COGNITIVE PSYCHOLOGY 21, 398–421 (1989)

Testing between the TRACE Model and the Fuzzy Logical Model of Speech Perception

DOMINIC W. MASSARO

University of California, Santa Cruz

The TRACE model of speech perception (McClelland & Elman, 1986) is contrasted with a fuzzy logical model of perception (FLMP) (Oden & Massaro, 1978). The central question is how the models account for the influence of multiple sources of information on perceptual judgment. Although the two models can make somewhat similar predictions, the assumptions underlying the models are fundamentally different. The TRACE model is built around the concept of interactive activation, whereas the FLMP is structured in terms of the integration of independent sources of information. The models are tested against test results of an experiment involving the independent manipulation of bottom-up and top-down sources of information. Using a signal detection framework, sensitivity and bias measures of performance can be computed. The TRACE model predicts that top-down influences from the word level influence sensitivity at the phoneme level, whereas the FLMP does not. The empirical results of a study involving the influence of phonological context and segmental information on the perceptual recognition of a speech segment are best described without any assumed changes in sensitivity. To date, not only is a mechanism of interactive activation not necessary to describe speech perception, it is shown to be wrong when instantiated in the TRACE model.

INTRODUCTION

Speech offers a viable domain for developing and testing models of perception, pattern recognition, and categorization. There has been a tradition of fairly elaborate models of speech perception (for recent reviews, see Jusczyk, 1986; Massaro, 1989). Several of the models make little direct contact with experimental results, however, and fall outside the mainstream of psychological inquiry. In addition, some models treat speech as a unique phenomenon and properties of the models have very little generality beyond speech itself. In contrast to these models, two

The research reported in this paper and the writing of the paper were supported, in part, by grants from the Public Health Service (PHS R01 NS 20314) and the graduate division of the University of California, Santa Cruz. The author thanks Jeff Elman for making the simulation program for TRACE available; Michael Cohen for eclectic assistance; Uli Frauenfelder for valuable discussions; and Anne Cutler, Jay McClelland, and an anonymous reviewer for comments on the paper. Requests for reprints should be sent to Dominic W. Massaro, Program in Experimental Psychology, University of California, Santa Cruz, CA 95064.

0010-0285/89 $7.50

other models both address experimental results directly and assume prototypical psychological processes that cut across several domains of inquiry. These two models are the TRACE model and the fuzzy logical model of perception (FLMP).

The goal of the present paper is to test between these two models, using the general research strategy proposed by Platt (1964) and developed more fully in the context of speech perception by Massaro (1987). Although the models share certain assumptions and make similar predictions in several experimental paradigms, they differ on one important attribute. The difference has to do with how multiple sources of information interact to jointly influence performance. The TRACE model is structured around the process of interactive activation (and conversely competition). Because of this process, the representation over time of one source of information is modified by another source of information. In contrast, the FLMP assumes that the representation over time of one source of information remains independent of another source of information.

Formulated within the theory of signal detection, the contrasting assumption of the two models leads to different predictions. Signal detection theory distinguishes between sensitivity (d') and bias (β). A d' measure can be computed to measure how sensitive a perceiver is to a given source of information (or in information–theoretic terms, how much information is transmitted by that source of information). The TRACE model predicts that the sensitivity to one source of information is influenced by another source. The FLMP predicts that the sensitivity to one source of information remains independent of the other. To test between these predictions, a top-down and a bottom-up source of information are independently varied in a speech identification task. The results are analyzed to determine whether sensitivity to the bottom-up source is modified by the top-down source. The two models are first described more fully before a more detailed description of the experimental procedure and data analysis is presented. Readers familiar with the models can proceed directly to the section *Testing between the Models*.

TRACE Model of Speech Perception

The TRACE model of speech perception (McClelland & Elman, 1986) is an interactive–activation model in which information processing occurs through excitatory and inhibitory interactions among a large number of simple processing units. These units are meant to represent the functional properties of neurons or neural networks. Three levels or sizes of units are used in TRACE: feature, phoneme, and word. Features activate phonemes which activate words, and activation of some units at a particular level inhibits other units at the same level. Given that multiple units at one

level simultaneously activate units at a higher level, the model provides a natural account for the integration of several bottom-up sources of information in speech perception. In addition, an important assumption of TRACE, an interactive-activation model, is that activation of higher-order units activates their lower-order units; for example, activation of a word containing a /b/ phoneme would activate that phoneme. Therefore, the model predicts that top-down sources at a higher level can also influence performance in addition to the infuence of bottom-up sources. These two properties of the model agree with the outcomes of several lines of research (Massaro, 1987; McClelland & Elman, 1986).

Fuzzy Logical Model of Perception

According to the FLMP, speech patterns are recognized in accordance with a general algorithm (Massaro, 1987; Oden & Massaro, 1978). The model assumes three operations in speech recognition: feature evaluation, feature integration, and decision. Continuously valued features are evaluated, integrated, and matched against prototype descriptions in memory, and an identification decision is made on the basis of the relative goodness of match of the stimulus information with the relevant prototype descriptions. The concept of fuzzy logic and how it has influenced the development of the model is discussed more fully in Massaro (1987).

Central to the FLMP are summary descriptions of the perceptual units of the language. These summary descriptions are called prototypes and they contain a conjunction of various properties called features. A prototype is a category and the features of the prototype correspond to the ideal values that an exemplar should have if it is a member of that category. The exact form of the representation of these properties is not known and may never be known. However, the memory of representation must be compatible with the sensory representation resulting from the transduction of the speech signals. Compatibility is necessary because the two representations must be related to one another. To recognize the syllable /ba/, the perceiver must be able to relate the information provided by the syllable itself to some memory of the category /ba/.

Prototypes are generated for the task at hand. In speech perception, for example, we might envision activation of all prototypes corresponding to the perceptual units of the language being spoken. For ease of exposition, consider a speech signal representing a single perceptual unit, such as the syllable /ba/. The sensory systems transduce the physical event and make available various sources of information called features. During the first operation in the model, the features are evaluated in terms of the prototypes in memory. For each feature and for each prototype, featural evaluation provides information about the degree to which the feature in the speech signal matches the featural value of the prototype.

Given the necessarily large variety of features, it is necessary to have a common metric representing the degree of match of each feature. The syllable /ba/, for example, might have visible featural information related to the closing of the lips and audible information corresponding to the second and third formant transitions (Massaro, 1987). These two features must share a common metric if they eventually are going to be related to one another. To serve this purpose, fuzzy truth values (Zadeh, 1965) are used because they provide a natural representation of the degree of match. Fuzzy truth values lie between zero and one, corresponding to a proposition being completely false and completely true. The value .5 corresponds to a completely ambiguous situation whereas .7 would be more true than false and so on. Fuzzy truth values, therefore, can represent not only continuous information, but also different kinds of information. Another advantage of fuzzy truth values is that information is couched in as quantitative form and, therefore, allows the natural development of a quantitative description of the phenomenon of interest.

