

ATTENTION AND PERCEPTION: AN INFORMATION-INTEGRATION PERSPECTIVE *

Dominic W. MASSARO

University of California, Santa Cruz, USA

Accepted July 1985

Attentional effects in perceptual processing are analyzed within a framework of a fuzzy logical model of perception. The recognition of a pattern is conceptualized as involving three stages of processing: featural evaluation, integration of features, and pattern classification. The model predicts no loss of resolution when multiple sources of information are integrated to recognize a perceptual event. This model is contrasted with a single-channel model in which only one source of information can be recognized at a time. The task involves a relatively novel situation of speech perception by ear and eye. No attentional decrement is observed when observers process both auditory and visual speech specifying a single speech event. This result contrasts with previous studies showing a loss when attention has to be divided between different events along auditory and visual modalities. The different results are interpreted in terms of the number of events that have to be processed. Processing two different modalities leads to an attentional decrement when the two inputs specify different events whereas no attentional decrement occurs when these inputs are integrated to recognize a single event. A distinction is made between detection that requires only the evaluation of a single source of information and recognition that requires the evaluation and integration of multiple sources of information. The current framework is also used to discuss previous empirical and theoretical work and the issue of early versus late selection.

The aspiration of this paper is to propose an information-integration account of selective attention in perceptual processing. The goal is to predict attentional effects as a function of the nature of the environ-

* This paper was written while I was a member of the project 'Perception and Action' at the Center for Interdisciplinary Research (ZIF) at the University of Bielefeld. I appreciate the extended interactions with members of the Project Group and the valuable discussions on all aspects of psychological inquiry. I thank Peter McLeod, Doug Mewhort, Odmar Neumann, Andries Sanders, and Peter Wolff for helpful comments on the paper. The writing of this paper and the research reported in the paper were supported, in part, by NINCDS Grant 20314 from the Public Health Service and Grant BNS-83-15192 from the National Science Foundation.

Send correspondence to: D.W. Massaro, Program in Experimental Psychology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA.

mental situation, the processing structures involved, and the intentions of the perceiver. In contrast to the paradigmatic concern for whether or not attentional effects occur in specific situations, the current analysis is in terms of the processes involved in perceptual recognition and the implications of these processes for attentional effects. This analysis is representative of my earlier call (Massaro 1975) for a theory of attention following rather than preceding a theory of information processing. I begin with a presentation of a general model of perceptual recognition, followed by a test of the model. The model is used as a framework for a review of related empirical and theoretical contributions. The distinction between detection and recognition processes and the issue of early versus late selection are also discussed within the context of the model.

A process model of perception

In the present paper, the concept of attention will be developed within a process model of perceptual recognition. According to the model, objects and events from a wide variety of domains are recognized in accordance with a general algorithm (Massaro 1979; Oden and Massaro 1978). The model postulates three operations in pattern recognition: featural evaluation, featural integration, and classification. Continuously valued features are evaluated, integrated with respect to prototype representations, and a classification decision is made on the basis of the relative goodness of match of the stimulus information with the relevant prototype. The model is called a fuzzy logical model of perception, and it will be helpful to discuss the concept of fuzzy logic and how it is used in the model.

In fuzzy logic, propositions are neither entirely true nor false but rather take on continuous truth values. Within this framework, it is possible to say that a team is relatively strong on offense but very weak in defense. Ordinary logical quantification would require that the team is either strong or not strong on offense and weak or not weak in defense. Fuzzy-logic theory (Zadeh 1965; Goguen 1969), on the other hand, provides a representation for propositions having a continuous nature. In fuzzy logic, the truth of a proposition is represented by values between 0 and 1 corresponding to the range between completely false and completely true. It should be noted that the concept of fuzzy

truth values is different from probability. If we say that a whale is a fish to degree 0.2, it does not mean that there is a 0.2 probability that any randomly selected whale is a fish. Rather, it is true that the concept whale represents the concept fish to degree 0.2. In general, the truth value is inversely related to the amount of stretching of the concept required to represent the object.

The present use of fuzzy logic theory requires the standard logical operations of negation and conjunction of truth values, $t(x)$. Consistent with standard logic, a reasonable definition for negation is the additive complement:

$$t(\sim x) = 1 - t(x), \quad (1)$$

where $t(\sim x)$ is the truth of not x .

Both the minimization rule and the additive rule have been proposed for the conjunction (\wedge) of two events a and b (Massaro and Cohen 1976; Zadeh 1965). Research in a number of domains, however, has supported the psychological reality of the multiplicative rule (Goguen 1969) over these other two forms of conjunction.

$$t(a \wedge b) = t(a) \times t(b). \quad (2)$$

Massaro and Cohen (1976) studied the conjunction of voice-onset-time and fundamental frequency as perceptual cues to the /si/-/zi/ distinction. A multiplicative combination of the cue values described the results about four times more accurately than did an additive combination. Oden (1977) investigated which set of definitions of fuzzy logical conjunction best fit judgments about logical combinations of pairs of statements about class membership functions (e.g., a bat is a bird, and a refrigerator is furniture). The data from the experiment were better explained by the multiplication rule than by the minimization or additive rules. Thus, the multiplicative rule is assumed in the present application of the model.

According to the fuzzy logical model of perception, perceptual recognition is carried out in three operations (Massaro and Oden 1980; Oden and Massaro 1978). The first operation is featural evaluation, during which the features transduced by the sensory systems are assigned truth values. The features are assumed to be continuous rather than discrete, and thus featural evaluation provides truth values, $t(x)$,

representing the degree to which each relevant feature is present. Taking an example from speech perception, the formant transitions of a stop consonant would be evaluated with respect to the degree to which the alternative /ba/ is supported.

The second operation is featural integration which involves the integration of the truth values of the features with respect to the prototype representations. A prototype defines a percept or concept in terms of an arbitrarily complex fuzzy logical proposition (Oden 1977). For example, a syllable prototype would represent the conjunction of both auditory and visual features defining the syllable. The integration operation would consist of replacing the respective features of each prototype with their corresponding truth values of the relevant speech event. The conjunction of these truth values determines to what degree each prototype is realized in the pattern. To distinguish between /ba/ and /da/, the truth values corresponding to the evaluation of the auditory formant transitions and the evaluation of the visual lip movements would be integrated and matched against the prototypes for the relevant alternatives /ba/ and /da/.

The third operation of recognition processing is pattern classification. During this stage, the merit of each relevant prototype is evaluated relative to the summed merits of the other relevant prototypes. The relative goodness of a prototype gives the proportion of times it would be selected as a response or its judged magnitude. This is similar to Luce's (1959) choice rule which is based on the relative strengths of the alternatives in the candidate set. In pandemonium-like terms (Selfridge 1959), we might say that it is not how loud some demon is shouting but rather the relative loudness of that demon in the crowd of relevant demons. With respect to our example, the likelihood of a /da/ identification would be equal to the goodness-of-match value to the alternative /da/ relative to the sum of the goodness-of-match values for /da/ and /ba/.

I now apply the fuzzy logical model of perception to a relatively novel situation of perceiving speech by ear and eye. The goal is to test its implications for attentional effects in the evaluation and integration of multiple sources of information. The central assumption of the model is that different sources of information specifying a single speech event are evaluated independently of one another and integrated without any loss of quality to achieve recognition. This model is contrasted with a single-channel model that assumes that only one source of informa-

tion can be evaluated at any given time to recognize a given event. Thus, the critical question relevant to the issue of attention is whether or not a deficit occurs during the perceptual recognition of a speech event consisting of two different sources of information.

Speech perception by ear and eye

It has now been well demonstrated that both auditory and visual sources of information contribute to face-to-face speech perception. A speaker's lip movements, as well as the speech sound, contribute to the perceptual recognition of a speech event (Massaro and Cohen 1983; MacDonald and McGurk 1978; McGurk and MacDonald 1976). This situation is unique since it seems to represent a truly bimodal (or amodal) perception of a speech event influenced by both sight and sound. Given the appropriate framework, it is an ideal situation for assessing how multiple sources of information are evaluated and integrated in perceptual recognition. The framework that we use is that of information integration theory (Anderson 1981, 1982) and the testing of quantitative process models of performance.

Consider the expanded factorial design illustrated in fig. 1. Subjects are asked to identify speech events consisting of just sound, just sight, or both audible and visible speech dimensions. The auditory event is chosen from a synthetic speech continuum between /ba/ and /da/. These syllables begin with voiced stop consonants that differ primarily in their place of articulation; /ba/ is articulated at the lips and /da/ further back in the mouth behind the alveolar ridge.

