A Connectionist Description of Saliency in Pre-Reading Children

Kevin C. Larson (klarson@uts.cc.utexas.edu)
Philip B. Gough (gough@psy.utexas.edu)
University of Texas, Austin
310 Mezes Hall, Austin, TX 78712 USA

Children and the Thumbprint

Gough (1993) showed that when pre-reading children are taught to read words that have a thumbprint nearby, the children are more likely to use the thumbprint than the letters of the word to identify the word.

32 children were shown either 4 similar words {bag, bat, rag, rat} or 4 dissimilar words {box, leg, sun, rat}. One of the words would have an ink thumbprint mark on the card in addition to the word. The children would learn the words until they could go through the complete set of words twice without making a mistake.

Gough found that children learn their initial words through selective association. Children are faster at learning dissimilar words than similar words, and much faster at learning words that have a thumbprint on the card than words that had no thumbprint. Children are better at learning visually distinct words. A thumbprints creates the greatest visual distinction and dissimilar words are more visually distinct than similar words.

Children are fastest at learning words when the salient clue of the thumbprint is present because they're using the thumbprint to recognize the entire word. When tested on a thumbprint presented by itself and the word previously paired with the thumbprint by itself, 26 of the 32 children could recognize the word by the thumbprint alone while only 13 of the 32 children recognized the word in the absence of the thumbprint.

Failed Model

A simple neural network model of reading does not imitate the children's performance. Two backpropigation networks (one for each of the two lists of words) with letter input units, a layer of hidden units, and word output units were tested. We found that dissimilar words were learned faster than similar words, but words with a thumbprint unit (represented by a unique input unit) were not learned significantly faster than words without the thumbprint unit.

The network fails to describe children's behavior because it doesn't account for the visual salience of the thumbprint. The network treats the thumbprint as another letter instead of a character with extra-special importance.

A Salient Thumbprint

We found that by increasing the thumbprint's importance at the input level, the network does a better job at simulating the children's behavior. The number of units representing the thumbprint was increased from 1 to 50. Figure 1 shows that the network is now much faster at learning words that contained a thumbprint than words that contained no thumbprint and is also faster at learning words in the dissimilar list than words in the similar list.

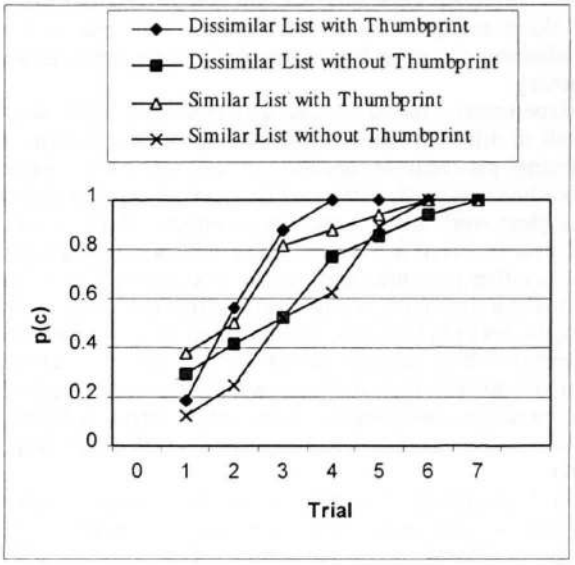

Figure 1: Learning rates in the network model with a salient thumbprint representation.

When the network is tested on a thumbprint presented by itself and the word previously paired with the thumbprint presented by itself, the network now performs similarly to children. The network recognizes the word by the thumbprint alone 25 of 32 times, and only 4 of 32 times when the word is presented without the thumbprint.

In the earliest stages of reading acquisition children use visual features to identify entire words. Neural network models of reading acquisition and visual attention can partially account for this by increasing the size of the input representations for salient features.

References

Gough, P. B. (1993). The beginning of decoding. *Reading and Writing,* 5, 181-192.

Olshausen, B. A., & Koch, C. (1995). Selective Visual Attention, in *The Handbook of Brain Theory and Neural Networks* (ed. Michael A. Arbib), MIT Press, 837-840.