AUDITORY MEMORY[*]

DIANA DEUTSCH[†]

University of California, San Diego

ABSTRACT

Current views of auditory memory are examined. The sensory buffer hypothesis is critically reviewed, and it is argued instead that auditory memory is the function of a system whose various subdivisions differ considerably in the persistence with which they retain information. It is further argued that, rather than being susceptible to general influences such as capacity limitation, acoustic memory is subject to highly specific interactive effects which vary systematically as a function of the relationships between the elements involved.

PSYCHOACOUSTICAL RESEARCH has accumulated a large body of evidence concerning the characteristics of the system that processes incoming auditory information. In contrast, very little is known about the system which retains such information once it has arrived. It is generally supposed that non-verbal information is held only transiently in a buffer store, while it awaits verbal encoding. The form in which this hypothesized buffer store is organized has received very little attention; the assumption usually being made that information is here subject to general influences such as decay (Broadbent, 1958), displacement on a simple first-in, first-out basis (Crowder & Morton, 1969), or channel capacity limitation (Posner, 1967).

This article is concerned with models of memory for non-verbal auditory information. The argument is made that, although some acoustic attributes may be retained only transiently, others are held for very long periods in the absence of verbal encoding. Further, it is argued that previous models made use of notions of categorical or precategorical information storage that are quite ambiguous. The more specific hypothesis is advanced that memory for an acoustic attribute is the function of an array whose elements are activated by specific values of this attribute. Interactions take place along this array which vary systematically as a function of the relationships between the interacting elements.

THE SENSORY BUFFER HYPOTHESIS AND THE DURATION OF AUDITORY MEMORY

The first detailed version of the sensory buffer hypothesis was suggested by Broadbent (1958), who proposed that information is first stored in a

large capacity sensory system (the S-system) where it decays very rapidly. The material is saved from obliteration only by verbal encoding, which allows it to enter short term memory (the P-system). Broadbent did not propose that independent systems exist for the storage of information arriving via different sensory modalities. However, later theorists (Sperling, 1963, 1967; Atkinson & Shiffrin, 1968; Crowder & Morton, 1969) amended Broadbent's general scheme to assume the existence of modality-specific stores.

Arguments for the sensory buffer hypothesis are based both on introspective and experimental evidence. Subjectively, there is a clear qualitative distinction between memory for events that occurred within a few seconds and for those that occurred less recently. The auditory image particularly possesses an initial richness and specificity which is lost rapidly, and memory for an acoustic stimulus appears relatively schematic after some time has elapsed. Neisser (1967) has aptly described the period of a few seconds following auditory stimulus presentation as the period of "echoic memory." Proponents of the sensory buffer hypothesis argue that auditory information is retained in a separate, large capacity store during this period; and that the residue which remains later has been verbally encoded and resides in a limited capacity short-term memory store. But this is not a necessary conclusion. One might alternatively suggest that a complex acoustic stimulus is represented in memory as a set of features, which differ in their decay characteristics and susceptibility to interference. The subjective richness of acoustic memory during the first few seconds following stimulus presentation reflects the presence of features that deteriorate rapidly, leaving a smaller set of more enduring features to represent the stimulus over a longer time span. The advantage of this alternative hypothesis is that it does not require the assumption that information is transferred from one store to another.

Experimental evidence for a rapidly deteriorating acoustic trace has been provided in various types of study. One group has been concerned with memory for acoustically presented verbal materials. For instance, when memory is tested for a list of verbal items presented either acoustically or visually, substantially fewer errors occur for items at the end of the list with acoustic than with visual presentation (Corballis, 1966; Murdock, 1967). Similarly, when a list of items is presented visually, vocalization at presentation favours recall, but again at the later serial positions (Murray, 1966; Conrad & Hull, 1968; Crowder, 1970). Other experiments confirm the superiority of acoustic over visual input modality during the first few seconds following stimulus presentation (Cooley & McNulty, 1967; Craik, 1969; Murdock & Walker, 1969; Morton, 1970; Morton, Crowder, & Prussin, 1971; Crowder, 1970, 1971, 1972). When a list of verbal items is presented

acoustically, recall is disrupted if the experimenter emits a single extra item at the end of the list, even though this item is redundant. The disruptive effect of the suffix is greatest at the terminal portions of the serial position curve and only occurs if the suffix is presented within one or two seconds following the memory list. Crowder and Morton interpret these results in terms of a rapidly decaying acoustic trace which is also subject to displacement by subsequent acoustic events.