Figure 1 gives a schematic diagram of the three operations of the FLMP. Feature evaluation provides the degree to which each feature in the syllable matches the corresponding feature in each prototype in memory. The goal, of course, is to determine the overall goodness of match of each prototype with the syllable. All of the features are capable of contributing to this process and the second operation of the model is called feature integration. That is, the features (actually the degrees of matches) corresponding to each prototype are combined (or conjoined in logical terms). The outcome of feature integration consists of the degree to which each prototype matches the syllable. In the model, all features contribute to the final value, but with the property that the least ambiguous features have the most impact on the outcome.

The third operation during speech recognition is decision. During this stage, the merit of each relevant prototype is evaluated relative to the sum of the merits of the other relevant prototypes. This relative goodness of match gives the proportion of times the syllable is identified as an instance of the prototype. The relative goodness of match could also be determined from a rating judgment indicating the degree to which the syllable matches the category. The decision operation is modeled after Luce's (1959) choice rule. In pandemonium-like terms (Selfridge, 1959), we might say that it is not how loud some demon is shouting but rather the relative

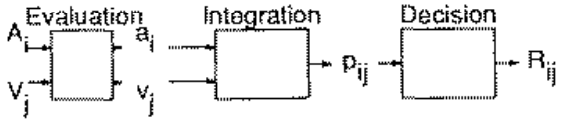

FIG. 1. Schematic diagram of the three operations assumed by the FLMP.

loudness of that demon in the crowd of relevant demons. An important prediction of the model is that one cue has its greatest effect when a second cue is at its most ambiguous level. Thus, the most informative cue has the greatest impact on the judgment.

Testing between the Models

The TRACE model and FLMP make similar predictions and definitive tests between the models might appear to be difficult, if not impossible (Massaro, 1987; McClelland & Elman, 1986). However, the two models make fundamentally different assumptions that can be tested hy a fine-grained analysis of the predictions of the models against empirical results. The predictions of the two models differ from one another when the models are conceptualized within the theory of signal detectability (TSD). A distinction is made between sensitivity and bias in TSD. Within this

That is, the question of interest is to what extent the top-down effects are reflected in sensitivity or bias (d' or β). The FLMP predicts no systematic effects of top-down context on sensitivity at the bottom-up level (Massaro, 1979). In contrast to the predictions of the FLMP, the TRACE model accounts for the top-down effects of phonological constraints by assuming interactive activation between the word and phoneme levels. Bottom-up activation of the phonemes activates words, which in turn, activate the phonemes that make them up. Interactive activation appropriately describes this model because it is clearly an interaction betwen the two levels that is postulated. The amount of bottom-up activation influences the amount of top-down activation, which then modifies the bottom-up activation, and so on.

More generally, other properties of TRACE might produce sensitivity differences. As noted by Massaro (1987) and J. L. McClelland (personal communication), changes in sensitivity also occur when units within a level interact with one another. The concern of the present test, however, is to what extent TRACE's account of context effects necessarily results in sensitivity differences. That is, an important empirical and theoretical question is whether top-down context produces sensitivity differences at a bottom-up level. Our analysis asks whether phonological context produces sensitivity differences, as well as whether TRACE and the FLMP predict sensitivity differences due to context.

The concept of sensitivity is tied to the discriminability of two different stimulus events, whereas the concept of bias refers to the direction of the perceptual judgment. As an example, consider the top-down effect evaluated in the present paper. A set of syllables along a /li/–/ri/ continuum is factorially combined with different initial consonant contexts /t/, /p/, or

/s/. Subjects asked to identify whether /li/ or /ri/ is present in each test syllable are influenced not only by the information specifying /l/ or /r/, but also by the initial consonant context (Massaro & Cohen, 1983). For example, subjects report /r/ more often in the context /t_i/ than in the context /s_i/. Without the appropriate experimental design and data analysis, we do not know if this result is due to sensitivity or bias. Usually, catch trials are included in the signal detection task, and this has been profitably exploited in a recent study of episodic priming (Ratcliff, McKoon, & Verwoerd, in press). Another technique is to have a continuum of stimulus conditions and to analyze performance across this continuum (Braida & Durlach, 1972). This analysis was used successfully by Massaro (1979) in arguing against top-down sensitivity effects in written word recognition. Sensitivity effects would be reflected in changes in the discriminability of two adjacent levels along the liquid continuum as a function of context. Bias would be reflected in a change in overall response probability as a function of context. *Nonindependence* refers to an effect of context on sensitivity and *independence* refers to top-down effects having an influence only on bias (Massaro, 1979, 1988).

Before evaluating the TRACE model and the FLMP, it is necessary to describe the concepts of sensitivity and bias in the present framework and experimental tasks. Consider a set of syllables along a /li/–/ri/ continuum factorially combined with different initial consonant contexts. Assume three different syllables /li/, /Li/, and /ri/ placed after the initial consonants /t/, /p/, or /s/. The syllable /li/ is a better /li/ than /ri/, the syllable /Li/ is halfway between /li/ and /ri/, and the syllable /ri/ is a better /ri/ than /li/. Each adjacent pair of syllables in a given context can be viewed as the two types of trials in a signal detection task. Sensitivity and bias are indexed by different dependent measures. If a subject responds /l/ or /r/ on each trial, a measure of sensitivity is reflected in the differential responding to the two types of trials. To the extent the subject responds /l/ to one member of the adjacent pair and /r/ to the other member, sensitivity is high. Using the concept of information within information theory, the subject transmits more information to the extent that there is differential responding to the two stimulus alternatives. The overall probability of responding with a given alternative, independently of the stimulus that was presented, reflects bias within the framework of signal detection. To the extent the subject responds with only one response alternative, there is a bias towards that alternative. We now proceed to test the empirical question whether top-down influences of phonological context change sensitivity or just bias in human listeners.

Experimental Test of Top-Down Effects on Sensitivity

It is claimed that the concept of interactive activation, as implemented

in TRACE, should predict sensitivity effects rather than just bias. Take as an example a liquid phoneme presented after the initial consonant /t/. The liquid would activate both /l/ and /r/ phonemes to some degree; the difference in activation would be a function of the test phoneme. There are many words that begin with /tr/ but none that begin with /tl/ and, therefore, there would be more top-down activation for /r/ than for /l/. Top-down activation of /r/ would add to the activation of the /r/ phoneme at the phoneme level. What is important for our purposes is that the amount of top-down activation is positively related to the amount of bottom-up activation. Now consider the top-down effects for the two adjacent stimuli along the /l/–/r/ continuum. Both of these test syllables activate the phonemes to some degree, and the phonemes activate words, which then activate phonemes. However, the two adjacent syllables have different patterns of bottom-up activation because they are different syllables. The bottom-up activation differs for the two syllables and, therefore, the top-down activation must also differ. The difference in the top-down activation contributes to differences in activation at the phoneme level. This nonlinear relationship between top-down and bottom-up activation should be reflected in sensitivity differences as a function of top-down context.

