Assignment 8, Introduction, Academic English 2

Philip Bouman (10668667)

In most cases in natural languase there is no obvious relationship between the form and the meaning of a word. However in contrast to this arbitrariness there are instances of similarity between word form and meaning. A well studied phenomenon of such a relationship of iconicity is the *bouba-kiki* effect, in which shapes are consistently labeled with non-words. The goal of this research is to find if such relationships also exist between the shape and the sound of individual characters (letters). To avoid language-specific relations, a wide variety of writing systems (scripts) will be taken into consideration to discover general similarities between form and sound. To classify the sounds (phonemes) in a consistent way, a phonological encoding scheme will be used. This scheme contains vector representations of the phonemes as described in the International Phonetic Alphabet (IPA). In the first part of the research a deep convolutional neural network will be trained to classify images of characters based on their related sounds. A deep convolutional neural network is a machine learning architecture consisting of multiple stacks of layers and has proved to be very succesfull for applications with two dimensional data. The second part of the project will focus on getting a better understanding of the representations learned by the network. This will mainly be done by investigating and visualizing features learned in individual layers and smaller combinations of layers. The results will be evaluated by measuring the level of accuracy on the test data.