		AUDITORY					
		BA		DA		NONE	
VISUAL	BA						
	DA						
	NONE						X

Fig. 1. Expanded factorial combination of auditory and visual speech to include single auditory and visual conditions.

The auditory representations of these two syllables differ primarily in the onset and direction of their second and third formant transitions between the consonant and the vowel. The consonant /ba/ has rising transitions and the consonant /da/ has slightly falling transitions. By varying the transitions, while keeping all other properties constant, it is possible to synthesize a continuum of auditory syllables equally spaced between the alternatives /ba/ and /da/. The visual event is a view of the speaker articulating /ba/ or /da/.

The bimodal speech events are made by a factorial combination of the auditory and visual events (see fig. 1). Subjects identify the three kinds of speech events as /ba/ or /da/. The results allow us to compare the fuzzy logical model and the single-channel model. The fuzzy logical model assumes that the auditory and visual dimensions are evaluated independently of one another in parallel. The single-channel model assumes that only one dimension can be evaluated at a time. Given a bimodal speech event, it is also assumed that the short duration and transient nature of the critical auditory and visual information would not provide sufficient time to switch from one dimension to the other. Given these assumptions, the fuzzy logical model predicts no decrement processing both modalities whereas the single-channel model predicts a decrement relative to the single-modality situation. I now derive the predictions made by the models and illustrate what implications these predictions have for selective attention.

Fuzzy logical model

Given an auditory speech event, the featural evaluation stage determines the degree to which /ba/ is supported and analogously for /da/. Using fuzzy truth values, a value between zero and one is assigned indicating the degree to which the auditory information supports /da/. The support for /ba/ would simply be one minus the support for /da/ given the definition for negation in fuzzy logic. Given the pattern classification operation based on relative truth values (Luce 1959; Oden and Massaro 1978), the probability of a /da/ decision given an auditory event A_i is equal to

$$P(/da/ : A_i) = a_i / (a_i + (1 - a_i)) = a_i, \quad (3)$$

where a_i is the truth value of the proposition that an auditory /da/ is present.

Following the same logic, we derive an exactly analogous description for the recognition of a visual speech event V_j

$$P(/da/:V_j) = v_j / (v_j + (1 - v_j)) = v_j, \quad (4)$$

where v_j is the truth value of the proposition that a visual /da/ is present. In both of the single-dimension cases, the probability of a given decision is equal to the truth value determined from the relevant dimension.

Applying the model to the bimodal task with both auditory and visual speech, both sources provide independent evidence for the alternatives /ba/ and /da/. Defining the important auditory cue as the onsets of the second and third formants (F2–F3) and the important visual cue as lip closure, the prototypes for /da/ and /ba/ are defined as

/da/: Slightly falling F2–F3 & Open lips,
/ba/: Rising F2–F3 & Closed lips,

where F2–F3 represent the second the third formant transitions. Given that the auditory and visual sources are specified independently of one another in the prototype representation, the value of one source cannot change the value of the other source at the prototype-matching stage. Using the additive-complement definition of negation, we can represent Rising F2–F3 as (1 – Slightly falling F2–F3) and Closed lips as (1 – Open lips).

/da/: Slightly falling F2–F3 & Open lips,
/ba/: (1 – Slightly falling F2–F3) & (1 – Open lips).

The integration of the features defining each prototype can be represented by the product of the feature values. The value a_i represents the outcome of featural evaluation indicating the degree to which the auditory source A_i has Slightly falling F2–F3. Analogously, the value v_j represents the outcome of featural evaluation indicating the degree to which the visual source V_j has Open lips. It follows that the

outcome of prototype matching would be

$$\begin{aligned} /da/ &: a_i \times v_j, \\ /ba/ &: (1 - a_i) \times (1 - v_j). \end{aligned}$$

The pattern classification stage determines the relative merit of the relevant alternatives, a determination leading to the prediction that

$$P(/da/ : A_i V_j) = a_i v_j / (a_i v_j + (1 - a_i)(1 - v_j)). \quad (5)$$

Eq. (5) predicts an American-football shape to the curves generated by varying the auditory and visual sources in the bimodal condition. This shape reflects the prediction that the contribution of one source to the judgment will increase as the other source is made more ambiguous. The model predicts a statistical interaction between the auditory and visual dimensions in the bimodal condition.

Prototype matching is also assumed to occur when only visual or only auditory speech is presented although, logically, no integration is necessary. If the missing source of information is given the completely ambiguous truth value of 0.5, however, integration and pattern classification could proceed identically to the bimodal situation. This formalization makes the same prediction as derived for the single-dimension cases. For example, given a visual speech event V_j

$$P(/da/ : V_j) = 0.5 v_j / (0.5 v_j + (1 - 0.5)(1 - v_j)) = v_j, \quad (6)$$

when the missing auditory dimension is evaluated to be 0.5. This prediction is equivalent to that given by eq. (4) when only visual information is considered.

Single-channel model

This model assumes that the subject is able to evaluate optimally the auditory or the visual but not both dimensions of the speech event. A similar all-or-none attentional model is described by Sperling and Melchner (1978) and by Kinchla (1980). Given an auditory or visual trial, the subject attends to the relevant dimension with probability one. This is a reasonable assumption because the presentation conditions

were blocked and subjects could prepare in advance by attending to the appropriate modality.

On bimodal trials, the subject attends to the auditory dimension with some probability p and attends to the visual dimension with probability $1 - p$. On proportion p of the trials, the judgment is determined by the auditory dimension and on proportion $1 - p$ of the trials, it is determined by the visual dimension. Thus, predicted performance given a bimodal speech event is a simple weighted average of the two identifications given the single-dimension speech events. If the probability of a /da/ response given an auditory stimulus A_i is a_i and the probability of a /da/ response given a visual stimulus V_j is v_j , then the probability of a /da/ response given a bimodal speech event $A_i V_j$ is equal to

$$P(\text{/da/} : A_i V_j) = p a_i + (1 - p) v_j. \quad (7)$$

The weighted average prediction given by the single-channel model implies that no statistical interaction should be observed between the auditory and visual dimensions in the bimodal condition.

Experimental test

To test the models, *Ss* were tested with auditory-alone, visual-alone, and bimodal speech events as illustrated in fig. 1. The *Ss* were instructed to decide whether the speaker said /ba/ or /da/ on each trial. Given five levels along the auditory continuum and two levels along the visual continuum in the present task, the predictions of the fuzzy logical model require seven parameters (five a_i values and two v_j values). Thus, the model predicts the five auditory-alone conditions, the two visual-alone conditions, and the ten bimodal conditions for a total of seventeen independent observations with seven free parameters. What is most relevant for the role (or nonrole) of attention is that the same truth value is given a particular level of a dimension at the feature evaluation stage in both the single-dimension and bimodal condition. The single-channel model also requires five a_i values and two v_j values and an additional parameter for p , the probability of attending to the auditory dimension on bimodal trials. Thus, the single-channel model requires one more free parameter than the fuzzy logical model to predict the same number of observations.

Method

Subjects

There were two groups of *Ss* with sixteen *Ss* per group. The two groups differed widely in age (6 decades) and were composed of fourth-grade children and senior

citizens, respectively. The fourth-grade children included seven males and nine females with an age range of 8-11 to 10-4 (mean = 9-7). The fourth-grade children were recruited from an elementary school and were tested during the school day. The senior citizens were five males and eleven females with an age range of 60 to 84 (mean = 72). The senior citizens were tested during the day at their neighborhood community center. The children were given a toy and the senior citizens a complementary meal ticket for participating.

Stimuli

The speech events were recorded on a videotape. The speaker was seated in front of a wood panel background, illuminated with fluorescent light. His head filled about two-thirds of the screen. On each trial, the speaker said /ba/ or /da/. An experimental tape was made by copying the original and replacing the soundtrack with synthetic speech. The synthetic speech was synchronized with the onset of the original audio, which gave the appearance that the speech was actually coming from the speaker's articulation. The five synthetic speech syllables were equally spaced along a /ba/ to /da/ continuum. A software formant serial resonator speech synthesizer was used to create the syllables. The first 80 msec of the syllables was altered to produce a set of five syllables covering the range from /ba/ to /da/. During the first 80 msec, the first formant went from 300 to 700 Hz following a negatively accelerated path. The second formant began with one of five values equally spaced between 1125 and 1625 Hz and followed a negatively accelerated or decelerated path to 1,199. The third formant followed a linear transition to 2,729 Hz beginning from one of five values equally spaced between 2325 and 2825 Hz. These syllables were identical to those used in the Massaro (1984) experiment except that the present syllables were made to be more ambiguous, by decreasing the range of the starting frequencies of the second and third formants.