The FLMP, on the other hand, predicts no effects of context on sensitivity. Context is treated as an additional independent source of information that is integrated with the bottom-up source. Thus, the goodness-of-fit of the FLMP provides a measure of sensitivity effects: a good fit indicates no sensitivity effects. To the extent that the FLMP gives a good description of empirical results and the TRACE model can be shown to predict sensitivity effects, then there is evidence against the interactive-activation assumption of TRACE. In order to carry out this test between the FLMP and the TRACE models, subjects were asked to identify a liquid consonant in different phonological contexts. Each speech sound was a consonant cluster syllable beginning with one of the three consonants /p/, /t/, or /s/ followed by a liquid consonant ranging (in five levels) from /l/ to /r/, followed by the vowel /i/. Therefore, there were 15 test stimuli created from the factorial combination of five stimulus levels combined with three initial-consonant contexts. Elementary school children were instructed to listen to each test syllable and to respond whether they heard /li/ or /ri/.

METHOD

Subjects

Eight subjects were tested three sessions each. The subjects were fourth-graders recruited from the Madison, Wisconsin School District. Each subject was paid $5.00 for participation.

Apparatus

All speech sounds were produced on-line during the experiment by a formant series resonator speech synthesizer (FONEMA-OVE-IIId) controlled by a DEC PDP-8/L computer (Cohen & Massaro, 1976). Segment durations were always multiples of 8 ms. The stimuli were defined as a series of parameter vectors, each specifying a target value and transition time, with linear, positively, or negatively accelerated transitions. Intermediate values were computed and fed to the synthesizer at 8-ms intervals. The output of the synthesizer was amplified (McIntosh MC-50), bandpass filtered 20 Hz–10 kHz (KROHN-HITE 3500R), and presented over headphones (Koss Pro-4AA) at a comfortable listening level (about 72 dB-SPL-A). Four subjects were tested simultaneously in separate sound attenuated rooms.

Stimuli

The stimuli were slight modifications of the syllables used in the Massaro and Cohen (1983) study, which also contains details of the speech synthesis. Each speech sound was a consonant cluster syllable beginning with one of the three consonants /p/, /t/, or /s/ followed by a liquid consonant ranging (in five levels) from /l/ to /r/, followed by the vowel /i/. The syllables were synthesized with the constraint that the bottom-up information specifying the liquid was identical in the three different top-down contexts. The five different levels along the /l/–/r/ continuum differed with respect to the third formant (F_3) of the liquid. The initial values of F_3 at the onset of the liquid were 2933, 2614, 2263, 1958, and 1695 Hz, from the sound most like /l/ to the sound most like /r/.

Procedure

In order to familiarize the subjects with synthetic speech, they first listened to the entire set of stimuli twice. The sounds were presented in a fixed order with the five levels of F_3 defining the /l/–/r/ continuum as the fastest moving variable. The subjects were told that these sounds were a subset of the sounds involved in the experiment and that the stimulus order in the experiment was entirely random.

On each test trial, a syllable was randomly selected without replacement from the set of 15 syllables generated from the factorial combination of the three initial consonants and the five F_3 levels of the following liquid. The subjects responded with one of two buttons labeled L and R. The computer waited until each subject responded. The response interval averaged between 1 and 2 s. An additional 1-s interval intervened before the start of the next trial. The subjects were told that there were three possible consonants in initial position followed by either /l/ or /r/ followed by /i/. Their task was to identify the second segment on the basis of what they heard. They were told that there was no correct response and simply to make the best judgment they could. The subjects were then given a practice session of 15 trials before the first test session. Each test session of 150 trials consisted of 10 blocks of the 15 stimuli. There were three test sessions giving a total of 30 observations for each subject on each of the 15 test stimuli.

RESULTS

Figure 2 gives the average probability of an /r/ response as a function of the two factors. As can be seen in the figure, both factors had a strong effect. The probability of an /r/ response increases systematically with decreases in the F_3 transition, $F(4, 28) = 38.69$, $p < .001$. Phonological context also had a significant effect on the judgments, $F(2, 14) = 16.43$,

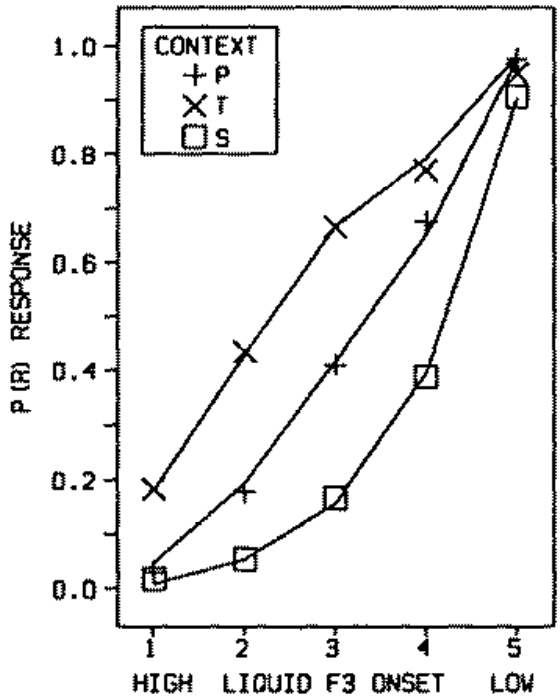

FIG. 2. Observed (points) and predicted (lines) probability of an /r/ identification as a function of the F_3 transition onset of the liquid; the initial consonant is the curve parameter. The predictions are for the FLMP.

$p < .001$. Finally, the significant interaction reflected the fact that the phonological context effect was greatest when the information about the liquid was ambiguous, $F(8, 56) = 8.25$, $p < .001$.

In terms of the signal detection framework, phonological context had a strong effect on bias. Subjects responded /r/ more often given the context /t/ than given the context /p/. Similarly, there were fewer /r/ responses given the context /s/ than given the context /p/. The issue of sensitivity effects will be addressed in the context of the FLMP and TRACE predictions of the results.

