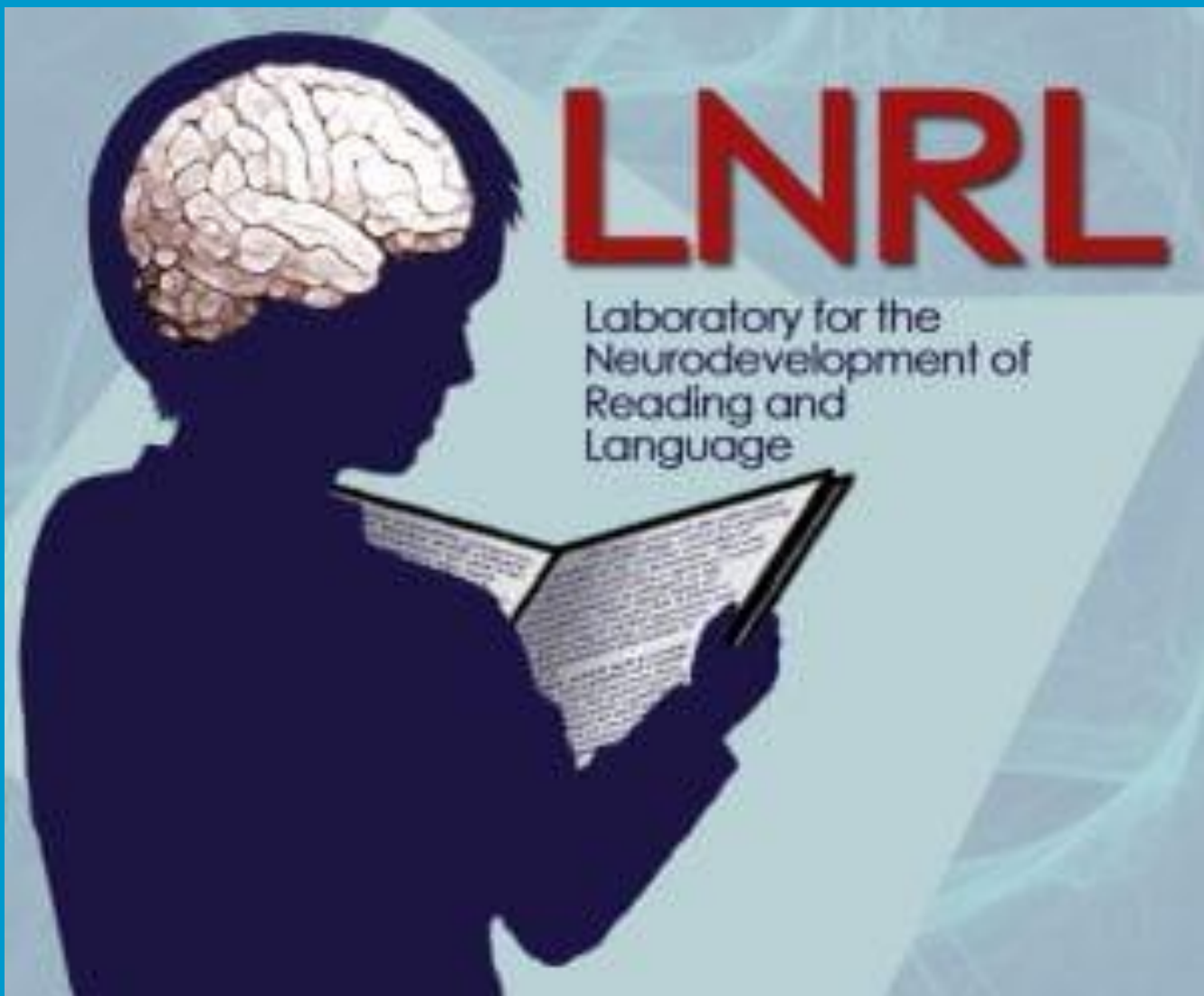


Image from lnrl.umd.edu. This image depicts the image used on the website of the lab that I worked in for this project.

Materials:

Fry, E. 2004. Phonics: A Large Phoneme-Grapheme Frequency Count Revised. *Journal of Literary Research* **36**: 85-98.

Hanna, P.R.. 1966. Phoneme-Grapheme Correspondences as Cues to Spelling Improvement. *Education Resources Information Center* CRP-1991: 1-144.

Methods:

Using Excel, I created a phoneme- grapheme mapping table, consolidated it with the findings of an academic study conducted by Fry (see ‘Materials’), and then created a grapheme-phoneme mapping table to track how probable certain phonemes are for a given grapheme. Additionally, I created a guide detailing how to go from the Hannah tables to the tables that I created, so anyone who uses my table would know how it was made.

Results:

It is still relatively early in the research process, so we do not have any results yet, but the next steps in the process include finding words with high phoneme-grapheme and low grapheme-phoneme probabilities and vice versa, and set up experiments to see if they are processed in different ways/ use different neural pathways.

	Grapheme	Hannah Symbol	Phoneme	Grapheme_Phoneme	GP_freq	GP_Prob
7783	A	A	eɪ	A + eɪ	1002	0.129
	A	A2	eə	A + eə	64	0.008
	A	A3	æ	A + æ	4192	0.539
	A	A5	ɔ	A + ɔ	821	0.105
	A	E3	e	A + e	262	0.034
	A	l3	ɪ	A + ɪ	4	0.001
	A	ə	ə	A + ə	1438	0.185
1226	A-E	A	eɪ	A-E + eɪ	790	0.644
	A-E	A2	eə	A-E + eə	50	0.041
	A-E	A3	æ	A-E + æ	147	0.120
	A-E	A5	ɔ	A-E + ɔ	48	0.039
	A-E	l3	ɪ	A-E + ɪ	187	0.153
	A-E	E3	e	A-E + e	1	0.001
	A-E	ə	ə	A-E + ə	3	0.002
283	AI	A	eɪ	AI + eɪ	208	0.735
	AI	A2	eə	AI + eə	46	0.163
	AI	A3	æ	AI + æ	1	0.004
	AI	E3	e	AI + e	4	0.014

This image demonstrates a small portion of the Grapheme-Phoneme table that I created from the modified mappings found in the Hannah paper, which presented multiple Phoneme-Grapheme mappings.

Discussion:

In this table, I took the data from the Hannah (and Fry) tables and reorganized it so that it mapped graphemes to phonemes (instead of the other way around). The leftmost column shows the total frequency of the given grapheme in all of the words analyzed in the studies, while the frequency column gives the frequency for that specific pairing. Finally, the last column shows how probable the phoneme-grapheme pairing is compared to the other phonemes that also can map to that phoneme.

Future Work:

The tables that I made will help us determine how correlated the G-P probability mappings for each monosyllabic word is to the lexical decision time/voice onset time in reading (i.e. how long it takes a person to say a word after they read it). Then, we will see if the consistency measure we derived better predicts reading performance compared to other preexisting measures.

Acknowledgments:

Many thanks to Dr. Holtz and Dr. Merck, my SGC advisors, Dr. Purcell, my research mentor, and to Dr. Bolger, who runs the lab.