The Ss were tested in a panel truck modified to serve as a research van (Mayer 1982). All Ss viewed a 12-inch TV monitor, which presented both the audio and the video. The loudness of the audio was at comfortable listening level (70 dB-A). The Ss sat two to three feet from the monitor.

Procedure

There were three experimental conditions tested between blocks of trials: bimodal, visual, and auditory. During each trial of the bimodal condition, one of the five auditory stimuli on the continuum from /ba/ to /da/ was paired with one of the two visual stimuli, a /ba/ or a /da/ articulation. Trials in the visual condition used the same video tape but without the speech sounds. In the auditory condition, the TV screen was covered so that only the auditory information was presented.

A 250-msec bell preceded each trial in all conditions. The silent interval between the bell and the onset of the speech sound was chosen randomly from the range of 1175 to 1375 msec. The Ss had about four seconds to make a response before the next trial.

In the bimodal condition, Ss were instructed to watch and to listen to the 'man on the TV' and to tell the experimenter whether the man said /ba/ or /da/. Before the visual condition, each child watched the experimenter's mouth as she demonstrated silent articulations of the two alternatives. In this condition, children were instructed to

report whether the speaker's mouth made /ba/ or /da/. The senior citizens were simply told to lipread. In the auditory condition, the Ss were instructed to listen to each test sound to indicate whether the man said /ba/ or /da/. In the visual and bimodal conditions, the experimenter watched the Ss to insure they were watching the screen at the time of the speech event. If this criterion was not met, the trial was disregarded. All Ss made their response by verbally reporting what they heard.

The fourth-graders were tested for a session of 40 trials under each of the three experimental conditions. The senior citizens were tested for 80 trials under each condition rather than 40. The order of the three conditions was counterbalanced across Ss. Each fourth-grader contributed four responses (minus discarded trials) to each of ten bimodal conditions, eight responses to each of the five auditory conditions, and twenty responses to each of the two visual conditions. The senior citizens contributed twice this number.

Results

Given that the task required a two-alternative forced choice response, a single dependent variable, the proportion of /da/ identifications, provides all of the information about choice performance. This proportion was computed for each S at each of the seventeen experimental conditions. Figs. 2 and 3 give the average results for the three conditions for the fourth-graders and senior citizens, respectively. Separate analyses of variance were carried for each of the two groups on the bimodal, visual, and auditory conditions. For the fourth-graders, the auditory and visual variables produced signifi-

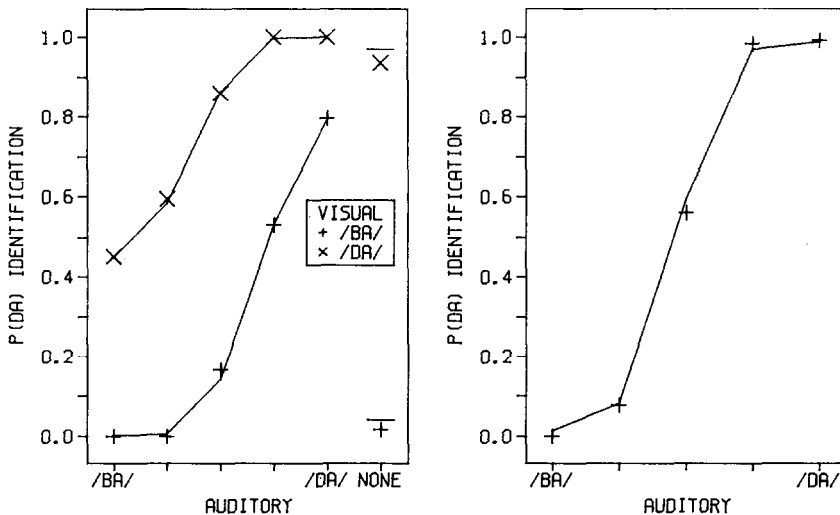


Fig. 2. Observed (points) and predicted (lines) proportion of /da/ identifications of the bimodal, visual-alone, and auditory-alone speech events for fourth-graders. The predictions are for the fuzzy logical model.

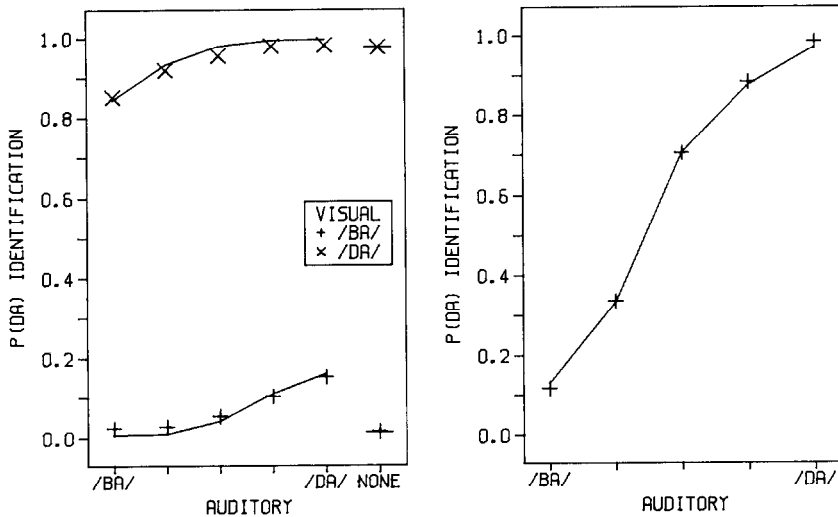


Fig. 3. Observed (points) and predicted (lines) proportion of /da/ identifications of the bimodal, visual-alone, and auditory-alone speech events for senior citizens. The predictions are for the fuzzy logical model.

cant differences as did the interaction between these two variables in the bimodal condition. Although the main effects were significant for the senior citizens, the interaction between the auditory and visual dimensions in the bimodal condition missed significance, $p = 0.16$.

The two models were fit to the results of individual *Ss* using the program STEPIT (Chandler 1969). The predictions of each model were determined by estimating the parameter values that minimized the squared deviations between the predicted and observed values for each *S*. Figs. 2 and 3 give the average predictions for the fuzzy logical model and figs. 4 and 5 for the single-channel model. Tables 1 and 2 give the average parameter values for the two models. It should be noted that the average parameter values for the fuzzy logical model in table 1 do not predict the average predictions shown in figs. 2 and 3. Given the mathematical formalization, the average prediction is not equal to the prediction generated from the average parameter values.

The fuzzy logical model gave a much better description of individual *S* performance than did the single-channel model for both groups of *Ss*. The goodness of fit is given by the average root mean squared deviation (RMSD) between the predicted and observed values. The average RMSD values were 0.034 and 0.031 for the fourth-graders and senior citizens for the fuzzy logical model whereas these same values were 0.121 and 0.049 for the single-channel model. Thus, the description given by the single-channel model was about four times poorer for the fourth-graders and about one-and-a-half times poorer for the senior citizens than was the description given by the fuzzy logical model. The much better description given by the fuzzy logical model is especially impressive given that this model required one fewer free parameter than did the

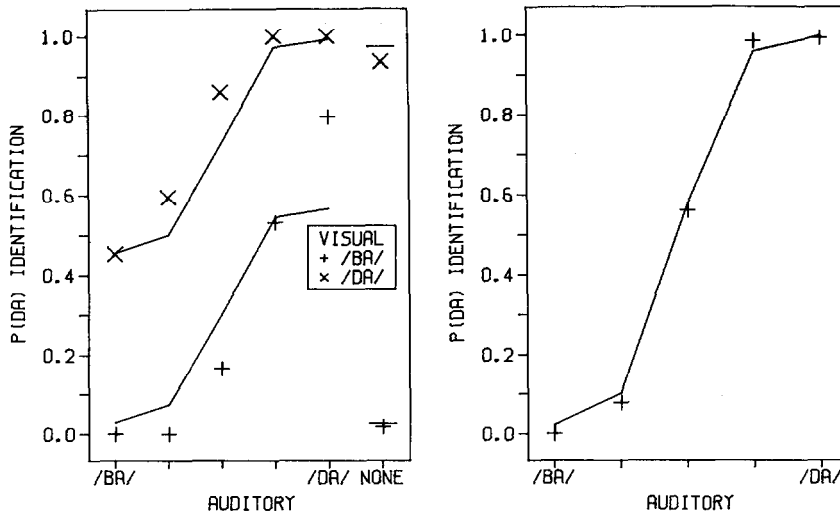


Fig. 4. Observed (points) and predicted (lines) proportion of /da/ identifications of the bimodal, visual-alone, and auditory-alone speech events for fourth-graders. The predictions are for the single-channel model.