Test of the FLMP

A critical assumption of the FLMP is that the featural information from the liquid and the phonological context provide *independent* sources of information. It is assumed that subjects adopt the prototypes R and L in the task, and evaluate and integrate the two sources of information with respect to these prototypes. The featural information supporting the R prototype can be represented by the truth value t_i, where the subscript i indicates that t_i changes only with the F_3 transition. For the /l/–/r/ identification, t_i specifies how much the critical F_3 transition feature supports the prototype R. This value is expected to increase as the onset frequency of the F_3 transition is decreased. However, t_i is assumed to be indepen-

dent of the phonological context. With just two alternatives along the continuum, it can further be assumed that the featural information supporting the L prototype is the compliment of t_i. Thus, the support for L is simply one minus the support for R. Therefore, if t_i specifies the support for R given by the F_3 transition, then $(1 - t_i)$ specifies the support for L given by that same transition.

The phonological context also provides independent evidence for R and L. The value c_j represents how much the context supports the prototype R. The subscript j indicates that c_j changes only with changes in phonological context. The value of c_j should be large when /r/ is admissible and small when /r/ is not admissible. Analogous to the treatment of the featural information, the degree to which the phonological context supports the prototype L is indexed by $(1 - c_j)$.

The listener is assumed to have two independent sources of information. The total degree of match with the prototypes R and L is determined by integrating these two sources. Feature integration involves a multiplicative combination of the two truth values. Therefore, the degree of match to R and L for a given syllable can be represented by

$$R = (t_i \times c_j) \tag{1}$$

$$L = [(1 - t_i) \times (1 - c_j)]. \tag{2}$$

The decision operation maps these outcomes into responses by way of Luce's choice rule. The probability of an /r/ response given test stimulus S_{ij} is predicted to be

$$P(r|S_{ij}) = \frac{t_i c_j}{t_i c_j + (1 - t_i)(1 - c_j)}. \tag{3}$$

The FLMP was fit to the proportion of /r/ identifications as a function of the F_3 of the liquid and the initial consonant context. Five levels of the liquid times three phonological contexts gives 15 independent data points to be predicted. In order to predict the results quantitatively, the model requires the estimation of eight free parameters. Five values of t_i are required for the five levels of the F_3 transition of the liquid. Unique c_j values are required for each of the three different initial consonant contexts. Fitting the model to the observed data, therefore, requires the estimation of $5 + 3 = 8$ parameters.

The model was fit to each of the children's results individually and also to the average results. The criterion of best fit was based on the root mean square deviation (RMSD) or the square root of the average squared deviation between predicted and observed points. The RMSD values ranged between .0157 and .0719 and averaged .0312. The RMSD for the fit of the average subject was .0264. The lines in Fig. 2 give the average predictions

of the FLMP. The estimated truth values averaged 0.0308, 0.1482, 0.3681, 0.6381, and 0.9752 for the five levels going from /l/ to /r/ along the /l/–/r/ continuum. The estimated truth values averaged 0.1780, 0.5672, and 0.7581 for the phonological contexts /s/, /p/, and /t/, respectively. These parameter estimates of the model are meaningful. The t_i values, representing the degree of match with the prototype R, increase systematically with decreases in the starting frequency of F_3. The c_j values change systematically with phonological context; the degree of match with R given by the context is much larger for initial /t/ than for initial /s/.

Simulations of TRACE

In contrast to the FLMP, the TRACE model cannot be tested directly against the results. It is necessary to simulate the experiment with TRACE and to compare the simulation with the observed results. Given this method, the goal is to test fundamental properties of TRACE rather than specific results that are primarily a consequence of the details of the implementation. Therefore, differences due to the makeup of the lexicon and specific parameter values are less important than systematic properties of the predictions. Within the current architecture of the TRACE model, the word level appears to play a fundamental role in the discrimination of alternatives at the phoneme level. The most straightfoward test of this observation is to simulate results with the standard TRACE model and compare this simulation with simulations in which the top-down connections from the word level to the phoneme level are eliminated. This contrast tests whether the top-down activation modifies the discriminability between alternatives at the phoneme level.

Two simulations of TRACE were run with and without top-down connections. If top-down activation modifies sensitivity, then we should get differences in a sensitivity measure with these two simulations. The results of the simulations were analyzed for sensitivity and bias effects. The first set of simulations used the lexicon, the input feature values, and the parameter values given in McClelland and Elman (1986, Tables 1 and 3). Three levels of information about the liquid (l, r, and L) were used as three levels of input information. The phoneme /L/ refers to the class of liquid phonemes; thus, the diffuse and acute feature specifications for this ambiguous phoneme are neutralized at intermediate feature values. The other feature specifications are identical for those given for the liquids /l/ and /r/. The input /L/ activates the two liquids more than the other phonemes, but activates /l/ and /r/ to the same degree. These three liquids were placed after initial /t/, /p/, and /s/ contexts and followed by the vowel /i/. The simulations, therefore, involved these nine stimulus conditions tested with and without top-down connections.

The TRACE simulation is completely deterministic; a single run is

sufficient for each of the three conditions. The activation of the /l/ and /r/ units at the phoneme level occurred primarily at the 12th time slice of the trace, and these values tended to asymptote around the 54th cycle of the simulation run. Therefore, we will take the activations at the 12th time slice after the 54th cycle as the predictions of the model. The activations of the /r/ and /l/ units as a function of the three syllables in the three phonological contexts are shown in Table 1. The first critical variable of interest is whether top-down connections are present.

The overall level of activation is fairly independent of whether top-down connections are present. Table 1 shows that top-down connections mainly modify the relative activation of the /l/ and /r/ phoneme units in TRACE rather than modify the overall level of activation. As expected, the relative activations of /l/ and /r/ phonemes appear to differ as a function of top-down connections. For example, given the context /t/ and the liquid /l/, /r/ is more active than the phoneme /l/ when top-down connections are present whereas /l/ is more active than /r/ when there are no top-down connections. These activations cannot be taken as direct measures of sensitivity or bias, however. In order to assess whether top-down connections modify sensitivity or bias in the TRACE model, it is necessary to map these activation levels into predicted responses.

