

# Why localist connectionist models are inadequate for categorization

In *The Behavior and Brain Sciences* (2000). Vol. 23, No. 4, p. 477.

Robert M. French  
Psychology Department (B32)  
University of Liège, Liège, Belgium  
rfrench@ulg.ac.be

Elizabeth Thomas  
Institut Léon Frédéricq  
University of Liège, Liège, Belgium  
ethomas@ulg.ac.be

## Abstract

We claim that two categorization arguments pose particular problems for localist connectionist models. The internal representations of localist networks do not reflect the variability within categories in the environment, whereas networks with distributed internal representations do reflect this essential feature of categories. Secondly, we provide a real biological example of perceptual categorization in the monkey that seems to require population coding (i.e., distributed internal representations).

In spite of Mike Page's bold frontal assault on distributed connectionism, we wish to point out what appear to us to be two significant problems with this type of localist network — namely:

### The problem of category variability

Consider two categories, “fork” and “chair.” The variability within the first category is very low: there just aren't that many different kinds of forks. Chairs, on the other hand, come in all different shapes, sizes and materials: they range from beanbag chairs to barstools, from overstuffed armchairs to rattan chairs, from plastic lawn chairs to that paragon of ergonomic design, the backless computer chair that you kneel on; some have four feet, some three, some none; some have backs, some don't; some are made of metal, others plastic, others wood, others, cloth and Styrofoam pellets, etc. In other words, the variability within the category of *chair* is enormous.

But in the localist model proposed by Page, and in localist models in general, *this information about category variability is lost*. In distributed models, it takes more hidden nodes to encode a category with high-variability than one with low variability. In other words, the internal representations reflect external category variability. However, the category nodes in localist networks are unable to reflect this differential variability-in-the-environment of various categories. The one-node internal “representation” corresponding to the extremely low-variability category “fork” is precisely the same as the one-node internal representation corresponding to the highly variable category “chair.”

Why is this a problem? Most significantly, because of the well-documented fact of category-specific losses: in general, naming of inanimate objects is found to be better preserved than naming of animate objects (Warrington and Shallice, 1984; Funnell & Sheridan, 1992; Farah, Meyer, & McMullen, 1996; etc.). A model with distributed internal representations can handle this problem quite simply: low-variance categories (e.g., many natural kinds categories, like *cat*, *horse*, etc.) are encoded over fewer “units” than high-variance categories (e.g., many artificial kinds categories, like *chair*, *tool*, etc.) Random lesioning of the model will be more likely, on average, to destroy the representation of a category with low-variability (e.g., natural kinds categories) that is coded over a small number of units than a high-variability category (e.g., artificial kinds categories) coded over a large number of units. Localist models in which all the category nodes are the same will have

considerable problems explaining category-specific deficits of this kind, especially when the featural inputs to the internal category representations remains intact. If, on the other hand, we assume differing degrees of variance associated with the internal encoding of different categories, these kinds of deficits can be predicted in a straightforward manner, as French (1997) and French & Mareschal (1998) have shown using a dual-network architecture based on the hippocampal-neocortical separation proposed by McClelland, McNaughton, & O'Reilly (1995).

As Page points out in his target article, we have argued for the necessity of “semi-distributed” representations in connectionist models for many years. But “semi-distributed” does not mean localist. “Semi-distributed” representations preserve category variance information; localist representations do not. Further, it seems crucial to us that these semi-distributed representations *emerge* as a result of learning.

### Biological category representations

Page is right in pointing out that some of what is called population or ensemble coding in biological systems can be viewed as localist. For example, even though broadly tuned, cells of the motor cortex have their maximum activity tuned to a particular direction (Georgopoulos et al, 1993). One should therefore be able to ascertain the direction being represented by looking at the activity of individual neurons (or very small groups of neurons). However, an example of a cognitively relevant task that cannot be achieved in this fashion can be found in the anterior temporal cortex. Vogels (1999) reports on the responses of cells in this area during a tree, non-tree categorization task by a monkey. Most of the cells were stimulus selective, (i.e., they did not respond to all of the presented stimuli) and responded to both trees and non-trees. The maximum response of these neurons was not tuned to either category. Even though it was the case that certain (category-selective) neurons responded to particular subsets of tree exemplars, *no individual neuron (or small set of neurons) responded to all of the presented trees*, while not responding to any non-tree. These category-selective neurons alone did not appear to play an important role in the categorization performance of the monkey (Thomas, Van Hulle, & Vogels, 1999). In other words, population coding was necessary for the monkey to correctly categorize all exemplars in the test set.

### **References**

- Farah, M., Meyer, M., & McMullen, P. (1996). The living/non-living distinction is not an artefact: Giving an a priori implausible hypothesis a strong test. *Cognitive Neuropsychology*, 13, 137-154.
- French, R. & Mareschal, D. (1998). Could category-specific semantic deficits reflect differences in the distributions of features within a unified semantic memory? In *Proceedings of the 20<sup>th</sup> Annual Conference of the Cognitive Science Society*, 1998, NJ:LEA, 374-379.
- French, R. (1997). Pseudo-recurrent connectionist networks: An approach to the “sensitivity–stability” dilemma. *Connection Science*, 9(4), 353-379.
- Funnell, E. & Sheridan, J. (1992). Categories of knowledge? Unfamiliar aspects of living and non-living things. *Cognitive Neuropsychology*, 9, 135-153.
- Georgopoulos G., Taira Masato and Lukashin A (1993) Cognitive neurophysiology of the motor cortex. *Science* 260: 47-51.
- McClelland, J., McNaughton, B., & O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*. 102, 419–457.
- Thomas E., Van Hulle M. and Vogels R (1999) The study of the neuronal encoding of categorization with the use of a Kohonen network. *Proceedings, 5<sup>th</sup> Neural Computation*

- and Psychology Workshop. Connectionist Models in Cognitive Neuroscienc*, pp. 218-227. Birmingham, England, Springer Verlag.
- Vogels R. (1999) Categorization of complex visual images by rhesus monkeys. Part 2: single cell study. *European Journal of Neuroscience* 11: 1239-1255.
- Warrington, E. K. & Shallice, T. (1984). Category-specific semantic impairments. *Brain*, 107, 829-859.