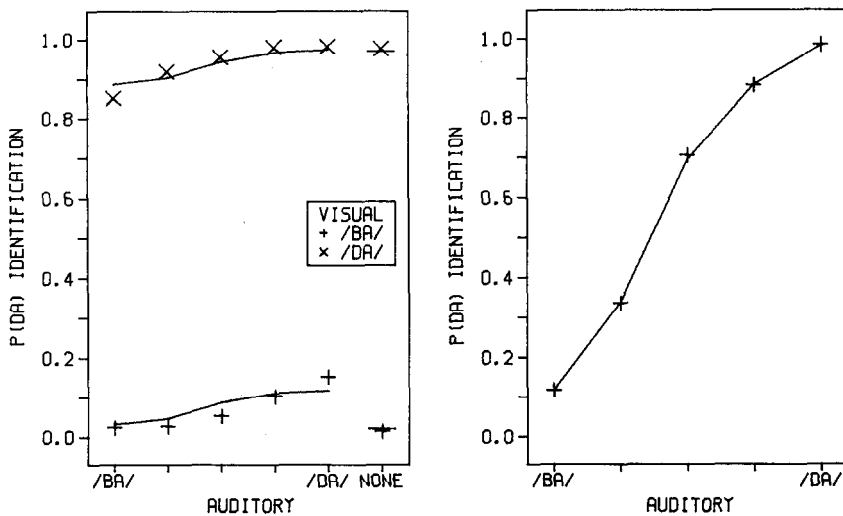


Fig. 5. Observed (points) and predicted (lines) proportion of /da/ identifications of the bimodal, visual-alone, and auditory-alone speech events for senior citizens. The predictions are for the single-channel model.

Table 1

Average parameter values indicating degree of /da/-ness given the auditory and visual sources for the fuzzy logical model for the two groups of subjects.

Group	Auditory (a_i)					(Visual (v_j))	
	/ba/	2	3	4	/da/	/ba/	/da/
Fourth grade	0.014	0.081	0.594	0.969	0.988	0.043	0.970
Senior citizens	0.130	0.332	0.706	0.874	0.969	0.012	0.977

single-channel model. The absolute goodness of fit of the fuzzy logical model was also very good supporting the assumption of no loss of efficiency in the evaluation and integration of two sources of information specifying a speech event, relative to a single source specifying the event.

The results for the fourth-graders provide unequivocal support for the fuzzy logical model of perception. The lack of a statistical interaction and the respectable description by the single-channel model for the senior citizens should not be considered as evidence in favor of this model. The senior citizens were much more influenced by the visual source than the auditory source which may simply reflect a hearing loss with age. Thus, the single-channel was able to capture the results by assuming that the auditory source was attended to only about 10% of the time in the bimodal condition (cf. table 2). If the auditory information were made more discriminable, the contribution of the visual source might be attenuated, and the single-channel model should fare much more poorly. This result actually obtains with senior citizens; when the auditory syllables are increased in discriminability, a significant interaction occurs (Massaro, unpublished results).

Although the two models are clearly different, they make fairly similar predictions when one of the dimensions is much more influential than the other. The models make very different predictions when the two dimensions are equally influential. Thus, the most informative test between the models is in the latter case, and it is best represented by the results of the fourth-grade Ss. The fourth-graders were influenced by the auditory and visual sources more equally and the results provide unambiguous support for the fuzzy locigal model over the single-channel model.

Table 2

Average parameter values indicating the probability of a /da/ decision to the auditory and visual sources for the single channel model for the two groups of subjects.

Group	Auditory (a_i)					Visual (v_j)		Bias (p)
	/ba/	2	3	4	/da/	/ba/	/da/	
Fourth grade	0.022	0.101	0.576	0.958	0.997	0.026	0.974	0.553
Senior citizens	0.116	0.338	0.693	0.883	0.982	0.021	0.970	0.102

As just noted, the relative effects of visual and auditory speech differed for the two groups of Ss. In terms of the fuzzy logical model, the differences between the two groups are completely accounted for by the information values assigned the auditory and visual sources during the evaluation operation. Table 1 shows that the auditory parameters cover more of the range between zero and one for the fourth-graders relative to senior citizens. The opposite result occurs for the visual values. A more extensive range along a given dimension reflects better discrimination along that dimension and a larger impact on identification performance. These results are consistent with previous findings that children are less influenced by visual information than are adults (Massaro 1984; McGurk and MacDonald 1976) and provide the new observation that senior citizens are less influenced by auditory information than are children. Testing these two populations of Ss represents a stronger test of attentional contributions to the perceptual recognition of bimodal speech, given the large differences in the identification functions.

Discussion

The results of the experiment with both senior citizens and fourth-graders provide strong support for the assumption of independent auditory and visual features which, in turn, illustrate no loss in processing two dimensions of a speech event relative to processing just one. One might predict that the contribution of a given source of information might be compromised in the bimodal condition relative to the single-dimension conditions. For example, Miller (1982: appendix) implemented an attentional contribution to the activation of auditory and visual modalities in a bimodal signal detection task. The assumption was that the rate at which activation accrues is modified as a function of attention. In the present task, we might expect that the information value of a given source would be neutralized in the bimodal relative to the single-dimension condition. In fact, no such compromise occurred. Thus, we have evidence that subjects can process two sources of information as well as one if the two sources are integrated to achieve recognition of a single perceptual event.

Reaction times to speech events

An important assumption of the information-integration view of attention is that the evaluation of one source of information remains independent of the presence and evaluation of other sources. Thus, the amount of /da/-ness assigned to the auditory source should be independent of the presence or absence of a visual source. This assumption was supported by the present experiment's analysis of the percentage of choice identifications. In addition to predicting choice, the model makes testable predictions concerning reaction time (RT) under various stimulus conditions. Consider RTs to an auditory /ba/, a

visual /ba/, and a bimodal /ba/ consisting of an auditory /ba/ and visual /ba/. If the assumption of independent evaluation holds, it should be possible to account for the RTs to a bimodal /ba/ in terms of simply distributing the RTs to a visual /ba/ and to an auditory /ba/. If the two dimensions are independent, we might expect RTs to the bimodal event to be somewhat faster than those to the single-dimension events, but the advantage should be completely accounted for by the minima distribution resulting from the combination of the distribution of RTs to the auditory stimulus and the distribution to the visual stimulus (Gielen et al. 1983; Raab 1962). If the two dimensions are not evaluated independently of one another, then the distribution of RTs to the bimodal stimulus will not be accounted for by the minima distribution.

To test this idea, Cohen (1984) used six types of trials in a choice RT task: auditory /ba/, auditory /da/, visual /ba/, visual /da/, auditory-visual /ba/, and auditory-visual /da/. Subjects were instructed to respond as quickly as possible without making too many errors by hitting buttons corresponding to the /ba/ and /da/ alternatives. The RTs were analyzed individually for each of the ten subjects in the experiment.

Fig. 6 shows the distributions of RTs of two representative subjects to the six stimuli. As can be seen in the figure, each subject is somewhat faster to the bimodal speech event but no faster than expected if the subject simply begins to initiate a response when either the auditory or visual dimension is identified. The advantage of bimodal trials can be accounted for simply in terms of the variability along each dimension allowing the average RT to bimodal events to be shorter than the average to either dimension presented alone. An analysis of variance showed that the mean latency of the bimodal trials did not differ from the mean of the predicted minima distributions across the ten subjects. Thus, we have strong evidence for parallel and independent evaluation of auditory and visual featural information in the perception of a speech event.