Crowder and Morton (1969) have proposed that the precategorical components of acoustically presented speech are retained in a Precategorical Acoustic Store (PAS) where they decay within two seconds. However, other findings are at a variance with this view. Murdock and Walker (1969), Murdock (1971), and Murdock (1972) have shown in various studies that non-verbal speech components are preserved for at least five to ten seconds following stimulus presentation. Cole, Coltheart, and Allard (1974) found that recognition judgments for spoken letters are facilitated when the letters to be compared are spoken in the same voice rather than different voices, even with an eight-second interval between presentations. Indeed, Craik and Kirsner (1974) found that this same voice facilitation did not decline over a two-minute retention interval. Such results argue strongly that the acoustic properties of speech survive in memory for substantially longer periods than originally proposed for the PAS.

Other studies have demonstrated a rapidly deteriorating memory for certain non-verbal acoustic attributes. Guttman and Julesz (1963) showed that the repetition of a segment of white noise becomes difficult to detect when the segment is more than a second in length. Pollack (1972) has demonstrated that recognition of a sequence of interval-pulse coded stimuli is optimal with a gap of .5 seconds between sequences, and deteriorates when this gap is lengthened. Treisman and Rostron (1972) presented two sets of tonal stimuli in rapid succession, and required subjects to judge for each set independently whether a probe tone presented after a delay had been included in the set. They found that performance deteriorated to an asymptote within two seconds (though it is unclear whether this deterioration was due to a loss of item or order information). Further, although memory for the pitch of a single tone decays only slowly with time (Harris, 1952), the interpolation of a set of tones during a retention interval of only a few seconds will produce severe memory deterioration (Deutsch, 1970).

These experiments all demonstrate that certain acoustic attributes deteriorate rapidly in memory, and may be cited in support of the sensory buffer model. However, they are also consistent with the alternative hypothesis that acoustic attributes vary in their rates of decay or suscepti-

bility to interference. On this view, the above studies are simply demon-
strations of the transience in memory of certain specific acoustic attri-
butes. The example of music may be cited in further support of this argu-
ment. It is clear from general considerations that music must involve a
heterogeneous memory system, whose subdivisions vary considerably in
the persistence with which they retain information. For instance, the
system that stores abstracted tonal relationships (forming the basis for
melody and harmony recognition) must be capable of retaining highly
specific information over long periods of time (Attneave & Olson, 1971).
This is not true, however, of the system that stores absolute pitch values,
since only rare individuals are able to reproduce or recognize a tone in
the musical scale, in spite of our daily exposure to a multitude of such
tones. We can, therefore, conclude that some non-verbal acoustic attributes
are retained in transitory fashion, and yet others are held very securely
in memory. Paivio (1969) and Shepard (1967) also argue from visual
studies that we must be capable of long-term non-verbal information
storage.

THE FORM OF AUDITORY INFORMATION STORAGE

Most discussions concerning the form of auditory information storage
have centred on the categorical status of the retained material. This
emphasis is also due to the sensory buffer hypothesis, which assumes that
information, after traversing the sensory pathways, is first stored transi-
ently in "raw, uncategorized" form; and later, more durably, in "cate-
gorized" form. Various arguments have been advanced to support this
hypothesis. However, the meaning of the term "categorization" has not
been clearly spelled out in these arguments; rather, they have relied on
indirect evidence, and so have neither confirmed nor refuted the hypothe-
sis in a satisfactory way.