Before evaluating the predicted responses for sensitivity and bias effects, one aspect of the activations in Table 1 deserves a brief comment. The input phoneme /L/ is halfway between /l/ and /r/ and, therefore, might be expected to activate the /l/ and /r/ units equally when there are no top-down connections. This result is, in fact, the case with the contexts /t/ and /s/. With the context /p/, however, /r/ is activated more than /l/. The

TABLE 1
The TRACE Activations of the /r/ and /l/ Phoneme Units as a Function of the Bottom-Up Information /l/, /L/, or /r/; and the Top-Down Information of /t/, /p/, or /s/ in Initial Position

		Test phoneme					
		/l/		/L/		/r/	
Top-down connections: Context	Unit	Present	Absent	Present	Absent	Present	Absent
/t/	/r/	.46	.20	.57	.34	.66	.52
	/l/	.39	.52	.12	.34	.00	.20
/p/	/r/	.31	.28	.56	.48	.65	.59
	/l/	.52	.48	.11	.22	.00	.12
/s/	/r/	.09	.20	.23	.33	.55	.51
	/l/	.59	.51	.34	.33	.17	.19

Note. Top-down connections are either present or absent.

same input from /L/ can produce different patterns of activation in TRACE, even without top-down connections, because of differences in the initial consonant. The phoneme detectors span overlapping slices of time, so that the /l/ and /r/ detectors have some input from the initial consonant. It follows that different initial consonants could give different patterns of activation for /l/ and /r/, even without top-down connections and without the connections between adjacent time slices that are assumed in TRACE I (Elman & McClelland, 1986). The initial consonant /p/ supports /r/ over /l/ more than do the initial consonants /t/ or /s/ because of their feature values for acute (McClelland & Elman, 1986, Table 1). The value for acute is .2 for /r/ and .4 for /l/. The acute value for /p/ is also .2, whereas the acute value is .7 and .8 for /t/ and /s/, respectively. The phoneme /r/ receives strong bottom-up support from initial /p/ and, thus, shows more activation than /l/.

The computation of sensitivity and bias within the framework of signal detection theory is based on the proportion of responses. However, the proportion of /l/ and /r/ responses are not given by the activations directly. McClelland & Elman (1986) assume that the activation a_i of a phoneme unit is transformed by an exponential function into a strength value S_i,

$$S_i = e^{ka_i}. \tag{4}$$

The strength value S_i represents the strength of alternative i. The probability of choosing a particular alternative, $P(R_i)$, is based on the activations of all relevant alternatives, as described by Luce's (1959) choice rule,

$$P(R_i) = \frac{S_i}{\Sigma}, \tag{5}$$

where Σ is equal to the sum of the strengths of all relevant phonemes, derived in the manner illustrated for alternative i. The activation values in Table 1 were translated into strength values by the exponential function given by Eq. (4). The constant k was set equal to 5. The probability of an /r/ judgment was determined from the strength values using Eq. (5).

The probability of an /r/ response for three levels of the liquid as a function of whether top-down connections are present is shown in Table 2. As can be seen in Table 2, TRACE predicts that top-down connections modify bias. The probability of an /r/ response given the context /t/ is greater with than without top-down connections. Activation from units representing words beginning with /tr/ bias activation at the phoneme level toward /r/ rather than /l/.

To determine if the presence of top-down connections modifies sensitivity, the proportions were translated in d' values as in Braida and

TABLE 2
The Probability of an /r/ Response as a Function of the Bottom-Up Information, Top-Down Information, and whether Top-down Connections Are Present or Absent

	Test phoneme					
	/l/		/L/		/r/	
Top-down connections: Context	Present	Absent	Present	Absent	Present	Absent
/t/	.59	.17	.90	.50	.96	.83
/p/	.26	.27	.90	.79	.96	.91
/s/	.08	.18	.37	.50	.87	.83

Note. Predictions of the TRACE model.

Durlach (1972) and Massaro (1979). The probabilities of responding /r/ are transformed to z scores. The d' between two adjacent levels along the /l/–/r/ continuum is simply the positive difference between the respective z scores. Given response probabilities of .05 and .15, for example, the respective z scores would be -1.65 and -1.04. The corresponding d' would be .61. A d' value was computed for each of the two pairs of adjacent levels along the /l/–/r/ continuum. Two d' values were computed for each of the three different phonological contexts, with and without top-down connections. These values are given in Table 3.

The d' values reveal whether the sensitivity changes as a function of the presence or absence of top-down information. There are six possible comparisons: two adjacent pairs of inputs times three contexts. As can be seen in the table, three of the six d' values with top-down connections differed from those without top-down connections. For the context /t/, the discriminability of /L/–/r/ was about twice as large without top-down connections than with top-down connections. In contrast, the context /s/

TABLE 3
The d' Values as a Function of the Bottom-Up Information /l/–/L/ or /L/–/r/, the Top-Down Information of /t/, /p/, or /s/ in Initial Position, and whether Top-down Connections are Present or Absent

	Test phonemes			
	/l/–/L/		/L/–/r/	
Top-down connections: Context	Present	Absent	Present	Absent
t	1.09	0.96	0.50	0.96
p	1.95	1.41	0.47	0.57
s	1.09	0.93	1.47	0.96

Note. Predictions of the TRACE model when k is equal to 5.

produced a significantly larger sensitivity between /L/–/r/ with top-down connections than without top-down connections. In terms of the present analysis, the presence of top-down connections influences sensitivity. Thus the simulation is consistent with the intuition that interactive activation between the word and phoneme levels in TRACE produces sensitivity changes at the phoneme level (Massaro, 1988).

Having demonstrated sensitivity effects in TRACE as a function of presence or absence of top-down connections, the next charge is to test for whether the type of context influences sensitivity when top-down connections are present. Top-down connections in TRACE should modify sensitivity at the phoneme level differentially as a function of different top-down contexts. The liquid phoneme presented after the initial consonant /t/ should activate words that begin with /tr/. Top-down activation from these words to the phonemes that make them up should activate /r/ and not /l/. This activation will, in turn, activate words with /r/ in second position, and so on. The question of interest is whether this context will influence sensitivity. If it does, then the listener's ability to discriminate the two adjacent levels along the liquid continuum should be modified relative to some other context condition.

The d' values in Table 3 reveal how the sensitivity changes as a function of top-down information. Consider the differences observed between the /t/ and /s/ contexts. The discriminability of two successive levels along the /l/–/r/ continuum appears to be a function of relative strength of the context supporting one alternative and the bottom-up information supporting that alternative. The context /t/ supports /r/ over /l/. When the two adjacent levels are on the /l/ side of the continuum, discriminability is better than when the two adjacent levels are on the /r/ side of the continuum (1.09 vs 0.50). The analogous result holds for the context /s/. The context /s/ supports /l/ over /r/. When the two adjacent levels are on the /r/ side of the continuum, discriminability is better than when the two adjacent levels are on the /l/ side of the continuum (1.47 vs 1.09).

At first glance, the effect of the context /p/ seems strange because there is a strong bias for /r/ rather than for /l/. One might have expected very little difference because initial /p/ activates both /pr/ and /pl/ words. However, the makeup of the lexicon used in the simulation favored /r/ much more than /l/. In this case, the /p/ context functions more like the /t/ context and enhances the discrimination of the two adjacent levels on the /l/ end of the continuum relative to the /r/ end of the continuum (1.95 vs 0.47).