These RT results might seem to contradict the earlier conclusion that the two sources of information are integrated before a response is initiated. The six stimuli used in the RT study were completely unambiguous, however, and the subjects were instructed to respond as quickly as possible. Given a single-modality stimulus, it is reasonable to assume that the missing modality is assigned the completely ambiguous

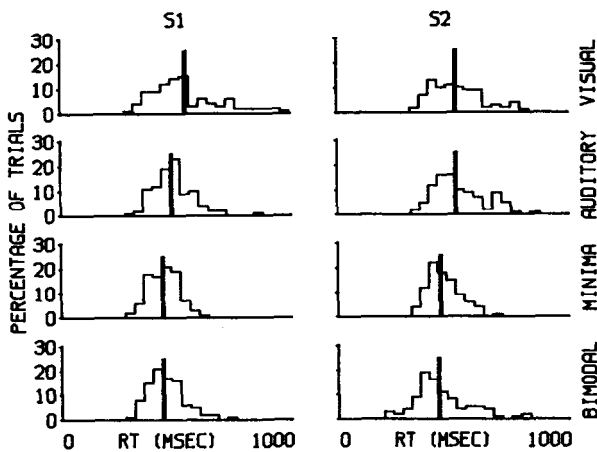


Fig. 6. Distribution of RTs for two representative subjects to visual-alone, auditory-alone, and bimodal speech events. The minima distribution is that predicted for bimodal trials based on the auditory-alone and visual-alone trials.

truth value 0.5 (see eq. (6) and its justification). Thus, the integration of the unambiguous truth value given the single modality with the ambiguous truth value of the missing modality would still provide unambiguous information. The results are consistent with the occurrence of all three operations assumed by the model, even when only a single modality is present.

Relation to other experimental findings

The finding of no divided attention loss between auditory and visual perception contrasts sharply with the large losses found across these dimensions when they do not specify the same event. Massaro and Kahn (1973), Massaro and Warner (1977), and Long (1975) all report substantial decrements when both auditory and visual recognition are required relative to just one. Visual letter recognition and recognition of tonal quality cannot be performed simultaneously without a performance decrement relative to the single modality control (Massaro and Warner 1977). In contrast to the bimodal speech situation, we find significant deficits in the simultaneous evaluation and integration of information specifying two unique events. That is, letter recognition

and tonal recognition require that attention be divided across two modalities. One integration process is needed for the visual task and another for the auditory task. In contrast, the auditory and visual speech information specify a single event, and only a single integration process is necessary to achieve bimodal speech recognition.

One might question the comparison between the speech identification task and the task of identifying tones and letters. The former requires only a single decision and the latter requires overt identification of both dimensions. Thus, the decrement found in the latter situation might be due to making two decisions rather than identifying separate events. One way to assess this possibility is to require overt identifications of two dimensions when these specify a single event. If the decision explanation is correct, then a decrement should also be found when the two dimensions are identified even though they make up a single event or belong to the same object. Moore and Masaro (1973) found only a small 1.5% decrement when subjects identified both the timbre and loudness of a single tone relative to identifying just one of these dimensions. The authors also provided strong evidence for independent evaluations of the two dimensions with the finding that correct identification of one dimension was independent of the other on a given test trial.

Duncan (1984) has recently provided evidence for an object-based theory of visual attention, as opposed to a discrimination-based or a space-based theory. Two judgments concerning the same object were made simultaneously without a decrease in accuracy relative to a single judgment. When the two judgments were about different objects, however, a significant decrease in accuracy occurred. A discrimination-based theory cannot predict this result since it is the number of separate discriminations that is important for this theory. A space-based theory cannot predict the results, since the author controlled for the spatial distribution of information in the same-object and different-object conditions. The result appears to be consistent with the idea developed in this paper that cross talk in the process of integrating multiple sources of information can degrade perceptual recognition when two or more objects or events must be processed simultaneously.

The integrity of integrating multiple sources of information in the perceptual recognition of an event appears to have a parallel in cognitive decision making. Brehmer and Slovic (1980) asked subjects to rate the attractiveness of various jobs described by either salary level,

commuting time, or both of these attributes. The scale values derived from the single-attribute conditions did not differ from those derived from the double-attribute condition. Thus, the process of integrating two attributes to determine job attractiveness did not compromise the evaluations of the two component attributes. This result is exactly analogous to our finding that the evaluations of auditory and visual speech in the single-dimension conditions were capable of predicting performance in the bimodal speech condition.

The ability to integrate multiple sources of information without loss is relevant to the issue of integrating top-down and bottom-up sources of information. In the framework presented here, perceivers should be capable of integrating these two sources when they specify a perceptual object or event. Massaro and Oden (1980) review evidence consistent with this proposal in the domain of speech perception and it is also supported in the domain of visual word recognition. As an example, Massaro (1979) independently varied the visual information specifying a letter and its orthographic context in a letter recognition task. Both variables influenced performance and, most importantly, it was demonstrated that the two sources can be considered independent contributions that are not compromised in the integration process. Thus the idea of independent evaluation and integration of multiple sources of information appears to be an appropriate description of a broad range of perceptual abilities.

Role of evaluation and integration processes

We now review some representative studies of attentional effects in information processing in terms of whether or not integration is necessary to perform the task. The present framework predicts no decrease in perceptual resolution with increases in the number of items that must be processed if only evaluation is necessary. If integration is also necessary, however, perceptual resolution should decrease with increases in the number of items to be processed. The reason is that cross talk should be much greater when integrating multiple sources of information specifying multiple events than when only evaluation of multiple sources is required. We will discuss only a few representative studies in terms of the role of evaluation and integration processes. Swets (1984) provides a more comprehensive review of the relevant literature.

Two experimental tasks investigated by Shaw (1984) can be interpreted in terms of evaluation and integration processes. One task required a luminance detection and the other required a letter detection. In terms of the present framework, there is only a single source of information relevant to the first task but multiple sources relevant to the second. Thus, luminance detection requires only evaluation and letter detection requires both evaluation and integration. Increasing the number of items that must be detected simultaneously, therefore, should have different consequences in the two situations. There should be no decrement in the evaluation process as the number of items is increased, but a decrement in the integration process with increases in the number of items.

Evaluation tasks

Shaw (1984) presents convincing evidence that increasing the number of locations over which detection must be divided causes *no* loss in quality of the information evaluation for luminance detection. Subjects monitored two or four continuously illuminated corner lights. The target presentation involved incrementing the luminance of one of the lights, and subjects reported the location 'giving the strongest impression that a light brightened'. The decrease in performance accuracy found with monitoring four relative to two positions was well described by a decision process and did not require any additional attentional explanation. That is, we should expect somewhat poorer performance when subjects have to choose among four relative to two positions (cf. Eriksen and Spencer 1969; Gardner 1973; Massaro 1975: ch. 15) based on decision processing alone. Shaw's findings were consistent with the conclusion that the quality of the information evaluation at each position was independent of the number of locations being monitored.

The role of location cueing in light detection task developed by Posner and his colleagues might be seen as evidence for attentional effects at the evaluation stage (Posner 1980). Subjects are cued that a target light will more likely occur to the left or right of fixation. Reaction times (RTs) to the light are faster when the light occurs in a predictable relative to an unpredictable position. This result might be interpreted as evidence for selective attentional effects in detection. That is, subjects might have speeded up the perceptual processing of the light in a predictable position by selectively allocating processing

capacity to that position. This interpretation would also be consistent with the additional finding of slower RTs to a target presented in an unpredictable position relative to one presented in a neutral position.

The RT advantage observed with a precue does not necessarily mean that the quality of the perceptual processing was enhanced, however. A decision process (occurring at the pattern classification operation in the fuzzy logical model) might be responsible for the advantage that is observed (Duncan 1980a; Shaw 1984; Sperling 1984). Subjects may simply set a more liberal criterion for initiating a response to a target presented in a predictable position relative to the criterion maintained for a target in an unpredictable position or even a neutral position. Thus, a considerable advantage in RT can be achieved with only a small (and possibly unnoticeable) change in accuracy. In terms of our process analysis, the 'attentional' effects that are observed would be due solely to the pattern classification operation: evaluation would not be influenced by cueing.

Pursuing the possible differences between the accuracy and RT paradigm, Shaw (1984) tested whether RT would be facilitated if a given position were made more predictable in terms of monitoring a luminance increment. She found positive evidence of the same form observed by Posner et al. (1980). To support the decision explanation, she reported a preliminary finding that subjects do seem to be able to use different criteria associated with two different locations in the display. As noted by Duncan (1981), the efficacy of a selection cue does not necessarily reflect how fully the item is analyzed.

