# Orthographic Representation Learning for Modeling Dyslexia

UCONN
UNIVERSITY OF CONNECTICUT

Henry George Wolf VII

Haskins Laboratories

*Abstract –* **Simulations of reading with dyslexia first used explicit mapping rules, but now typically use learned mappings. Such models still use hand-coded input and output. This project proposes using convolutional neural networks to learn the orthographic input from images of text. This could provide insights into the computational, cognitive, and neurobiological development of orthographic representations in children with dyslexia. Further, it would allow for a direct comparison of learning to read across languages.**

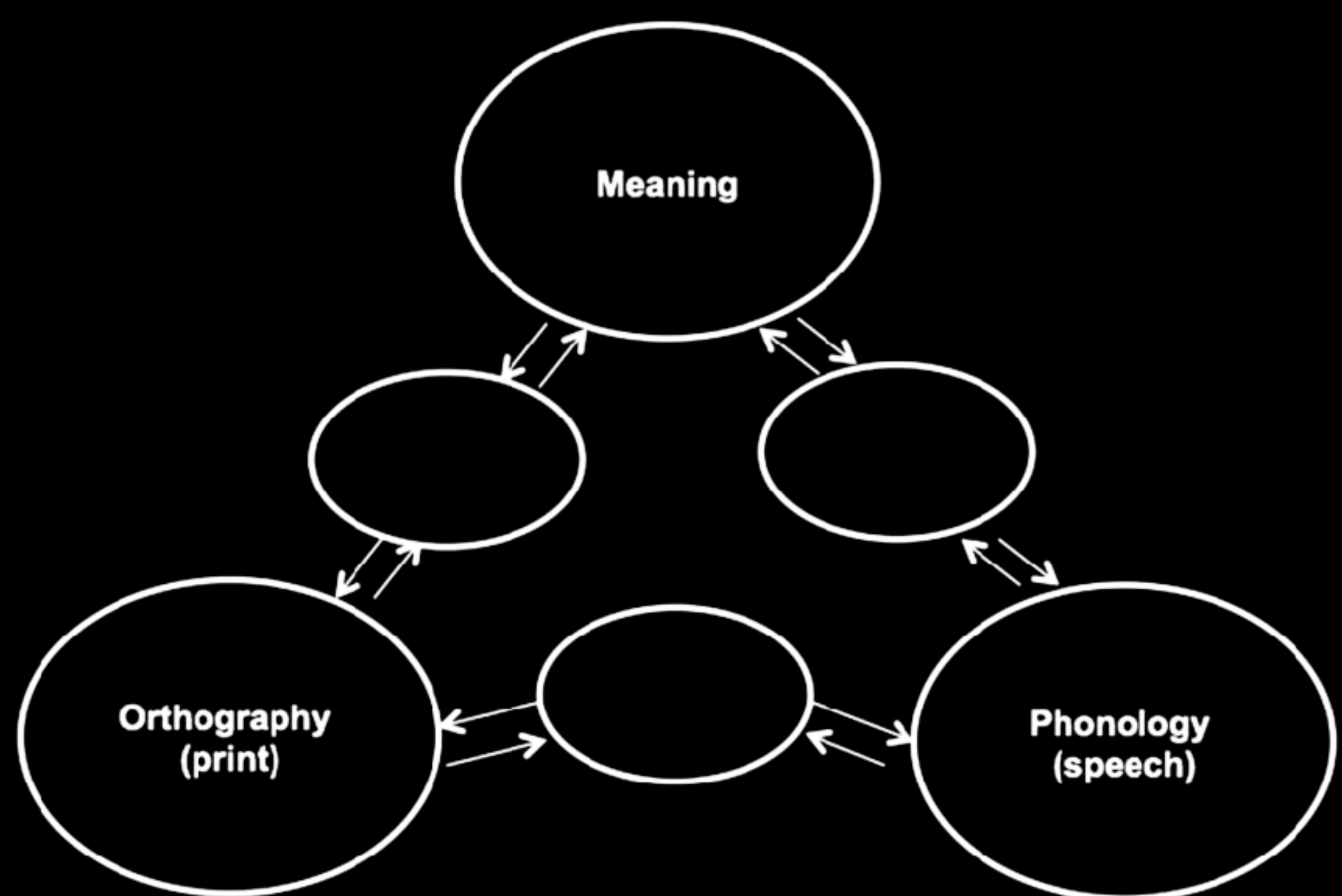
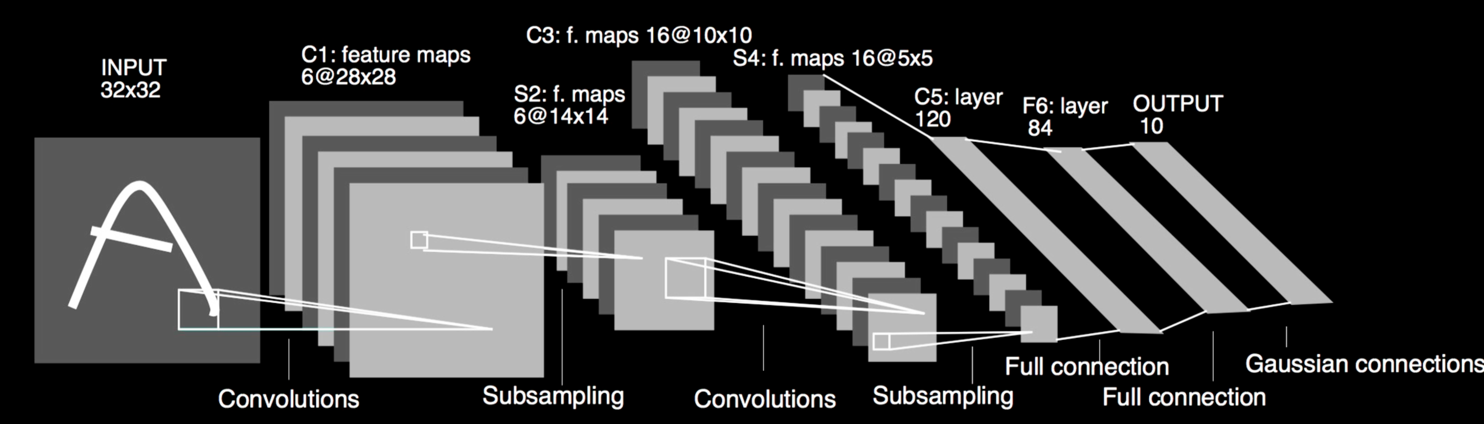