It is of interest whether TRACE predicts sensitivity differences for other values of the constant k that maps activation into strength values. Eight values of k were used, giving a total of eight simulated subjects. The values of k were 0.5, 1, 2, 3.5, 5, 7.5, 10, and 15. The critical result of

interest is whether the context influenced sensitivity or just bias. For each simulated subject, a d' value was computed for each of the two pairs (/l/–/L/ and /L/–/r/) of adjacent levels along the /l/–/r/ continuum. These two d' values were computed for each of the three different phonological contexts /t/, /p/, and /s/. Table 4 gives these d' values as a function of the three contexts and two pairs of levels. The d' values for the two pairs of levels along the /l/–/r/ continuum differed from one another for all values of k. For each value of k, there was a large effect of context and the nature of the context effect interacted with the two levels along the continuum. Beginning at zero and adding these successive d' distances gives a cumulative d' discrimination function. Figure 3 plots the average predicted cumulative d' values as a function of context and level along the place continuum. These values were computed from the average proportions of /r/ judgments. As can be seen in the figure, there are systematic effects of context on sensitivity as measured by d'.

Given that TRACE has been shown to predict d' differences, it is of interest whether phonological context influenced sensitivity in the same manner in human subjects. Accordingly, the same d' analysis was performed on the results obtained with the children subjects. To make the analysis directly comparable to the one with the simulated subjects from TRACE, only the middle three levels of the /l/–/r/ continuum were included. Excluding the endpoint stimuli was also reasonable because many of the proportions were 0 or 1, precluding a direct z score transformation. An analysis of variance was carried out on these d' values with the three contexts and two pairs of levels as factors. There was no effect of context

TABLE 4

The d' values, Representing TRACE's Predicted Discriminability between Two Adjacent Levels along the /l/–/r/ Continuum, for /l/–/L/ and /L/–/r/ as a Function of Context /t/, /p/, or /s/ when Top-Down Connections Are Present

	Place /l/–/L/			Place /L/–/r/		
	Context			Context		
k value	/t/	/p/	/s/	/t/	/p/	/s/
0.5	0.12	0.21	0.12	0.07	0.06	0.15
1.0	0.24	0.41	0.24	0.13	0.12	0.31
2.0	0.47	0.82	0.48	0.25	0.24	0.61
3.5	0.80	1.40	0.80	0.39	0.37	1.05
5.0	1.09	1.95	1.09	0.50	0.47	1.47
7.5	1.51	2.79	1.49	0.62	0.59	2.11
10.0	1.86	3.52	1.80	0.71	0.68	2.69
15.0	2.40	4.78	2.27	0.85	0.81	3.70

Note. The k values correspond to the constant in Eq. (4) that translates activations into strength values.

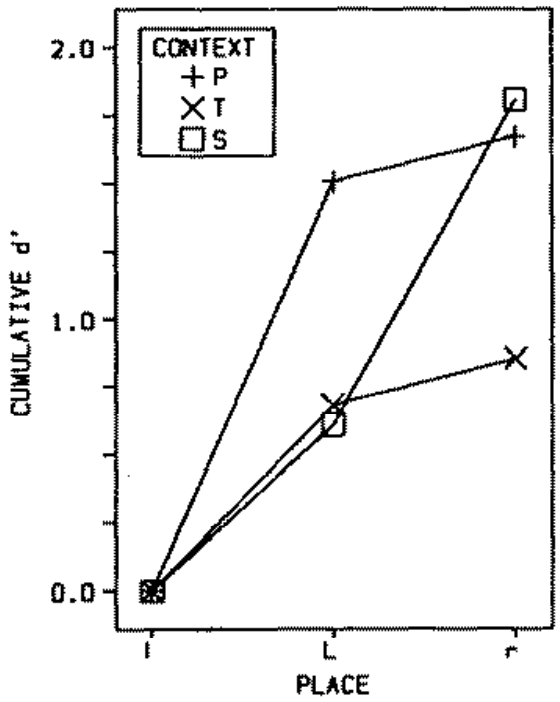

FIG. 3. Cumulative d' values, based on the average of the proportions generated from the eight k values, as a function of the place of the liquid; the initial consonant is the curve parameter. Predictions of the TRACE model.

and context did not interact with the two levels along the continuum ($p > .25$). Figure 4 plots the cumulative d' values as a function of context and level. These values were computed from the average proportions of /r/ judgments. In agreement with the statistical test, there is no systematic effect of context on sensitivity as measured by d'. In contrast to the predictions of TRACE, and consistent with the predictions of the FLMP, there is no systematic effect of phonological context on sensitivity.

An advocate of TRACE might remark that the predictions in Fig. 3 and the empirical observations in Fig. 4 are not all that different. One might say that the results differ significantly only with respect to the context /p/ at the middle level of the /l/–/r/ continuum. However, the more important point is that TRACE predicts *systematic* effects of context on sensitivity whereas no systematic effects are observed in the empirical results. As noted in the discussion of Table 3, TRACE predicts differences in discriminability as a function of the relative strength of the context and the bottom-up information. Discriminability is best when the stimulus information is on the opposite end of the continuum from the support given by the context. Given the context /t/, the two levels at the /l/ end of the continuum are better discriminated than the two levels at the /r/ end of the continuum. Given the context /s/, the two levels at the /r/ end of the continuum are better discriminated than the two levels at the /l/ end of the continuum.

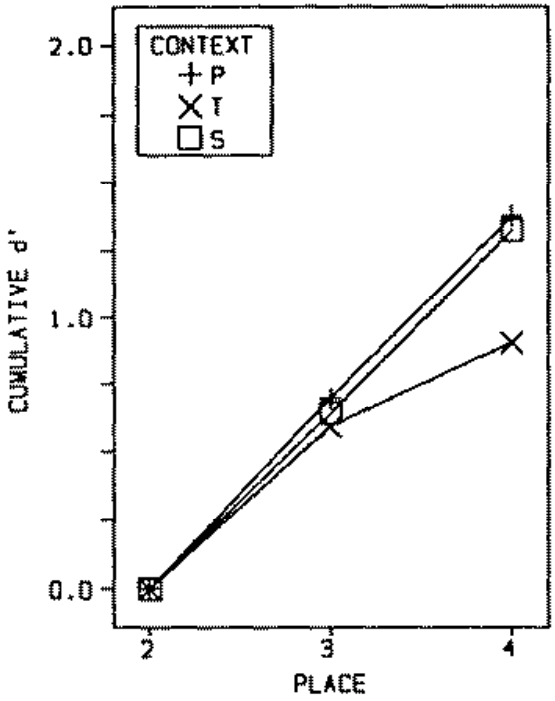

FIG. 4. Cumulative d' values as a function of the F_3 transition onset of the liquid; the initial consonant is the curve parameter. The d' values were computed from the average results of the eight human subjects (see Fig. 2).