A finding by Tsal (1983) is also consistent with a postperceptual interpretation of the facilitation given a precue in the Posner et al. (1980) task. Subjects named a target letter presented at varying distances away from the fixation point. A peripheral cue was presented some time before presentation of the target letter. RTs decreased with increases in the cue-target interval and showed more facilitation with relatively distant as compared to relatively near targets. The difficulty of the target-naming task was also varied; the two alternatives were the letters O and X in the easy task and O and D in the hard task. Although difficulty increased RT, it did not interact with the cue target interval. Following additive-factor logic, there should have been a significant interaction if both variables were affecting the same stage of processing. Since the discriminability of the test alternatives is an important variable for the recognition process, it appears that the

cue-target interval is having its effect at some later stage, such as response selection and execution. Posner (1980) also appears to reject a sensory quality (evaluation) explanation of the RT advantage to predictable targets. His attentional interpretation involves a 'system for routing information and for control of priorities' (Posner 1980: 9). Thus, we conclude that the evaluation of information responsible for detection at different spatial locations is not under attentional control. This conclusion stands in marked contrast to the conclusion warranted with respect to the integration of multiple sources of information in recognition.

Evaluation and integration tasks

Shaw (1984) carried out a letter detection experiment exactly analogous to the luminance detection study. Subjects judged the location of a target letter among the occurrence of nontarget letters. The performance decrement found with a display size of four relative to a size of two was considerably larger than what could result from only a decision process. Identical results were found for automatic as well as controlled search as defined by Schneider and Shiffrin (1977). Thus, some sort of attentional explanation is needed; one possibility is based on the poorer quality of information integration for displays of four relative to displays of two items.

The apparently discrepant results in luminance and letter detection are nicely accounted for by the distinction between evaluation and integration. In the former case, no integration of multiple sources of information is necessary in the detection of a single light. In the latter case, multiple sources of information are necessary to detect a single letter and therefore integration is necessary. A decrement is found when integration is necessary and two or more separate events must be processed. On the basis of the luminance detection results, the quality of the evaluation of a particular source of information remains independent of the number of other sources that must be evaluated. On the basis of the letter detection results, the quality of the integration process decreases with increases in the number of simultaneous integrations that are performed.

Lupker and Massaro (1979) offer a quantitative picture of selective attention in letter recognition given multiple letter displays. Each display contained four items positioned on the corners of an imaginary square centered around the fixation point. A single target letter chosen

from the response set E, I, F, or T was presented in each display of four items. The other three characters were all zeros or were all hybrid letters made up of a combination of some of the features of the target letters. The display was presented for a 10-msec duration and preceded or followed by a 200-msec cue indicating the position of the target letter. The cue was either an arrow pointing to the target from the center of the display or a pattern mask positioned at the same location as the target.

Fig. 7 gives the observed target recognition in d' values as a function of the stimulus onset asynchrony (SOA) between the target and cue. Positive SOA values refer to the cue following the display. Given the difference in the cue and display durations, the cue both preceded and followed the test display at the negative SOAs (but did not occur during the display). First, consider the no-cue condition as a function of the nonconfusable (zero) versus confusable (hybrid) background letters. Performance is, at least, three times better with a nonconfusable relative to a confusable background. This result is similar to that reported by Gardner (1973) and has usually been interpreted as due to a decision process. The decision process explanation is based on all four items

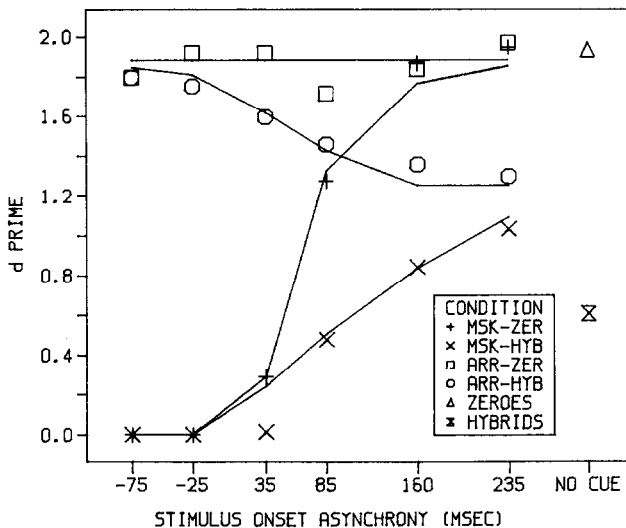


Fig. 7. Observed (points) and predicted (lines) performance in d' values for the cueing and masking conditions as a function of SOA and confusable (hybrids) and nonconfusable (zeros) backgrounds (from Lupker and Massaro 1979).

competing for the decision with the confusable display and only the target item being considered for the nonconfusable display. However, the decrement in performance accuracy in the Lupker and Massaro study is simply too large to be accounted for by differences in a decision process.

Another result easy to comprehend is the flat function for the arrow cue given the nonconfusable background. The arrow cue is redundant in this condition since the nonconfusable targets provide the cueing information needed to selectively process the target. Thus, performance with an arrow cue is no better at any SOA than performance with no cue. It turns out that the nonconfusable targets allow attention to be switched more quickly than presentation of the arrow cue (Jonides 1981), which is why no additional advantage is obtained when the cue arrow precedes the test display at a very short SOA. With nonconfusable background letters, the cue can be considered to reside in the target letter itself since subjects only have to discriminate which location contains linear as opposed to the curvilinear features.

Although an arrow cue does not facilitate performance with a nonconfusable background, a mask presented at these same SOAs is still capable of interfering with performance out to about a 160-msec SOA. Thus, there is a significant time course to the perceptual process even when selective attention can be switched immediately to the target letter.

With the confusable background, delaying the arrow cue reduces performance accuracy. As pointed out by many authors, this result could represent a selective attention effect on perceptual processing or a forgetting of categorized information from short-term memory (Estes and Taylor 1964; Massaro 1975; Sperling 1960). The increasing backward masking function with increases in SOA with the mask cue, however, eliminates the possibility of a forgetting explanation. If the target had already been categorized, as must be assumed by the forgetting explanation, then no masking should have occurred. If the target had been recognized, the mask should have functioned to facilitate performance at short SOAs as was the case with the arrow cue. The mask was about effective as the arrow cue in improving performance relative to the no-cue condition at the longest SOA. This result illustrates that the mask served both a masking function and a cueing function in the task. The predicted results also shown in fig. 7 show that Lupker and Massaro (1979) achieved a good quantitative descrip-

tion of the results by extending a perceptual processing model (Massaro 1970, 1975) to incorporate attentional effects in the process of recognition. The description of the model and its fit to the results are given in the original report. The results provide strong support for selective attention in letter recognition since explanations in terms of short-term memory limitations and decision processes are inadequate.

With respect to the current thesis, attentional effects were observed in perceptual recognition because recognition of the target letter required the integration of multiple sources of information. These sources would be the features comprising the target set of letters. Without an attentional cue, the subject would be required to integrate information specifying four different targets. This would lead to a performance decrement relative to being able to integrate information specifying just a single target when an attentional cue is given and attention is switched before perceptual processing is complete.

Detection and recognition processes

The current framework is related to previous work on attentional effects in detection and recognition processes (Massaro 1975: chs. 15 and 16), in which some evidence was provided to support the idea of attentional effects in the process of recognition but not the process of detection. In many respects, the differences between these two stages of processing can be accounted for by the role of evaluation and integration. The process of integration is required for recognition in addition to the evaluation process required for both detection and recognition. Evaluating multiple sources of information might produce little cross talk, whereas integrating the sources specifying different events might produce significant cross talk. Although this proposal probably only approximates the distinction between detection and recognition, it seems to make sense of previous experimental findings. We distinguish between the acts of detection and recognition in terms of the processes that are involved. Detection answers the question, 'Was any stimulus presented?' recognition answers the question, 'Which stimulus was presented?' Detection provides information about whether something occurred whereas recognition provides information concerning what particular pattern occurred.

One can substantiate two important processes in detection: Sensory evaluation and decision. The evaluation process can be considered to provide some output indicating the degree to which it is likely that a

signal was presented. Utilizing the concept of fuzzy logic (Zadeh 1965), we can describe the output of evaluation as a truth value, $t(x)$, of the proposition that a stimulus was present. The truth values range from zero representing completely false to one representing completely true. A completely ambiguous state would have the value of 0.5. The decision process can be conceptualized as establishing a criterion value such that truth values greater than the criterion are mapped into 'yes' judgments. Truth values less than the criterion are translated into 'no' judgments.