*Figure 1 –* Triangle Model of Reading

*Modeling Reading –* The triangle model of reading (Figure 1) is based on the principles of the connectionist framework. The input is typically hand-coded orthographic representations. The output can be phonological or semantic representations. Each input presentation causes a flow of activation through the network until a stable output representation is achieved. This model simulates psycholinguistic phenomena related to reading, such as frequency and consistency effects. Modifications of the hyperparameters of the network (e.g. learning rate, Gaussian noise, layer size) allow for simulations of typical and dyslexic reading development, whereas modifications of the training corpus allow for simulations that run the gamut from impoverished reading exposure to targeted intervention. Using hand-coded representations, researchers make the assumption they are the same for all readers. This also makes cross-language models difficult, as representations likely differ across writing systems.

*Figure 2 –* Convolutional Neural Network (example)




*Orthographic Representation Learning –* This project makes use of deep convolutional neural networks (Figure 2) to learn orthographic representations from images of text. In addition to the advance in psycholinguistic modeling theory by limiting the assumptions made by researchers, this system of learning representations provides more hyperparameters (e.g. depth, kernel size, number of feature maps) that can be modified and which may better characterize individual differences between typical and atypical reading.

*Model Structure –* Input to the model consists of approximately 6,000 images of size 64x16 pixels. Each image has a typed word in the center. There are three convolutional layers with 3x3 kernels and ReLU activation. Each is followed by a 2x2 max pooling. There is a single fully-connected layer with 500 hidden units and ReLU activation. The output consists of a slot-based representation for the 42 phonemes of English and softmax activation. There are ten slots, for a total of output layer size of 420. Accuracy is determined as the correct unit in each slot being the one most activated. All slots must be correct.

