

# 11

## Capacity Limitations in Auditory Information Processing

Dominic W. Massaro

*Department of Psychology  
University of Wisconsin  
Madison, Wisconsin, United States*

### ABSTRACT

An information processing model serves as a heuristic to understand capacity limitations and attentional effects in auditory-processing tasks. Experiments in backward recognition masking show no selective perception along the dimensions of spatial location, loudness, or sound quality. Some capacity limitation is shown, however, when an auditory perception task must be performed concurrently with visual recognition. Selective attention also occurs when observers are required to integrate and count a sequence of sounds although these same sound sequences can be monitored on a perceptual level without an attentional effect. Attentional effects in processing dichotic speech inputs are also described in terms of the model.

### I. INTRODUCTION

The goal of this chapter is to utilize an information processing model in order to understand capacity limitations and attentional effects in auditory processing tasks. This approach rests on the idea that a model of attention can only follow rather than precede models of information processing. Accordingly, we begin by describing the development of an auditory processing model and the empirical support for the model. We then consider the implications of the model for attentional effects at successive stages of auditory processing.

Figure 1 presents a flow diagram of the auditory processing model utilized in our research. Auditory input is funneled through the ears so that although the



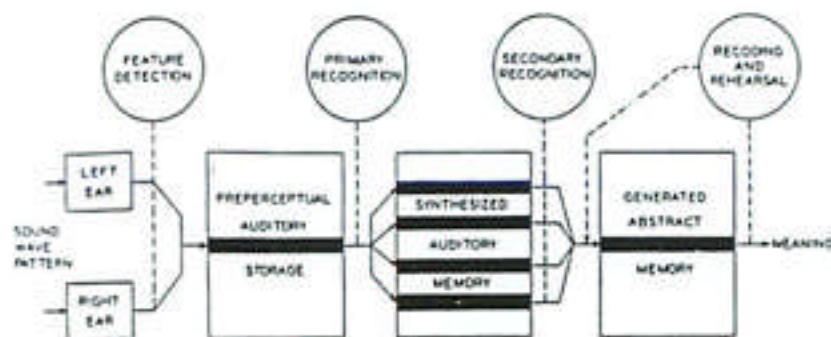


FIG. 1 Successive stages of auditory information processing.

atmosphere is filled with sound, our contact with it is limited to these two organs. The transduction of the sound-pressure fluctuations passes on featural information which is detected and stored centrally in preperceptual auditory storage. Preperceptual auditory storage holds the featural information given by the detection process for a short time after a stimulus is presented. The primary recognition process involves a resolution of this information, producing a synthesized percept held in synthesized auditory memory. Primary recognition accomplishes the phenomenological outcome of perceiving sound of a definite loudness and quality at some location in space. Preperceptual auditory storage appears to be centrally located with a capacity of just one sound; a second sound appears to replace the featural information stored there about the earlier sound. The information about the first sound passed on by primary recognition is limited to that processed between the onset of the first sound and the onset of the second. The quality of the information in synthesized auditory memory is, therefore, a direct function of the processing time between the onsets of the two sounds.

## II. RECOGNITION MASKING

Evidence for this conceptualization of preperceptual auditory storage and primary recognition comes from a number of results in recognition masking. Massaro, Cohen, and Idson (1976) provided a direct comparison of forward and backward recognition masking. On each trial the subject was required to identify a sine-wave test tone as high or low. The high and low test tones were 860 and 790 Hz, respectively. In forward masking, the test tone was preceded by a masking tone, whereas the masking tone followed the test tone in backward masking. The masking tone was either high (900 Hz), middle (825 Hz), or low (750 Hz). The test and masking tones were 20 msec in duration and were presented at a normal listening intensity.

Each trial began with the visual presentation of a Cue 1 or 2, signifying whether the test tone would be presented first or second. The cue was presented 1 sec before the onset of the test tone presentation. On masking trials, the test tone was preceded or followed by a masking tone with a 0-, 20-, 40-, 80-, 160-, 250-, or 350-msec silent interval between the test and masking tones. On 1/8 of the trials no masking tone was presented. These trials were cued with the digit "1." The subject had 1.5 sec to make a response after the offset of the test tone. Subjects pushed 1 of 2 buttons labeled "H" and "L," respectively, indicating whether the test tone was high or low. Feedback was given after this interval by illuminating for 250 msec a visual display with the symbol H or L depending on whether the high or low tone was presented. The intertrial interval was 1.5 sec.