To provide another test of TRACE, another set of simulations exactly analogous to the previous set was carried out. To evaluate TRACE when more of the lexicon (or a larger lexicon) is activated, the vowel context was changed from /i/ to /V/. The vowel /V/ is an input that activates all of the vowels in TRACE to the same degree (analogous to the liquid /L/). All details of the simulations were identical to the first set of simulations with the vowel /i/. Tables 5, 6, and 7 give the results.

The results with the vowel /V/ replicate exactly the effects found with the vowel /i/. Table 7 shows that the d' values with top-down connections differed from those without top-down connections. The results of eight simulated subjects were generated using the eight k values given earlier. Figure 5 gives the average predicted cumulative d' values. As can be seen in a comparison between Fig. 3 and 5, the effects are exactly analogous to those found with the vowel context /i/.

In summary, context had a significant influence on the relative sensitivity to changes along the liquid continuum as a function of phonological context. Consistent with the logic of interactive activation, top-down context modifies the relative discriminability of bottom-up information.

Hypothetical Results Based on Independence

To verify the signal detection analysis and to explicate the FLMP, some

TABLE 5
The Activations of the /r/ and /l/ Phonemes in TRACE as a Function of the Bottom-Up Information /l/, /L/, or /r/; and the Top-Down Information of /t/, /p/, or /s/ in Initial Position

		Test phoneme					
		/l/		/L/		/r/	
Top-down connections: Context	Unit	Present	Absent	Present	Absent	Present	Absent
/t/	/r/	.46	.19	.58	.36	.67	.55
	/l/	.45	.55	.15	.36	.00	.20
/p/	/r/	.29	.28	.58	.50	.65	.61
	/l/	.56	.52	.13	.25	.00	.11
/s/	/r/	.06	.19	.32	.36	.58	.55
	/l/	.61	.55	.42	.36	.12	.19

Note. Top-down connections are either present or absent. Predictions for the vowel /V/ that activates all of the vowels in TRACE to the same degree.

hypothetical results based on the assumption that context does not modify sensitivity were generated. We assumed that the two d' values were both 1.037 for all three contexts. These d' values generate the probabilities shown in Table 8. The TRACE model was not designed to fit results directly, which seems to preclude testing whether the model can predict these results generated from independence. On the other hand, the FLMP can be fit to results directly. The model can be fit to both the independence data and the predictions generated by the TRACE model. If the FLMP does not predict sensitivity effects, the FLMP should give a good description of the independence predictions. In agreement with this expectation, the RMSD was .001 for the fit of the independence data.

TABLE 6
The Probability of an /r/ Response as a Function of the Bottom-Up Information, Top-Down Information, and whether Top-Down Connections Are Present

	Test phoneme					
	/l/		/L/		/r/	
Top-down connections: Context	Present	Absent	Present	Absent	Present	Absent
/t/	.51	.14	.90	.50	.97	.85
/p/	.21	.45	.90	.78	.96	.92
/s/	.06	.14	.38	.50	.91	.86

Note. Predictions of the TRACE model for the vowel /V/ that activates all of the vowels in TRACE to the same degree.

TABLE 7
The d' Values as a Function of the Bottom-Up Information /l/–/L/ or /L/–/r/, the Top-Down Information of /t/, /p/, or /s/ in Initial Position, and whether Top-Down Connections Are Present or Absent

	Test phonemes			
	/l/–/L/		/L/–/r/	
Top-down connections: Context	Present	Absent	Present	Absent
/t/	1.23	1.07	0.57	1.05
/p/	2.13	1.50	0.47	0.67
/s/	1.24	1.07	1.65	1.07

Note. Predictions of the TRACE model for the vowel /V/ that activates all of the vowels in TRACE to the same degree.

A reader might conjecture that the free parameters in the FLMP allow the model to predict either no sensitivity differences or sensitivity differences depending on the nature of the results to be described. This conjecture can be shown to be false both formally and empirically. The FLMP is formally identical to Bayes theorem, which maintains complete

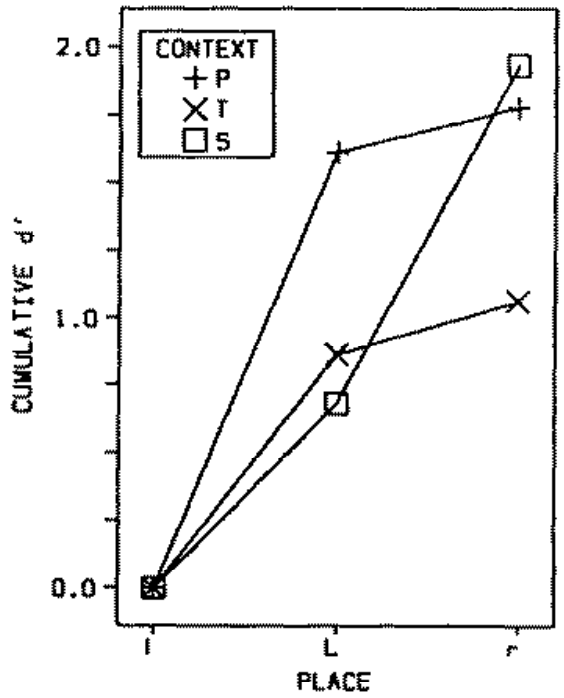

FIG. 5. Cumulative d' values, based on the average of the proportions generated from the eight k values, as a function of the place of the liquid; the initial consonant is the curve parameter. Predictions of the TRACE model for the vowel /V/ that activates all of the vowels in TRACE to the same degree.

TABLE 8
The Probability of an /r/ Response as a Function of the Bottom-Up Information /l/, /L/, or /r/; and the Top-Down Information of /t/, /p/, or /s/ in Initial Position

Context	Test phoneme		
	/l/	/L/	/r/
/t/	.25	.64	.92
/p/	.15	.50	.85
/s/	.05	.27	.67

Note. Hypothetical results generated with the constraint that the context does not influence sensitivity.

independence among the different sources of information (Massaro, 1987). To illustrate that the FLMP does a poor job of describing sensitivity differences, the model was fit to the hypothetical results of the eight simulated subjects of TRACE generated by assuming eight different k values. The FLMP gives a relatively poor description of these results with an average RMSD value of .0646. Thus, the FLMP gives a poor fit of results generated from an interactive-activation model in which context influences sensitivity and gives an accurate description of results generated from a signal-detection model based on independence.