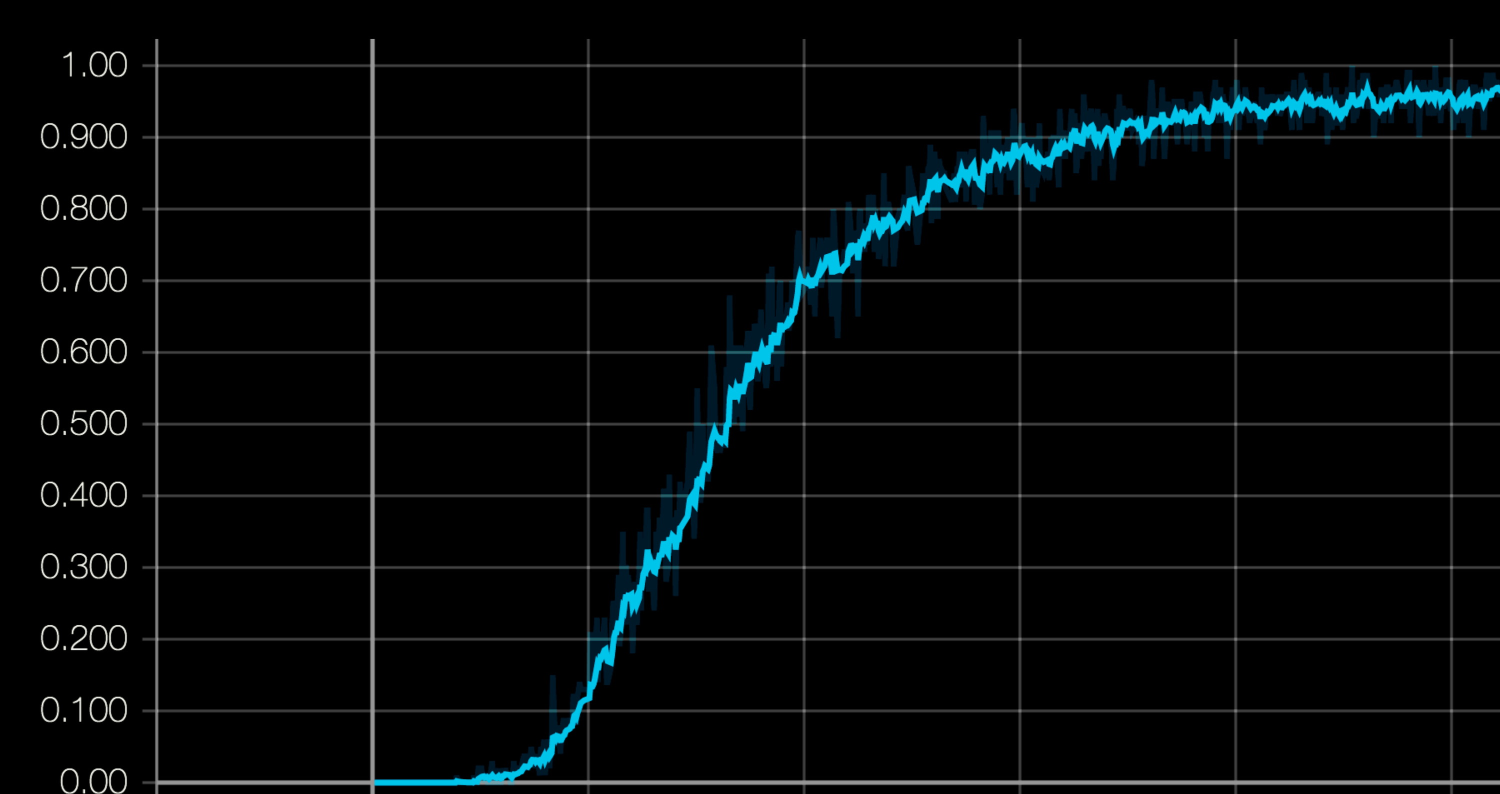



*Figure 3 –* Image to Orthography Accuracy over Time (Trials)

*Preliminary Results –* Early versions of this network mapped from images to orthographic representations and achieved over 95% accuracy after 100k presentations (Figure 3). Due to the complexity of the relationships in English orthography and phonology, the current version achieved 90% accuracy was achieved after 1m trials (Figure 4).