Figure 2 presents the forward and backward masking functions for eight subjects. The results show an insignificant amount of forward masking at silent intervals longer than 20 msec whereas significant backward masking occurs out to intervals 250–350 msec. Asymptotic performance under forward and backward masking averaged about 5% less than performance under the no-mask condition. This result contrasts with a number of previous results showing that backward masking at 250 msec is equivalent to the no masking case (Hawkins & Presson, this volume; Massaro & Cohen, 1975; Massaro & Kahn, 1973). One explanation might be that the subjects missed or forgot the visual cue on a small proportion of trials in this study. When this occurred on masking trials, the subject might have categorized the masking tone rather than the test tone. This mixup could not occur on no-masking trials, leading to the slight asymptotic advantage of the no-mask condition.

The results given in Fig. 2 can be interpreted in terms of the properties of preperceptual auditory storage and the primary recognition process. Recognition of a short sound may not be complete at the end of the stimulus and can continue to occur during the silent period after the sound presentation. Backward masking occurs because a second sound interferes with the recognition process by replacing the first sound in preperceptual auditory storage. Very little forward masking is found since the test sound replaces the masking sound in preperceptual storage and recognition can continue during the silent interval after the test sound presentation.

The recognition masking results appear to require some assumption of a preperceptual storage before recognition has taken place. Crowder (1975) suggested that backward masking may occur because subjects confuse the order of the test and masking tones and judge the pitch of the masking rather than the test tone. Increasing the silent interval between the tones decreases the likelihood of the temporal confusion and, therefore, performance will improve. Crowder's temporal confusion hypothesis predicts that backward and forward masking should be equally effective since temporal confusions must be bidirectional. The large asymmetry in backward and forward masking in Fig. 2 argues against a simple temporal confusion hypothesis.



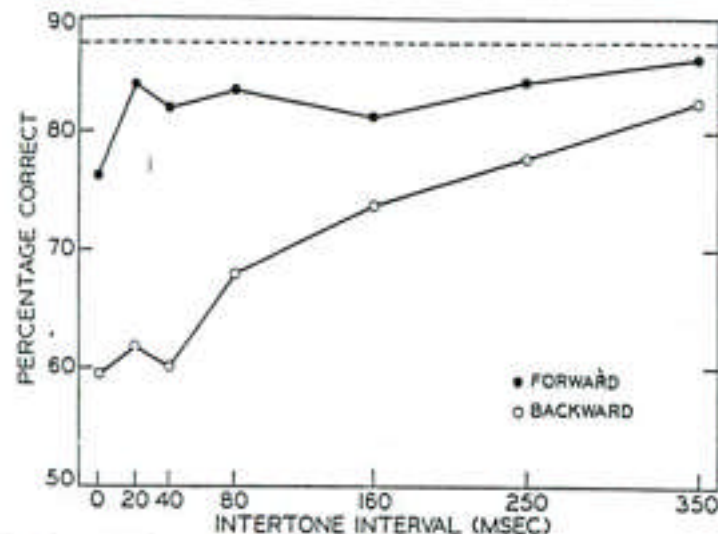


FIG. 2. Percentage of correct recognitions as a function of intertone interval under forward and backward masking. The dotted line gives performance when no masking tone was presented. (After Massaro, Cohen, & Idson, 1976.)

### III. SELECTIVE PERCEPTION

The following experiments asked to what extent capacity limitations and attentional effects occur at the initial stage of primary recognition. The experiments are carried out in a backward recognition masking task for a number of reasons. Since the masking stimulus terminates processing of the test stimulus, attentional effects can be measured at different processing times of the test stimulus. This procedure tests for attentional effects across the whole range of performance levels, allowing the experimenter to avoid floor and ceiling effects. The significant improvement in performance with processing time will provide a measure of the dynamic aspect of attentional effects if they are found, and will help silence potential criticisms of an inadequate test if no attentional effects are observed. If attentional effects are shown to interact with processing time, the results can be safely located at the primary recognition, rather than some later stage of information processing.

#### A. Spatial Location

The results of recognition masking support the idea that a second sound can replace an earlier sound held in preperceptual auditory storage. There is also good evidence that the storage is located centrally after the inputs from the two ears are combined in the auditory system. (see Hawkins & Presson, this volume;

Massaro, 1975b; Massaro & Cohen, 1975). A masking tone presented to the opposite ear of a test tone presentation produces as much interference as one presented to the same ear as the test tone.