Although we have equated the FLMP with the TSD prediction based on normal distributions, this assumption is not strictly true. The FLMP predicts exact independence of context on adjacent levels of a continuum when the underlying distribution is logistic rather than normal (Massaro & Cohen, 1987). The logistic distribution is very close to the normal, but is not identical to it. The analyses of the FLMP and TRACE predictions and the actual results were repeated with the assumption of an underlying logistic distribution. Equivalent results were found, and the d' analysis is presented here because of its greater familiarity to the reader.

Relationship to Other Research

There are very few experiments addressing sensitivity and bias effects in language processing in the literature. As an exception, Samuel (1981) employed a signal detection framework in his study of phonemic restoration (Warren, 1970). In the original type of study, a phoneme in a word is removed and replaced with some stimulus, such as a tone or white noise. Subjects have difficulty indicating what phoneme is missing. Failure to spot the missing phoneme could be a sensitivity effect or a bias effect. Samuel addressed this issue by creating signal and noise trials. Signal trials contained the original phoneme with superimposed white noise. Noise trials replaced the original phoneme with the same white noise. Subjects were asked to indicate whether the original phoneme was

present. Sensitivity is reflected in the degree to which the two types of trials can be discriminated. Bias would be reflected in the overall likelihood of saying that the original phoneme is present.

To evaluate the top-down effects of lexical constraints, Samuel compared performance on test words relative to performance on the phoneme segments presented in isolation. Subjects were more likely to respond that the phoneme was present in the word context than in the segment context. This result reveals a bias. In addition, subjects discriminated the signal from the noise trials much better in the segment context than in the word context. The d' values averaged about two or three times larger for the segment context than in the word context. In contrast to the present results, there appears to be a large effect of top-down context on sensitivity. However, the segment versus word comparison confounds stimulus contributions with top-down contributions. An isolated segment has other advantages over a segment presented in a word. Forward and backward masking degrades the perceptual quality of a segment presented in a word relative to being presented alone. In addition, the word context might provide co-articulatory information about the critical phoneme, which would not be available in the isolated segment.

Samuel carried out a second study that should have overcome these possible limitations in comparing words and segments. In this study, a word context was compared to a pseudoword context. Supporting the argument for stimulus differences between the words and segments, the d' value for unprimed pseudowords dropped below those for primed words. However, there was an advantage of primed psuedowords over primed words, which Samuel interpreted as a sensitivity effect. Stimulus confounding mught also be responsible for this difference. Natural speech was used and, therefore, the equivalence in stimulus information in the words and pseudowords could not be insured. In fact, Samuel observed that the pseudowords averaged about 10% longer in duration than the words. Longer duration is usually correlated with a higher-quality speech signal, which might explain the advantage of the pseudowords over the words. Until additional research is carried out, it seems premature to conclude that phonemic restoration produces sensitivity effects, and not just bias. More generally, top-down effects on sensitivity have yet to be convincingly demonstrated, making the concept of top-down activation unnecessary to explain speech perception.

DISCUSSION

Both the TRACE model and the FLMP can be characterized as information-processing models. They characterize the transmission and transformation of information between some speech signal and its identifica-

tion. In the TRACE model, the activation of units in the trace determines which alternative is identified. The input activations generated from bottom-up sources of information are eventually obliterated by activation at the word level. In contrast, the FLMP postulates separate stages of information processing and representation. The output from feature evaluation is made available to feature integration, but the bottom-up information is not obliterated by the integration of word information. Thus, subjects have been shown to integrate two sources of information in identification and also be able to discriminate the properties of each source. The influence of phonological context on bias, but not sensitivity, is evidence in favor of the FLMP over TRACE.

The different levels in the TRACE model might be thought of as stages of processing. They differ from stages in terms of feedback from a later stage on an earlier stage.

From a decision-making perspective, interactive-activation models are nonoptimal because they allow the processing system to distort the environmental input more than is reasonable. Given a movie review from two friends, discrepencies in the reviews lead to an eventual distortion of the original reviews within interactive activation. The fact that John's review differs from Mary's, however, should not necessarily question the validity of either review. Given opposing reviews, the receiver of the reviews, however, might want to conclude very little about the value of the movie, but yet would be well-informed about each of the separate reviews. That is, the evaluation of each review is informative for the system but the integration leads to some ambiguity. In other cases, each of two sources will be somewhat ambiguous but pointing in the same direction. Integration in this situation provides more certainty than contained in either of the two sources. The integrity of the two reviews is preserved by each one separately. The stage representation of the FLMP allows for lower-level information to remain independent of higher-level information, although a *decision* about lower-level information will reflect the contribution of the higher-level information. In interactive-activation models, however, the contribution of higher-level sources of information to lower-level decisions must come at the expense of modifying the representation of the lower-level information. This aspect of the TRACE model was falsified in the present study, and it is noted that there is no unambiguous evidence for interactive activation in the literature.

REFERENCES

Braida, L. D., & Durlach, N. I. (1972). Intensity perception II: Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, **51**, 483–502.

Cohen, M. M., & Massaro, D. W. (1976). Real-time speech synthesis. *Behavior Research Methods & Instrumentation*, **8**, 189–196.

Elman, J., & McClelland, J. (1986). Exploiting lawful variability in the speech wave. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes*. Hillsdale, NJ: Erlbaum.

Jusczyk, P. W. (1986). Speech perception. In Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds.), *Handbook of perception and human performance: Vol. II. Cognitive processes and performance*. New York: Wiley.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Massaro, D. W. (Ed.). (1975). *Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics*. New York: Academic Press.

Massaro, D. W. (1979). Letter information and orthographic context in word perception. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 595–609.

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1988). Some criticisms of connectionist models of human performance. *Journal of Memory and Language*, **27**, 213–234.

Massaro, D. W. (1989). Experimental psychology: An information processing approach. San Diego: Harcourt Brace Jovanovich.

Massaro, D. W., & Cohen, M. M. (1983). Phonological context in speech perception. *Perception & Psychophysics*, **34**, 338–348.

Massaro, D. W., & Cohen, M. M. (1987). Process and connectionist models of pattern recognition. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Massaro, D. W., & Oden, G. C. (1980). Speech perception: A framework for research and theory. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 3, pp. 129–165). New York: Academic Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1–86.

Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, **85**, 172–191.

Platt, J. R. (1964). Strong inference. *Science*, **146**, 347–353.

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 378–387.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110**, 474–494.

Selfridge, O. G. (1959). Pandemonium: A paradigm for learning. In *Mechanization of thought processes* (pp. 511–526). London: Her Majesty's Stationery Office.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, **61**, 401–409.

Thompson, M. C., & Massaro, D. W. (1973). Visual information and redundancy in reading. *Journal of Experimental Psychology*, **98**, 49–54.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, **167**, 392–393.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, **8**, 338–353.

(Accepted March 9, 1989)