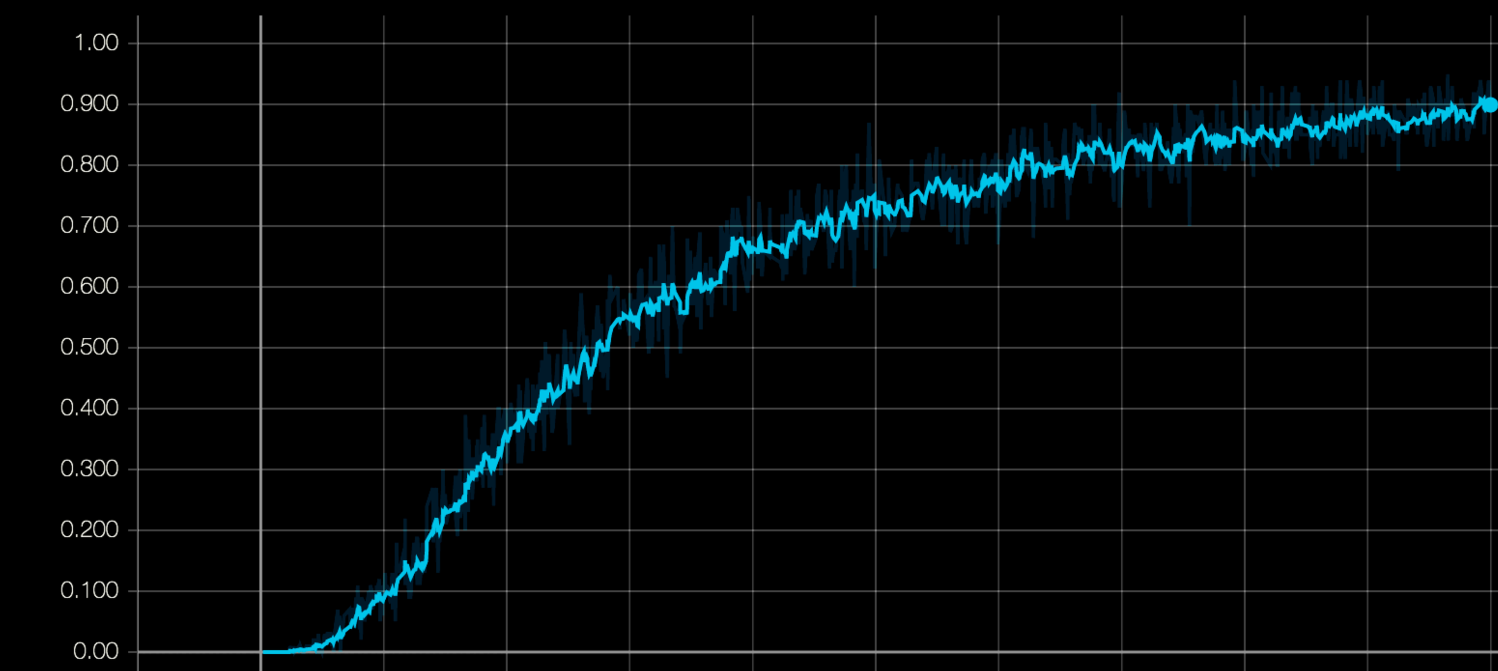



*Figure 4 –* Image to Phonology Accuracy over Time (Trials)

*Future Directions –* The model will be tested on a set of nonwords to determine its ability to generalize. It will then be compared with previous models in generating psycholinguistic phenomena. A wide hyperparameter search will be used to determine their effects on the size of psycholinguistic phenomena and reading accuracy. Additional languages will use used as input for cross-language analysis.

*References –*

Harm, M., & Seidenberg, M. S. (1999). Reading acquisition, phonology, and dyslexia: Insights from a connectionist model. Psychological Review, 106, 491-528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. Psychological Review, 111, 662-720.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Hannagan, T., Ziegler, J. C., Dufau, S., Fagot, J., & Grainger, J. (2014). Deep learning of orthographic representations in baboons. PloS one, 9(1), e84843.

*Contact –* **henry.wolf@uconn.edu**
**chaoticneural.com**
**@chaoticneural**