If the structure of preperceptual auditory storage is a single channel that receives inputs from both ears, primary recognition should not be capable of selective attending to one spatial location or blocking out another. To test this hypothesis, subjects were asked to recognize a test tone as high or low (Massaro, 1975a). The test tone was a 20-msec sine wave of either 720 or 780 Hz. The test tone could be presented to either of the two ears followed after a variable silent interval by a masking tone presented to the opposite ear. The masking tone was a 100-msec square wave tone of 750 Hz. Before each trial, subjects were cued that the test tone would be presented to the right, to the left, or to either ear. These practiced subjects also knew that the masking tone would be presented to the opposite ear. Therefore, given the cues right or left, subjects attempted to attend to the designated ear and to block out the unattended ear. Given the either cue, however, subjects had to divide their attention between the two ears as the test tone could occur on either ear.

If subjects can selectively attend to spatial location at the level of primary recognition, we should find significantly less masking in the selective than the divided attention conditions. On the other hand, the spatial location of a sound may not be available until primary recognition of spatial location has occurred. In this case, sounds coming in different ears cannot be differentiated and

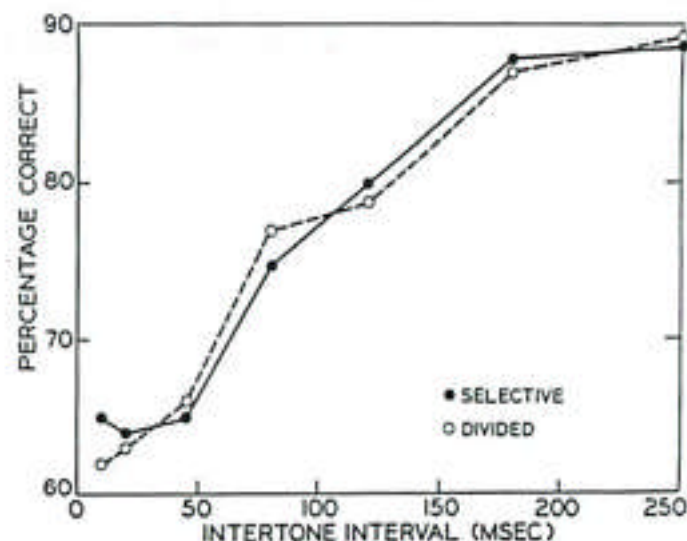


FIG. 3. Percentage of correct recognitions as a function of intertone interval under selective and divided attention. (After Massaro, 1975a.)



selective attention to spatial location should not be possible. Therefore, primary recognition of the pitch of the test tone should not be facilitated by a precue indicating the spatial locations of the test and masking tones. Figure 3 plots performance under the selective- and divided-attention conditions as a function of the intertone interval. Supporting earlier dichotic masking studies, pitch recognition improved with increases in processing time before the masking tone. And in agreement with the experimental hypothesis, no differences were found between the selective- and divided-attention conditions.

### B. Loudness and Sound Quality

Spatial location information is not immediately available at the level of preperceptual storage but must be resolved by primary recognition. Similarly, loudness and sound quality are not differentiated at the level of preperceptual storage. Moore and Massaro (1973) asked whether observers could selectively process one dimension of an auditory stimulus. On each trial, subjects identified either the loudness, sound quality, or both dimensions of a short test tone in a backward recognition masking task. Performance improved with increases in the silent intertone interval, but was not dependent on whether subjects identified one or two dimensions of the test tone. This result shows that the primary recognition process cannot selectively attend to certain features of sound quality or loudness to improve its performance relative to the situation in which both dimensions must be identified.

The results of Massaro (1975a) and Moore and Massaro (1973) show that subjects may not be able to selectively attend to an auditory dimension of spatial location, loudness, or sound quality at the level of primary recognition. The reason for this, in our model, is that these dimensions are not differentiated at the level of preperceptual auditory storage. An alternative explanation would be that the dimensions are differentiated but that the primary recognition process is automatic and unlimited in capacity. This hypothesis also predicts that multiple dimensions can be processed in parallel without a performance deficit. One way to distinguish between these two explanations is to measure auditory recognition when subjects are required to perform some other nonauditory task. If auditory recognition were unlimited in capacity, a second task should not limit its efficiency. Given a limited capacity, however, a second task might interfere with auditory recognition even though subjects cannot selectively attend to certain dimensions of a sound at the level of primary recognition.

