Abstract: The results of a series of gating experiments (1) are modeled using a spreading-activation connectionist model, the Integration-Competition model (2). In this model, phoneme and feature representations compete with one another such that featural distinctions correspond to different weights assigned to the connection between feature and phoneme layers. The model simulates the increase in identification scores across duration conditions found in the perception experiments by an increase in activation of the target phoneme with successive timesteps. The overall increase in identification scores for the Place group in the gating experiments is simulated by eliminating from the calculations those weights corresponding to the place feature in the model. The particular consonant confusion patterns observed in the gating experiments are also predicted by the model.

PERCEPTION EXPERIMENT

The perception of voicing, manner, and place of articulation was investigated in a series of gating experiments in which subjects identified English consonants in six different duration conditions. The stimuli were created from natural speech tokens excised from C[a] syllables. The consonants were then gated to include ten to fifty percent of the total consonant duration from the onset, to which vowel babble was appended. In Experiment 1, subjects identified eight consonants [p, t, b, d, f, s, v, z] and results were analyzed in terms of correct identification of the following features: voiced or voiceless, stop or fricative, and labial or alveolar. In Experiment 2, subjects were presented with two mutually exclusive sets of the same stimuli such that the information for a particular feature was given (e.g. voicing given in the sets [p, t, f, s] and [b, d, v, z]). A comparison of the results of the two experiments showed that the subjects given information about place of articulation (that is, the Place group - subjects who have no place decision to make) had significantly higher identification scores across the duration conditions than did either Experiment 1 subjects or the Manner group. Information transmission analysis confirmed that for the Place group, more information about voicing and manner was transmitted in the 20, 30, and 40% duration conditions. This analysis shows that the pattern in the identification scores cannot be attributed to a response bias on the part of the subjects. Overall, the results suggest a privileged status for place information such that ambiguities in voicing and manner are reduced, thereby facilitating identification.

MODEL CHARACTERISTICS

In the Competitive Integration model (2) phoneme representations and feature representations compete with one another in such a way that the different featural distinctions differ only in the weights assigned to the connection between the featural nodes and the phoneme layer. The featural arrays included for competition are those corresponding to the consonants included in the previously described perception experiments: voiced-voiceless, labial-alveolar, and stop-fricative. Each feature array consists of a vector of 8 nodes corresponding to the featural distinctions of the phonemes [p, t, b, d, f, s, v, z]. Thus the target phoneme [p] is represented using the manner array [1 1 1 0 0 0 0 0] indicating that it is a stop consonant, the voicing array [1 1 0 0 1 1 0 0] indicating that it is voiceless, the place array [1 0 1 0 1 0 1 0] indicating that it is labial. The target phoneme will have high activation from all three featural arrays, resulting in the highest activation in the phoneme array. Other distracter phonemes which share features will also be activated, such that a phoneme which shares two features will have a higher activation than a phoneme which shares only one feature with the target phoneme.

Competition among phonemes and features takes place in a four step cycle using the Normalized Recurrence (3) competition algorithm. First, the activation of the featural arrays for a given target phoneme are normalized to a sum activation of one by dividing each value by the vector sum. Next, the
activation values from the featural arrays are multiplied by fixed weights (determined for each feature array, and summing to 1.0 across all features) and summed to give the probabilistic featural input for each node in the integration (phoneme). These activation values are then normalized across the phoneme array. Finally, the activation values in the phoneme array are fed back to the feature nodes, which allows the activation patterns of the different features to influence one another. The resulting activation levels in the phoneme array indicate the probability that each is identified in response to the target phoneme.

SIMULATION RESULTS

As shown in Table 1, the activation level for the target phoneme at the tenth timestep corresponds roughly to the percent correct identification for that phoneme predicted at the 20% duration condition in the perception experiment (correct identification scores at the 10% duration condition are approximately at chance level). Each additional timestep corresponds to an additional 10% of consonant duration, and the resulting activation for a target phoneme increases linearly as does correct identification with increasing consonant duration. That is, the most active phoneme increases in activation relative to the distracter phonemes as the number of timesteps increases.

TABLE 1. Comparing target phoneme activation values in the model to identification scores from different subject groups in the perception experiments.

<table>
<thead>
<tr>
<th>Timestep</th>
<th>Place</th>
<th>Manner</th>
<th>All 8</th>
<th>Duration Condition</th>
<th>Place</th>
<th>Manner</th>
<th>All 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>88</td>
<td>79</td>
<td>66</td>
<td>20%</td>
<td>80</td>
<td>58</td>
<td>47</td>
</tr>
<tr>
<td>11</td>
<td>93</td>
<td>86</td>
<td>75</td>
<td>30%</td>
<td>95</td>
<td>80</td>
<td>76</td>
</tr>
<tr>
<td>12</td>
<td>96</td>
<td>92</td>
<td>83</td>
<td>40%</td>
<td>95</td>
<td>89</td>
<td>83</td>
</tr>
<tr>
<td>13</td>
<td>98</td>
<td>96</td>
<td>89</td>
<td>50%</td>
<td>97</td>
<td>92</td>
<td>89</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100%</td>
<td>100</td>
<td>98</td>
<td>96</td>
</tr>
</tbody>
</table>

* Scores for the Place group were significantly higher than the other two groups. Scores for the Manner group were not significantly different from the All 8 group.

The model was most successful in simulating the results of the Place group relative to those subjects given all 8 consonants and the Manner group. The weight chosen for the place array was .4, with manner and voicing given a weight of .3 each. These weights result in a lower probability of place errors (thus higher activations for competing phonemes sharing place) which is consistent with confusion patterns from the perception experiments. The increase in correct identification across the duration conditions for the Place group, as compared to the Manner group, is achieved in the model by removing the weights for place from the calculations. This could be analogous to focusing entirely on voicing and manner information for the identification task.

The elimination of place information from the competition algorithm when the experiment set islabials or alveolars represents the privileged status of the place feature for the identification of these eight consonants evident from the perception experiment results. Thus the model can account for the general findings of the perception experiments, although the fit of the model to the data could be improved.

REFERENCES