Recognition includes evaluation and decision as does detection, but recognition also includes integration of a number of independent evaluations. Consider the detection and recognition of the uppercase letter G. Detection requires evaluation of only one source of information, e.g. the relative amount of figure to ground or black to white in the display. Recognition of the letter, on the other hand, requires the evaluation of a number of component attributes or features. To recognize G and not Q, both the degree of openness of the right side of the circle and the obliqueness of the straight line must be evaluated.

Successful recognition also requires combining or integrating the information obtained from each of the evaluations. In the letter recognition example, the truth values representing the evaluations of the two attributes would be conjoined to obtain the overall evidence of each of the two alternatives. The decision process in recognition which follows integration is analogous to that in detection and is easily extended to situations with more than two alternatives (Luce 1959; Oden and Massaro 1978).

This process analysis reveals that the primary difference between detection and recognition concerns the number of relevant sources of information. Each source of information can be considered to require a separate evaluation. Detection requires evaluation of only a single source whereas recognition requires the evaluation and integration of multiple sources. The integration operation is central to the ability to categorize an event as an instance of one class of events as opposed to another. In terms of the present framework, therefore, the crucial difference between detection and recognition is that the latter requires the integration of multiple sources of information, while the former does not. If detection and recognition are defined in this way, we predict increasing the number of items, events, or objects that must be processed should have little consequence for detection, but should degrade recognition substantially.

Early versus late selection

The central implication of the fuzzy logical model of pattern recognition is that the evaluation and integration of multiple sources of information occurs without any attentional deficit. An event defined by multiple sources of information can be classified in such a way that the evaluation and integration of the multiple sources leads to optimal performance. An event defined by a single source of information is processed no more efficiently than the same event defined by multiple sources of information. This is, the contribution of a given source of information to pattern classification is identical in the two cases. There is no compromise of the information processing of a given source when it must be evaluated and integrated with other sources of information defining the event.

Given a single event, no attentional contribution is necessary to describe its recognition. A major controversy in attentional theory is whether an attentional contribution is necessary during the recognition of multiple events. The two major classes of theory have been based on early and late selection, respectively. According to early-selection theories, a subset of co-existing events can be processed selectively and, therefore, resolved more efficiently relative to those items not selected. Using relatively discriminable cues such as spatial location or color, it is possible to select an item from a display of items and enhance its processing. This is the traditional interpretation of the partial-report task. According to late-selection theory, on the other hand, multiple events are processed in parallel to the point of recognition without any attentional contribution. Recognition of each of the events proceeds independently without a contribution of attention. Selection of one item at the expense of other items does not enhance its perceptual processing or compromise the processing of the other items presented. For late selection theories, the results previously described as evidence for early selection are explained by short-term memory, decision, and/or sensory quality differences with increases in the number of items simultaneously presented (see Lupker and Massaro 1979).

The present research does not speak directly to the controversy between early and late selection. It is encouraging, however, that evidence from both camps might be interpreted in the framework of the model. The important variables appear to be the number of sources of information specifying the test items and the cross talk that can occur

when multiple items must be processed in parallel. Supporting late selection, the results from the automaticity literature show that practiced subjects can process multiple items in a display in parallel without any attentional deficit. In the Schneider and Shiffrin (1977) experiments, for example, increasing the number of items in the target set does not increase the reaction time in a situation of automatic search. Automatic search involves the case in which the targets are held constant for many trials of practice. Other convincing results come from Moray, Sorkin, and their colleagues (Moray et al. 1976; Ostry et al. 1976; Sorkin et al. 1976). These experiments show that with practice subjects can apparently monitor two channels of information as well as one for targets whether they be pure tones differing in frequency or words differing in semantic class.

LaBerge (1981) offers an explanation of these results in terms of only a single feature (source of information) being necessary for recognition of any one item. In terms of the present framework, automatic recognition would involve featural evaluation and decision, but not integration. Thus, we would expect that increasing the number of items to be processed would produce results like those found in evaluation tasks rather than in tasks involving both evaluation and integration. Another possibility is that the perceptual learning leading to automaticity has the primary consequence of reducing cross talk during the processing of multiple sources of information.

In favor of early selection, the conjunction of form and color of an object is said to require attention in that reaction times increase systematically with increases in the number of objects presented in a display. To recognize a green T, the observer must search the display serially to evaluate each item individually for the appropriate conjunction (Treisman and Gelade 1980). The results reviewed in the previous section of this paper provide some evidence that early selection is possible, and from an information processing point of view it is necessary to define the nature of the processing responsible for early selection. In the Lupker and Massaro experiment, for example, the subject must process the four items in a display in parallel until a cue is presented and the appropriate target location can be selected. The advantage of performance with decreases in the display-cue interval seems to reflect a deficit when multiple locations must be processed for meaning in parallel. A similar interpretation can be given for the letter detection experiments of Shaw.

One possibility is that it is not possible to integrate multiple sources of information defining multiple events in parallel without some loss. A reasonable interpretation might be cross talk in that the sources of information specifying one event might be integrated in the perception of another event. Thus, increasing a number of events that have to be identified can compromise the integration of the sources defining those events. Multiple sources can be integrated without mutual interference only when they specify a single event.

Another possibility that presents itself, however, is the fact that multiple decisions have to be made given multiple events. Duncan (1980b) presents an experiment supporting the idea that requiring multiple decisions degrades performance over and beyond that found for processing multiple events. In one condition subjects were required to make only one decision in terms of whether or not a given target was present in a display. Increasing the number of items from two to four decreased performance only minimally. In a second condition the same displays were presented but now subjects had two responses and had to make independent decisions about whether the target occurred on the horizontal limb of the display or on the vertical limb of the display. The deficit observed when the size of the display was increased from two to four items was much larger than in the first condition. Duncan's (1980b) point was that the attentional deficits are due to multiple decisions rather than recognizing multiple events. However, Duncan's (1984) findings, those of Moore and Massaro (1973), and the luminance detection study of Shaw (1984) demonstrate that multiple decisions are not necessarily sufficient for a performance decrement.

There is also a confounding in the contrast studied by Duncan in that location information was necessary to perform the second task but not the first. Requiring location information may have been responsible for the additional deficit that was required when two decisions were made as opposed to just one. It is interesting that location information was also central to the experiments of Shaw and those of Lupker and Massaro. It could be the case that attentional effects are found in visual information processing primarily when location information must be conjoined with item information. This would be an intriguing situation given the recent findings of apparently qualitatively different systems being responsible for the processing of location information and item information (Leibowitz and Post 1982).

Related theoretical work

The current approach to perceptual attention has promising similarities to other recent developments. The term sources of information has also been used by Kinchla (1980) in place of stimuli since memory might also be thought of as a source of information. However, he did not distinguish between sources of information specifying the same object and those specifying different objects. Kahneman and Treisman (1984) emphasize the important difference between selection of objects (or inputs) and selection of properties (cf. also Treisman 1969). In their view, people do not easily ignore properties of an attended object whereas irrelevant objects can be ignored quite easily. As an example, Kahneman and Henik (1981) found a decreased Stroop effect when the colored ink and the incompatible name were spatially separated (see also Van der Heijden et al. 1984). Analogously, Treisman et al. (1983) asked subjects to read a word as quickly as possible and to locate the position of a gap in a frame placed either around the word or adjacent to it. Both the reading task and the gap detection task were better when the frame surrounded the word. In the framework presented here, two sources of information are more efficiently processed when they can be integrated into a single event. As stated by Neisser (1976), we pay attention to objects and events, not sensory inputs.

Treisman and Gelade (1980) propose a feature-integration theory in which attention must be directed to an object to conjoin its separate features for recognition. This view is compatible with the current proposal that multiple sources of information can be integrated efficiently only if they specify a single event. Two or more events cannot be recognized in parallel without some cost if they each are specified by multiple features. The reason offered at the present time is that of cross talk in that the multiple sources of information specifying different events cannot be integrated independently. The important role of cross talk has been acknowledged previously (Broadbent 1982; Neumann 1985), although there is little direct evidence for its role during integration. Long (1975) found cross talk in simultaneous pitch and brightness judgments. Subjects tended to identify the pitch of the tone as high when the light was bright. The current view provides a specific computational model of the integration process and offers the potential for determining the nature of the cost that is produced when observers are required to recognize two or more events simultaneously.