### C. Auditory and Visual Recognition

Massaro and Kahn (1973) asked whether auditory recognition could be disrupted by an additional visual processing task. Simultaneously identifying the duration of a light decreased tone identification relative to the selective atten-

tion case in which just the tone had to be processed. This result is consistent with the idea that the process of primary recognition itself is limited in capacity and identifying the duration of the light can subtract from the process capacity available for primary auditory recognition. Within the process of auditory recognition, however, no further division of processing capacity may be possible because of the structure of preperceptual auditory storage. In this case, it may not be possible to selectively process one auditory dimension at the expense of another.

Shiffrin and Grantham (1974) have argued that duration judgments require short-term memory and decision capacity which cannot be divided without loss. Given the difficulty of the duration task, processing the duration of the visual stimulus may delay the decision about the pitch of the test tone until the appropriate information is lost. It is difficult to see, however, why the subject would forget the pitch of the test tone since it can be categorized as high or low. If making two decisions instead of just one produces a short-term memory loss, we might also have expected an attentional effect in the Moore and Massaro (1973) study.

If one argues that recognition is automatic and does not have capacity limitations, any deficit observed in some divided-attention situation might be claimed to occur at some later stage of processing. The evidence indicates, however, that recognition masking measures the temporal development of auditory recognition and does not depend on decision or short-term memory limitations (Hawkins & Presson, this volume; Massaro, 1976a). Accordingly, the recognition masking paradigm is ideal for studying attentional effects and capacity limitations at the primary recognition stage of processing. If some attentional manipulation is observed to interact across the temporal course of recognition, we are somewhat confident in assuming that its effect is occurring at the recognition stage of processing. Shiffrin (1976) agrees that backward masking occurs before short-term memory; if capacity limitations or attentional effects interact with backward masking, their effect should be located at the recognition, not the short-term memory, stage of processing.

Massaro and Warner (unpublished manuscript, 1976) asked whether primary recognition was limited in capacity. Subjects were asked to selectively attend to a visual recognition task, an auditory recognition task, or to divide their attention between the two tasks. The visual task required the subject to recognize a letter as "U" or "V." The subjects identified the test tone as high or low in the auditory task. Both the visual and auditory test stimuli were followed by masking stimuli after a variable interval. Each trial was initiated with a visual cue indicating whether the subject should identify the tone, the letter, or both the tone and the letter. All other experimental conditions were exactly analogous to the Moore and Massaro (1973) study.

Figure 4 presents the results of four subjects under the selective- and divided-attention conditions. Performance improved with increases in the processing



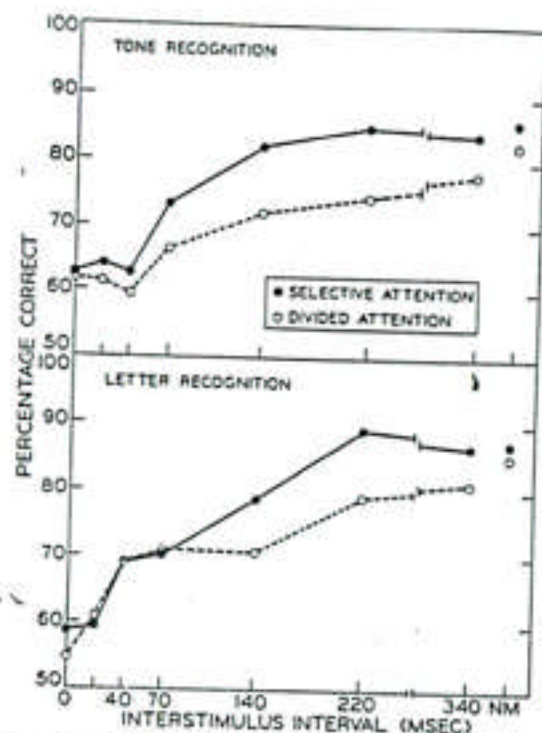


FIG. 4 Percentage of correct recognitions as a function of interstimulus interval under selective and divided attention. (After Massaro & Warner, unpublished manuscript, 1976.)

interval between the test and masking stimuli in both letter and tone recognition and under selective and divided attention. Overall performance was better in the selective- than the divided-attention condition in both visual and auditory recognition. These results provide some support for the hypothesis that auditory recognition may be limited in capacity. Even though auditory recognition has a capacity limitation, however, it may not be possible to selectively enhance processing of one auditory dimension at the expense of processing another dimension.