One argument that has been made for the existence of a transient un-
categorized acoustic store involves memory for unattended input. Such
memory has been shown to deteriorate more rapidly than memory for
attended materials (Erickson & Johnson, 1964; Treisman, 1964a, 1964b).
Some have reasoned here that since these materials had not been attended
to, they could not have been categorized by the listener (Neisser, 1967;
Treisman, 1964a, 1964b). However, this argument rests on a specific
theory of selective attention, which maintains that categorization does
not occur prior to the attention process (Broadbent, 1958; Treisman,
1964a, 1964b). This theory has been challenged on experimental grounds
by Deutsch and Deutsch (1963), who proposed instead that all incoming

information, whether attended to or not, is categorized; and that selective attention results from such categorization. Since this alternative theory was presented, further strong evidence has accumulated that unattended materials are indeed categorized at high cognitive levels (Lewis, 1970; Corteen & Wood, 1972). Studies of memory for unattended input cannot therefore be taken as *de facto* evidence for uncategorized acoustic memory storage.

A further attempt to infer the existence of a transient preperceptual acoustic store was made by Massaro (1970, 1972). The argument used here was rather complicated. Massaro presented subjects with one of two tones, of 20 msec duration, which they had already learned to classify as "high" or "low" (though these were actually close in pitch). He then presented a second tone of 500 msec duration which followed the first after a varying time interval. Massaro reports that judgment of the first tone was disrupted by the second as a function of their temporal proximity, and that this interference effect was not reduced when the second tone was presented to a different ear. The disruptive effect of the second tone declined with increasing inter-tone interval, and no disruptive effect occurred when the two tones were 250 msec apart. Thus, during this 250 msec period, classification of the first tone became increasingly resistent to interference.

Massaro infers from these observations that, during the first 250 msec following its presentation, auditory information is retained in a preperceptual acoustic store where it decays during this period. Information is fed from this store into a perceptual store at a rate determined by its amount in the preperceptual store. However, this is by no means a necessary conclusion. Memory consolidation, as manifest by increased resistance to disruption, has been shown to occur over very long time periods (Deutsch & Deutsch, 1973); and in such cases we would certainly not want to infer the existence of a slowly decaying preperceptual store which gradually transfers its contents into a categorical store. Although it could still be argued that short-term consolidation is due to such a process, there are no good reasons for accepting this theoretical formulation in preference to a more simple and direct explanation in terms of a single time-dependent consolidation process.

Further, Massaro's experimental observations have been challenged by Leshowitz and Cudahy (1973) in a careful parametric study. They found that the interference tone significantly degraded performance only with a 10 msec signal delivered to the same ear. In conflict with Massaro's claim, they found no performance decrement when the second tone was presented to the contralateral ear. This is in accordance with other studies

on contralateral masking (Zwislocki, 1972). Leshowitz and Cudahy there-
fore conclude against an interpretation of this interference effect in terms
of perceptual or mnemonic processes.

Another approach to the study of very short-term acoustic storage has
been to use materials which are in essence uncategorized at the level of
stimulus generation. This was achieved by Guttman and Julesz (1963)
who juxtaposed segments of random noise where these segments were
either identical or generated independently Listeners were able to detect
an iteration of the same segment with ease when it was less than a second
in length, but found this task more difficu t with longer segments. This
study has been quoted as providing an estimate of the duration of uncate-
gorized acoustic memory. However, the basis for this inference has never
been clearly stated and is rather obscure. Perhaps it is reasoned that since
the stimulus to be recognized was here generated by a random process,
some tape recorder form of memory would be needed to store it success-
fully. However, this is not a necessary conclusion. The fact that a stimulus
consists of randomly generated components shows nothing about the
organization of the system which retains it The results of Guttman and
Julesz (1963) are therefore consistent with a variety of models of very
short-term acoustic storage.

