From Speech to Print A Neurocomputational Model of the Predictors of Reading Based on Learned Speech Representations

Martín Dutra
Centro Interdisciplinario en Cognición para la Enseñanza y el Aprendizaje (CICEA), Universidad de la República, Montevideo, Uruguay

Daniel Zangrando
Centro Interdisciplinario en Cognición para la Enseñanza y el Aprendizaje (CICEA), Universidad de la República, Montevideo, Uruguay

Juan Valle-Lisboa
Sección Biofísica y Biología de Sistemas, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay
Centro Interdisciplinario en Cognición para la Enseñanza y el Aprendizaje (CICEA), Universidad de la República, Montevideo, Uruguay

Abstract

We present a reading acquisition model including state of the art architectures that reduce to a minimum ad hoc assumptions. We apply this model to compare early reading acquisition between opaque and transparent orthographies. We use speech representations obtained from wav2vec 2.0 as the auditory part of our model and the pre-trained convolutional neural network CORnet, to represent visual processing. Both modules converge onto a Long-Short-Term Memory (LSTM) model that we train to recognize words. We train the LSTM to recognize auditory presented words, letter sounds and visually presented letters. In Spanish, further training the model to recognize a word from a sequence of visual letters and letter sounds is enough to reach appropriate levels of visual word recognition, generalizing to words it has not seen visually, whereas the same does not happen in French or English. The model explains the difference in relevant predictors in opaque and transparent orthographies.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*