#### IV. MONITORING AND COUNTING SOUNDS

We have argued that, although primary recognition may be limited in capacity, it is not possible to selectively attend to certain auditory dimensions at this stage of processing. The outcome of primary recognition is the perception of some sound at some location in space. Synthesized auditory memory holds the

outcome of primary recognition. The storage capacity of synthesized auditory memory is larger than preperceptual auditory storage. Synthesized auditory memory can hold a number of items over the time span of seconds. Given that a sound is stored in synthesized auditory memory with perceptual information about spatial location, loudness, and pitch quality, it should be possible to selectively attend to some dimension at this stage of processing. Secondary recognition involves an analysis for meaning and this analysis should be dependent on the structural properties of the information in synthesized memory. The following experiments were aimed at showing that, although selective attention to spatial location and tone frequency are not possible at primary recognition, these dimensions can be selectively attended to at the level of secondary recognition.

##### A. Spatial Location

Previous studies of primary recognition have required subjects to process just a single test tone. In order to compare primary and secondary recognition, we asked observers to process sequences of test tones. First, the experiment on selective attention to spatial location (Fig. 3) was extended to monitoring a sequence of test tones (Massaro, 1976b). Subjects heard a sequence of 5, 6, 7, or 8 20-msec tones. The tones in a sequence were either presented to the same ear or alternated between the ears. The tones could be presented at any of 8 rates of presentation ranging between 20 and 3.5 per second. All of the tones but one were 800 Hz. The frequency of the different tone, called the probe tone, was either slightly higher or lower than 800 Hz. The task was to indicate whether the probe tone was higher or lower than the other tones in the sequence regardless of the rate of presentation or whether the tones alternated between ears. The subjects were practiced and were given feedback after each trial.

Subjects should be able to perform this task at the level of primary recognition. The sequence of tones can be monitored by listening for the higher or lower probe tone. If subjects cannot selectively attend to spatial location at this stage of processing, no decrement in performance should be observed when the tones are alternated between ears relative to being presented to the same ear.

The bottom panel of Fig. 5 plots the percentage of correct recognitions of the frequency of the probe tone as a function of the rate of presentation and whether the tones are presented to the same ear or alternated between ears. The rate of presentation is described in terms of processing time defined as the time between the onsets of the successive 20-msec tones. Performance improved about 20% with increases in the silent interval between successive test tones. This result demonstrates that backward recognition masking also occurs with a sequence of successive tones. Each tone in the sequence functions as a masking tone terminating any further resolution of the pitch of the previous tone. More importantly, probe recognition is not disrupted when the tones are alternated



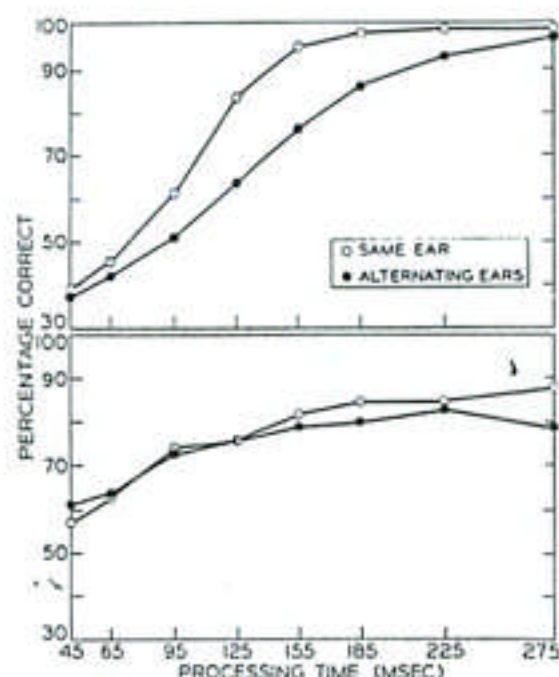


FIG. 5 Percentage of correct probe recognitions (bottom panel) and correct counts (top panel) as a function of processing time for each tone. The parameter indicates whether the tones were presented to the same ear or alternated between the ears. (After Massaro, 1976b.)

between the ears relative to being presented to the same ear. (The advantage of the same ear condition at the 275-msec processing interval will be discussed subsequently.)

The goal now was to take these same sequences of tones and require them to be processed at the level of synthesized auditory memory and secondary recognition. The task we chose was a counting task. The sequences were the same as in the probe monitoring task except that all tones were 800 Hz. Accordingly, subjects now heard a sequence of 5, 6, 7, or 8 20-msec tones, reported which of the four sequences occurred, and were informed of the number of tones actually presented. The top panel of Fig. 5 presents the percentage of correct counts as a function of processing time under the two types of presentation. Counting performance improved with increases in processing time but at a much faster rate when the tones were presented to the same ear than when the tones were alternated between the ears. The deficit in counting performance when the tones are alternated between the ears is consistent with the idea that perceptual dimensions at the level of synthesized auditory memory can be utilized for

selective processing during secondary recognition. Given that no decrement was found in the monitoring task, alternating the tones between the ears did not interfere with the perceptual resolution of each of the test tones. The decrement in the counting task with alternating tones must, therefore, mean that subjects have trouble counting or integrating sequences of tones that are perceived at different locations in space. Harvey and Treisman (1973) reached similar conclusions from an analogous set of studies.