References

- Anderson, N.H., 1981. Foundations of information integration theory. New York: Academic Press.
- Anderson, N.H., 1982. Methods of information integration theory. New York: Academic Press.
- Brehmer, B. and P. Slovic, 1980. Information integration in multiple-cue judgments. *Journal of Experimental Psychology: Human Perception and Performance* 6, 302–308.
- Broadbent, D.E., 1982. Task combination and selective intake of information. *Acta Psychologica* 50, 253–290.
- Chandler, J.P., 1969. Subroutine STEPIT – Finds local minima of a smooth function of several parameters. *Behavioral Science* 14, 81–82.
- Cohen, M.M., 1984. Processing of visual and auditory information in speech perception. Dissertation, University of California, Santa Cruz.
- Duncan, J., 1980a. The demonstration of capacity limitation. *Cognitive Psychology* 12, 75–96.
- Duncan, J., 1980b. The locus of interference in the perception of simultaneous stimuli. *Psychological Review* 87, 272–300.
- Duncan, J., 1981. Directing attention in the visual field. *Perception & Psychophysics* 30, 90–93.
- Duncan, J., 1984. Selective attention and the organization of visual information. *Journal of Experimental Psychology: General* 113, 501–517.
- Eriksen, C.W. and T. Spencer, 1969. Rate of visual processing in visual perception. Some results and methodological considerations. *Journal of Experimental Psychology Monograph* 79, (2, Part 2).
- Estes, W.K. and H.A. Taylor, 1964. A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Sciences* 52, 446–454.
- Gardner, G.T., 1973. Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology* 4, 130–155.
- Gielen, S.C., R.A. Schmidt and P.J.M. Van Den Heuvel, 1983. On the nature of intersensory facilitation of reaction time. *Perception & Psychophysics* 34, 161–168.
- Goguen, J.A., 1969. The logic of inexact concepts. *Synthese* 19, 325–373.
- Jonides, J., 1981. 'Voluntary versus automatic control over the mind's eye's movement'. In: J. Long and A. Baddeley (eds.), *Attention and performance IX*. Hillsdale, NJ: Erlbaum.
- Kahneman, D. and A. Henik, 1981. 'Perceptual organization and attention'. In: M. Kubovy and J.R. Pomerantz (eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum.
- Kahneman, D. and A. Treisman, 1984. 'Changing views of attention and automaticity'. In: R. Parasuraman and D.R. Davies (eds.), *Varieties of attention*. New York: Academic Press.
- Kinchla, R.A., 1980. 'The measurement of attention'. In: R.S. Nickerson (ed.), *Attention and performance VIII*. Hillsdale, NJ: Erlbaum.
- LaBerge, D., 1981. 'Automatic information processing: a review'. In: J. Long and A. Baddeley (eds.), *Attention and performance IX*. Hillsdale, NJ: Erlbaum.
- Leibowitz, H.W. and R.B. Post, 1982. 'The two modes of processing concept and some implications'. In: J. Beck (ed.), *Organization and representation in perception*. Hillsdale, NJ: Erlbaum.
- Long, J., 1975. Reduced efficiency and capacity limitations in multidimension signal recognition. *Quarterly Journal of Experimental Psychology* 27, 599–614.
- Luce, R.D., 1959. Individual choice behaviour. New York: Wiley.
- Lupker, S.J. and D.W. Massaro, 1979. Selective perception without confounding contributions of decision and memory. *Perception & Psychophysics* 25 60–69.
- MacDonald, J. and H. McGurk, 1978. Visual influences on speech perception processes. *Perception & Psychophysics* 24, 253–257.
- McGurk, H. and J. MacDonald, 1976. Hearing lips and seeing voices. *Nature* 264, 746–748.

- Massaro, D.W., 1970. Perceptual processes and forgetting in memory tasks. *Psychological Review* 77, 557–567.
- Massaro, D.W., 1975. *Experimental psychology and information processing*. Chicago, IL: Rand-McNally.
- 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., 1984. Children's perception of visual and auditory speech. *Child Development* 55, 1777–1788.
- Massaro, D.W. and M.M. Cohen, 1976. The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction. *Journal of the Acoustical Society of America* 60, 704–717.
- Massaro, D.W. and M.M. Cohen, 1983. Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 9, 753–771.
- Massaro, D.W. and B.J. Kahn, 1973. Effects of central processing on auditory recognition. *Journal of Experimental Psychology* 97, 51–58.
- Massaro, D.W. and G.C. Oden, 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. New York: Academic Press.
- Massaro, D.W. and D.S. Warner, 1977. Dividing attention between auditory and visual perception. *Perception & Psychophysics* 21, 569–574.
- Mayer, M.J., 1982. A mobile research laboratory. *Behavioral Research Methods & Instrumentation* 14, 505–510.
- Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology* 14, 247–279.
- Moore, J.J. and D.W. Massaro, 1973. Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology* 99, 49–54.
- Moray, N., M. Fitter, D. Ostry, D. Farveau and V. Nagy, 1976. Attention to pure tones. *Quarterly Journal of Experimental Psychology* 28, 271–283.
- Neisser, U., 1976. *Cognition and reality*. San Francisco, CA: W.H. Freeman.
- Neumann, O., 1985. 'Beyond capacity: a functional view of attention'. In: H. Heuer and A.F. Sanders (eds.), *Tutorials on perception and action*. Hillsdale, NJ: Erlbaum. (In press.)
- Oden, G.C., 1977. Integration of fuzzy logical information. *Journal of Experimental Psychology: Human Perception and Performance* 3, 565–575.
- Oden, G.C. and D.W. Massaro, 1978. Integration of featural formation in speech perception. *Psychological Review* 85, 172–191.
- Ostry, D., N. Moray and G. Marks, 1976. Attention, practice, and semantic targets. *Journal of Experimental Psychology: Human Perception and Performance* 2, 326–336.
- Posner, M.I., 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology* 32, 3–25.
- Posner, M.I., C. Snyder and B. Davidson., 1980. Attention and the detection of signals. *Journal of Experimental Psychology: General* 109, 160–174.
- Raab, D.H., 1962. Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences* 24, 574–590.
- Schneider, W. and R.M. Shiffrin, 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review* 84, 1–66.
- Selfridge, O.G., 1959. 'Pandemonium: a paradigm for learning'. In: *Symposium on the mechanisms of thought processes*. London: Her Majesty's Stationary Office.
- Shaw, M.L., 1984. 'Division of attention among spatial locations: a fundamental difference between detection of letters and detection of luminance increments'. In: H. Bouma and D.G. Bouwhuis (eds.), *Attention and performance X: control of language processes*. Hillsdale, NJ: Erlbaum.

- Sorkin, R.D., L.E. Pohlmann and D.D. Woods, 1976. Decision interaction between auditory channels. *Perception & Psychophysics* 19, 290–295.
- Sperling, G., 1960. The information available in brief visual presentations. *Psychological Monographs* 74, (11, Whole No. 498).
- Sperling, G., 1984. 'A unified theory of attention and signal detection'. In: R. Parasuraman and D.R. Davies (eds.), *Varieties of attention*. New York: Academic Press.
- Sperling, G. and M.J. Melchner, 1978. 'Visual search, visual attention, and the attention operating characteristic'. In: J. Requin (ed.), *Attention and performance VII*. Hillsdale, NJ: Erlbaum.
- Swets, J.A., 1984. 'Mathematical models of attention'. In: R. Parasuraman and D.R. Davies (eds.), *Varieties of attention*. New York: Academic Press.
- Treisman, A., 1969. Strategies and models of selective attention. *Psychological Review* 76, 282–299.
- Treisman, A. and G. Gelade, 1980. Feature-integration theory of attention. *Cognitive Psychology* 12, 97–136.
- Treisman, A., D. Kahneman and J. Burkell, 1983. Perceptual objects and the cost of filtering. *Perception & Psychophysics* 33, 527–532.
- Tsal, Y., 1983. Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance* 9, 523–530.
- Van der Heijden, A.H.C., R. Hagenaar and W. Bloem, 1984. Two stages in postcategorical filtering and selection. *Memory & Cognition* 12, 458–469.
- Zadeh, L.A., 1965. Fuzzy sets. *Information and Control* 8, 338–353.