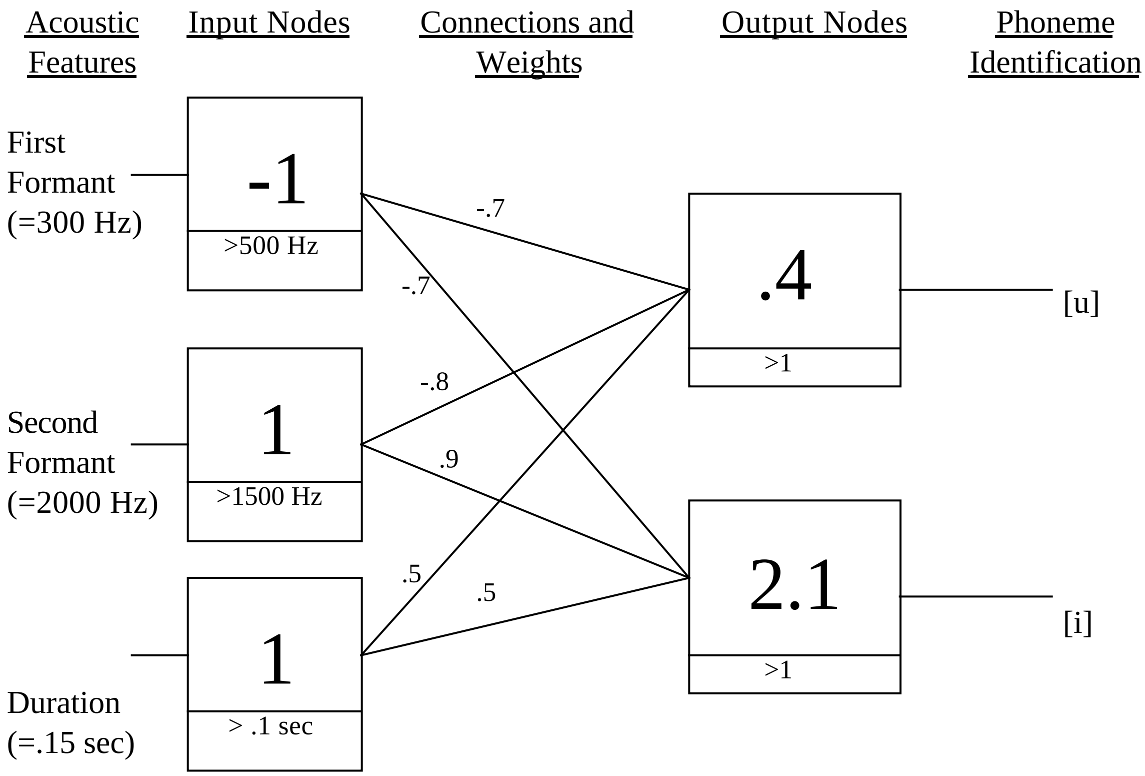

This is a neural network that can recognize either the vowel [u] or the vowel [i]. It takes in acoustic data from the environment (first formant, second formant and duration values for the vowel) and relates them through a series of nodes and connections to derive activation levels for both the [u] and the [i] nodes.

In this case, the relevant acoustic values are 300 Hz, 2000 Hz and .15 seconds. These input values yield an activation of 2.1 for the [i] node and .4 for the [u] node. Since the [i] node has an activation level which exceeds its threshold (1), the model is considered to have "perceived" the incoming sound as an [i].

Ling/Psych 371
April 12, 2000

Neural Networks: Training machines to perceive speech

Neural networks are models that are based on a metaphor for how the human brain works. They have become very influential in everything from electronic engineering to cognitive psychology. In the study of speech perception, neural networks have emerged as a practical alternative to modularity theory. The TRACE model that is mentioned in the book is an example of a neural network.

On the following page is a very basic neural network exemplifying how a computer-based model may be trained to identify phonemes from an acoustic signal.

This neural network includes the basic features of any neural net--
1. Nodes. Each box in the network is a node; these are metaphors for individual nerve cells in the human brain.
2. Connections. Nodes may be connected with one another (just as individual nerve cells are connected in the brain); the strength of these connections is represented with a numerical weight.
3. Activation. Individual nodes may be either on or off (represented by values of 1 or -1 in the node), just as nerve cells in the brain may be either electrically activated or unactivated.
4. Thresholds. A node will only activate when input to that node (from either outside of the system or from other nodes) reaches a certain level, called a threshold.

· Once a node is activated, it will discharge its activation to other nodes in the system through its weighted connections.
· The amount of activation an active node sends to another node is the product of its activation (1 or -1 in this case) and the weight of the connection between the nodes.
· The total amount of activation a node receives is the sum of all the products of connection weights and activation levels of the nodes it has connections with.
· Connections may be excitatory (positive) or inhibitory (negative).

<u>Other Common Features of Neural Networks</u>

· Competition may exist between nodes on the same level. For example, if there is an inhibitory connection between the [i] and the [u] nodes in our basic model, then a high level of activation for [i] will drive down the activation level for [u]. Establishing such connections can make a "winner take all" method of categorization possible (which is also, apparently, one of the aspects of human speech perception).

· Most neural networks have more than two levels of nodes. The TRACE model, for instance, has a feature level, a phoneme level, and a word level. Neural nets may also have levels that don't specifically represent any kind of linguistic or categorical structure. These are "hidden levels" of nodes.

· Top-down effects can be captured in neural networks by establishing connections that go from more abstract levels (like word or sentence nodes) to less abstract levels (like phoneme or feature nodes). Activation at higher levels may, therefore, excite or inhibit activation at lower levels.

· Another top-down effect is "bias". A node at some abstract level may have a high level of activation even before any bottom-up information comes into the system. Such predispositions may reflect attitude, expectation, likelihood or frequency effects. If such biases cannot be overcome by evidence coming in from the outside world, then the network will miscategorize the stimulus and "perceive whatever it wants to perceive" instead of facing up to the truth, as it were.

· The neural network approach to modelling speech perception presents a challenge to the modularity thesis, which maintains that modules must be domain-specific and not built up from more fundamental mental processes (like calculating associations between nodes in a system). If neural nets are taken to represent the way the entire brain functions, then the perception of speech would also interact with information from any other brain processing that was going on (i.e., there is a lack of domain-specificity). Neural networks also allow for top-down effects of attention and memory, etc.