## B. Frequency

Two additional experiments explored the possibility that the frequency dimension would operate similar to spatial location. These experiments replicated the previous two experiments, but the 20-msec tones were presented binaurally, and were either presented at the same frequency or alternated between two different frequencies. The different frequencies were slightly over one octave apart; the notes A<sub>4</sub> (440) and B<sub>3</sub> (988 Hz) were used.

The monitoring task required subjects to listen for a probe tone that was either slightly longer or shorter than the other 20-msec tones in a sequence. (A change in duration produces a timbre difference.) On each trial, subjects indicated whether the probe tone was longer or shorter and were given feedback on their decision. The bottom panel of Fig. 6 shows that recognition improved with increases in processing time and was somewhat poorer at the longer processing intervals when the tones alternated in frequency. This deficit, which contradicts the experimental hypothesis is discussed subsequently.

The counting task required the subjects to count the tones in the sequence regardless of whether they were presented at the same frequency or alternated between frequencies. The top panel of Fig. 6 plots counting performance as a function of the rate of presentation under the same and alternating frequency conditions. Counting performance was significantly poorer when the tones alternated in frequency relative to being presented at the same frequency but this effect did not emerge until the tones were presented at 8 per second or slower. Analogous to the experiment in which tones can alternate between the ears, subjects show a deficit in counting sounds that have been presented at alternating frequencies.

The monitoring results in Figs. 5 and 6 show a slight decrement at the slowest rate of presentation in the alternating relative to the same ear or frequency condition. This result shows that subjects may have switched to a comparison judgment between successive test tones at the slowest rate of presentation. In this case, the subjects would be at a disadvantage in the alternating relative to the single frequency or ear condition. Consider the monitoring experiment in which the tones could alternate in frequency. The change in duration of the probe tone produces a change in timbre and the timbre also depends on the frequency of the probe tone. To perform the monitoring task correctly, subjects



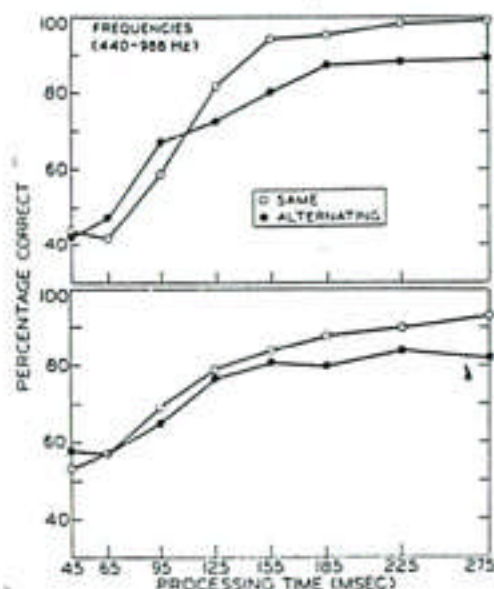


FIG. 6 Percentage of correct probe recognitions (bottom panel) and correct counts (top panel) as a function of processing time for each tone. The parameter indicates whether the tones were presented at the same frequency or alternated between frequencies. (After Massaro, 1976b.)

would have to remember the timbre of the longer and shorter probe tones at both frequency values. If subjects made successive comparison judgments, however, it would be easier to compare the timbre of successive notes when the notes are at the same frequency than when the notes alternate in frequency. An analogous argument can be made when the tones alternate between the ears since a tone of a given frequency may give a different pitch impression to one ear than to another. This argument is not critical for our purposes since no monitoring deficit was observed at those rates of presentation that gave the largest counting deficit.

Primary recognition involves the readout of a single preperceptual store that receives information from both ears and different frequencies. In the monitoring task, subjects can evaluate the output of primary recognition to observe whether a probe frequency or duration was presented. The critical variable is the processing time available for primary recognition. Given the central preperceptual store, primary recognition shows very little decrement when sounds are alternated between the two ears or between different frequencies. Secondary recognition, on the other hand, is necessary to abstract meaning from synthesized auditory memory. Given that sounds are perceived at different spatial locations or different pitches in the alternating conditions, secondary recogni-

tion has more difficulty integrating and counting these sounds than sounds perceived at the same pitch and spatial location (cf. Fig. 1). This analysis illuminates the differences found in the monitoring and counting experiments and provides a processing model of the stages of processing at which selective attention will and will not occur.