All the above lines of argument suffer from a lack of clear statement as
to what constitutes "categorization" of a sensory stimulus, and so from a
lack of clear procedure for determining whether or not this has taken
place. Two definitions of categorization may here be advanced. The first
is the placement of the information in its appropriate position along an
appropriate stimulus dimension. Evidence for such categorization may
then be obtained by finding specific interactive effects between this and
other information as a function of their proximity along this dimension.
The second proposed definition involves the partitioning of such a sensory
dimension into discrete segments such that stimuli falling within a seg-
ment are discriminated with difficulty, but stimuli falling on either side of
a boundary between adjacent segments are discriminated more easily.
This type of categorization has been demonstrated in perception of pho-
nemes (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967;
Studdert-Kennedy, Liberman, Harris, & Cooper, 1970). It has also been
shown to occur in perception of musical intervals by trained musicians
(Burns & Ward, 1973; Siegel, Siegel, Harris, & Sopo, 1974) and of pitch
stimuli by a few such individuals, especially those with absolute pitch
(Siegel, Siegel, Harris, & Sopo, 1974).

Experimental evidence is here advanced that, over time periods of a
few seconds, acoustic information is stored categorically according to the
first of the two above definitions. These experiments show that pitch mem-

ory is subject to highly specific interactive effects, which vary systematically as a function of the pitch relationships between the interacting elements; and so argue that it is the function of a specifically and systematically organized continuum. These experiments are not addressed to the question of categorization according to the second of the above two definitions, i.e., the partitioning of the continuum into discrete segments as manifest in identification and discrimination tasks.

The first experiment explored the effect on recognition performance produced by a tone forming part of a sequence interpolated between a standard and comparison tone, as a function of its pitch relationship to the standard tone (Deutsch, 1972a). Subjects were instructed to judge whether two tones were the same or different in pitch when these were separated by a five-sec retention interval during which six other tones were played. The standard and comparison tones were either identical in pitch, or they differed by a semitone. All tones were 200 msec in duration, and were separated by 300 msec pauses, leaving a 2-sec pause before the comparison tone. In all conditions but one, there was placed in the second serial position of the intervening sequence a tone whose pitch bore a critical relationship to the pitch of the standard tone. This relationship varied in ⅙ tone steps from identity to a whole tone separation on the equal-tempered scale. The other tones in the intervening sequence were all at least 1½ tones removed from the standard tone. In the final condition, the pitch of the tone in the second serial position of the intervening sequence was chosen in the same way as were the other intervening tones. This condition thus provided a baseline for estimating the effects of the critical intervening tones. The entire group of sequences was presented in random order, with no separation by condition. Twelve subjects were selected for this experiment, on the basis of obtaining a score of at least 90 per cent correct on a short tape designed as in the baseline condition.

The results of this experiment are shown on Figure 1, which plots percentage errors in pitch comparison as a function of the pitch separation between the standard tone and the critical intervening tone. It can be seen that the error rate did indeed vary systematically as a function of the pitch relationship between the standard tone and the critical interpolated tone. When these two tones were identical in pitch, memory facilitation was produced. Errors rose progressively as the pitch of the critical interpolated tone moved away from the pitch of the standard tone, peaked at a separation of ⅔ tone, and returned roughly to baseline at a whole tone separation. The error rate in the condition where the critical interpolated tone was identical in pitch with the standard tone was significantly lower than baseline ($p < .02$, two-tailed, on a Wilcoxon test). Further, the error rates were significantly higher than baseline in the conditions where the critical
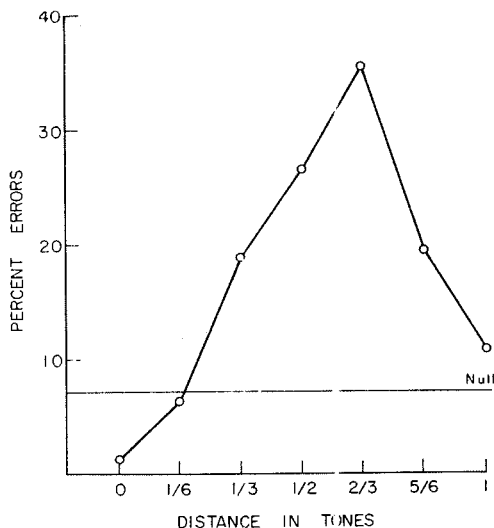
FIGURE 1. Percentage errors in pitch compari-
sons plotted as a function of the pitch separation
between the standard tone and the critical inter-
vening tone. The line labelled *Null* shows per-
centage errors in the baseline condition where
no tone closer in pitch to the standard tone than
1½ tones was included in the intervening se-
quence.

intervening tones were ⅙ tone, ½ tone, ⅔ tone, or ⅚ tone removed from the
standard tone ($p < .01$, two-tailed, on Wilcoxon tests, for all comparisons).