Information is most easily processed along a temporal dimension in synthesized auditory memory. The perceptual dimensions such as spatial location or pitch cannot be used to organize auditory inputs. If the inputs from the different ears could be processed by location rather than by temporal order, then we wouldn't have expected such a substantial counting deficit in the alternating ear condition. Subjects would have been able to count the number of tones along one location and then those along the other location. Evidently, synthesized auditory memory did not maintain all of the items along their spatial locations for the 1 or 2 seconds necessary to perform the task in this way.

## V. SPEECH PROCESSING

The tone experiments and this analysis have implications for the processing of speech stimuli presented simultaneously to the two ears. Previous investigators have assumed that the analysis for meaning of a list of items presented to one ear can be delayed for 1 or 2 seconds until the list on the other ear is processed (Broadbent, 1958, 1971; Darwin, Turvey, & Crowder, 1972). One test of the capacity and structure of synthesized auditory memory has been to present simultaneous lists of items to different spatial locations preceded or followed by a probe cue asking the subjects to recall some subset of the items. If the items are indeed maintained along separate locations in synthesized auditory memory and semantic analysis is delayed for 1 or 2 seconds, cuing recall of the items according to spatial location should produce better performance than cuing recall along some semantic dimension such as category name. Semantic analysis is not delayed for 1 or 2 seconds in our model but occurs on the order of every syllable (see Massaro, 1975c). Therefore, we predict that spatial location will not be a more effective partial report cue than category name even if the cue is given immediately after the test lists.

In one experiment (Massaro, 1976c), two lists of four items each were recorded by different speakers at a rate of 4 items per second. These lists were presented simultaneously to the two ears. Each list on each ear contained two one-syllable words and two letters chosen randomly without replacement from respective master sets of 25 one-syllable words and 25 letters. When subjects were asked to recall by location the cue indicated whether they should report the items presented to the left or right ears. When subjects were asked to recall by category, the cue indicated whether they should report the words or letters. The exact same lists and cues (pure tones) were used in both report conditions.



The cue was presented immediately after the last items in the lists. In both report conditions there were two possible report cues and the subjects were required to report four items. Therefore, if the items are maintained along separate auditory dimensions according to spatial location and semantic analysis is delayed for 1 or 2 seconds, recall cued by location should be superior to recall cued by category name.

The results showed no advantage of a report by spatial location over that by category name when the cue was given immediately after the test list (Massaro, 1975b, 1976c). This result shows that the separate lists are not maintained for 1 or 2 sec in a precategorical form according to spatial location of presentation. The spatial location of an item does not appear to be preserved in a precategorical form but must be remembered by way of association (the letter "E" comes from the left).

On the other hand, we might expect that recall by spatial location would be much better if the report cue is given before the test lists. In this case, subjects can utilize the report cue during the processing of the lists. Figure 7 presents the mean number of items recalled when the location and category cues were given before or after the test list. Recall by spatial location was much better than

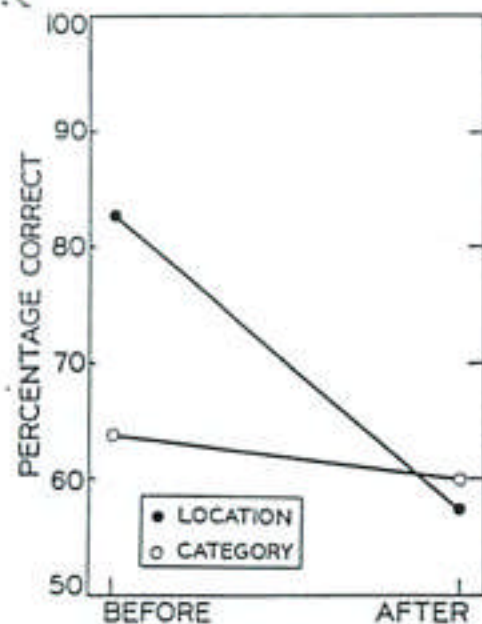


FIG. 7 Percentage of items correctly recalled as a function of whether the report cue was given seconds before or immediately after the test list. The parameter specifies whether recall was cued according to spatial location or category name. (After Massaro, 1976c.)

recall by category name when the cue was given before the list presentation. Delaying the report cue until immediately after the test list decreased recall by spatial location by 25% but had an insignificant effect on recall by category name. Replicating the earlier study, no difference in recall occurred when the cues were given after the test list.

The results of the partial report experiments clarify the temporal course of successive stages of auditory information processing. Simultaneous items are first stored centrally in preperceptual storage. Primary recognition resolves the items at a perceptual level and locates them at different locations in space. Secondary recognition then processes the items for meaning. If the report cue for spatial location is given before the test list, secondary recognition can devote most of its processing capacity (but not all of it) to items perceived at a given location in space. In recall by category name, the cue before the list does not reduce the processing required by secondary recognition. Each item must still be processed for meaning to determine if it belongs to the appropriate category for recall. Accordingly, more items can be reported correctly in recall by spatial location than in recall by category name when the cue is given before the test list. If subjects are not cued until after the test list, the items have been processed by both primary and secondary recognition. In this case, retrieval by spatial location no longer shows an advantage over recall by category name.

## VI. SUMMARY AND CONCLUSION

A simplified version of the model proposed earlier is shown in Fig. 8. The model distinguishes between three stages (levels) of auditory processing: detection, primary recognition, and secondary recognition. The detection of sound refers to the listener's experience that some sound was present. In this case, detection means the observer is able to state when sound is present or absent. After a sound is detected, the listener does not necessarily have any information about

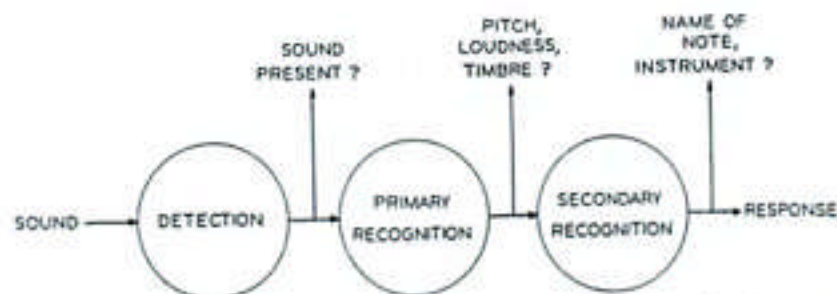


FIG. 8 Detection, primary recognition, and secondary recognition stages of auditory processing.



the nature of that sound. Having heard a sound, primary recognition refers to the listener's ability to resolve some attribute of the sound, such as its pitch, timbre, loudness, or spatial location. Finally, the listener's ability to identify a particular pitch as middle C is called secondary recognition. It is assumed that the onsets of these three stages of processing are successive although the stages themselves may overlap in time. Given the successive nature of detection and primary and secondary recognition, a listener might detect a tone but not resolve its pitch, or resolve its pitch but fail to identify its name. Furthermore, a later stage of processing usually does not occur unless each earlier stage is successful. For example, the pitch of a sound could not be resolved without detection of the sound.

The stage analysis appears to provide a worthwhile heuristic for studying capacity limitations and selective attention effects in auditory processing. Although primary recognition may be limited in capacity, it may not be possible to selectively attend to certain dimensions of sound at this stage of processing. On the other hand, integrating successive sounds at the level of secondary recognition is dependent on selective attention to dimensions of the sound sequence. Therefore, selective attention to some auditory dimension can facilitate processing for meaning. Given that the analysis for meaning is not delayed for 1 or 2 seconds, however, retrieval along an auditory dimension shows no advantage over retrieval along a semantic category when the retrieval cue is given immediately after the auditory input.

#### ACKNOWLEDGMENTS

The preparation of this paper was supported by National Institute of Mental Health Grant MH-19399 and a grant from the Wisconsin Alumni Research Foundation.

#### REFERENCES

- Broadbent, D. E. *Perception and communication*. New York: Pergamon Press, 1958.
- Broadbent, D. E. *Decision and stress*. London: Academic Press, 1971.
- Crowder, R. G. Inferential problems in echoic memory. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. London: Academic Press, 1975.
- Darwin, C. J., Turvey, M. T., & Crowder, R. G. An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, 1972, 3, 255-267.
- Harvey, N., & Treisman, A. M. Switching attention between the ears to monitor tones. *Perception & Psychophysics*, 1973, 14, 51-59.
- Hawkins, H. L., & Preston, J. C. Masking and perceptual selectivity in auditory recognition. This volume.
- Massaro, D. W. Backward recognition masking. *Journal of the Acoustical Society of America*, 1975, 58, 1059-1065. (a)