One must conclude that the pitch memory system is precisely and sys-
tematically organized. Further, the function shown on Figure 1 was ob-
tained by superimposing plots derived from sequences where the standard
and comparison tones were placed at different points along the pitch con-
tinuum, ranging from C_4(259 Hz) to B_4(483 Hz). Since the musical scale
is a logarithmic function of waveform frequency, an identical pitch rela-
tionship on this scale is based on an increasing difference in waveform
frequency as the scale is ascended; this difference doubling over an octave.
Thus, if the pitch memory system were arranged in some way other than
logarithmic, the peak of errors should shift systematically with a shift in
the standard tone pitch. However, no such peak shift could be discerned
in this experiment, which indicates strongly that this system is the function
of an array along which elements are tonotopically organized on a log
frequency dimension. Attneave and Olson (1971) have also argued from
general considerations of music that pitch interactions must occur along
a log frequency continuum.

Given that specific disruptive interactions take place within the pitch memory system, the question arises as to their underlying basis. A further experiment explored systematically the effects on pitch recognition of including in an interpolated sequence a tone which was a semitone removed from the standard tone (Deutsch, 1973). This experiment demonstrated the presence of at least two separable disruptive effects. First, interpolating a tone that is either a semitone higher or lower than the standard tone produces a significant disruptive effect which cumulates in size when two such tones, one a semitone higher and the other a semitone lower, are both interpolated. Second, a significantly and substantially larger disruptive effect occurs when the critical interpolated tone is identical in pitch with the comparison tone, and the standard and comparison tones differ in pitch.

It was theorized that the second of the two above effects is due to a deterioration of information along a temporal or order continuum (Deutsch, 1972*b*). Errors would then be due to the subject recognizing correctly that a tone of the same pitch as the comparison tone had occurred, but assuming incorrectly that this had been the standard tone. If this hypothesis were correct, then the amount of memory disruption produced by a tone of identical pitch with the comparison tone should vary as a function of its position in the intervening sequence. The closer the position of this tone relative to the standard tone, the more difficult it should be to discriminate their two positions, and so the greater should be the number of false recognition judgments.

This prediction was tested in the following experiment. Subjects were instructed to compare two tones for pitch when these were separated by a sequence of six intervening tones. The temporal parameters were identical with those in Deutsch (1972*a*). In half of the sequences the standard and comparison tones differed in pitch by a semitone. A tone of the same pitch as the comparison tone was interpolated either in the second serial position of the intervening sequence, or in the fifth serial position, or not at all. In the other half of the sequences the standard and comparison tones were identical in pitch. And analogously, a tone which was a semitone removed from this pitch was interpolated either in the second serial position of the intervening sequence, or in the fifth serial position, or not at all. The entire group of sequences was presented in random order, with no separation by condition. Subjects were selected for obtaining a score of at least 85 per cent correct on a small tape containing sequences in which no tone a semitone removed from the standard tone was included among the interpolated tones. Sixteen subjects were selected for this experiment, the selection rate being about one in four.

It was found that, in sequences where the standard and comparison tones differed in pitch, the critical interpolated tone produced substantially more

errors when it was positioned close to the standard tone than when it was positioned farther away. The error rates at each serial position differed significantly from baseline, and they also differed significantly from each other ($p < .01$, two-tailed, on Wilcoxon tests for all comparisons). This finding is in accordance with the hypothesis that errors result from a deterioration of temporal or order information. In contrast, in sequences where the standard and comparison tones were identical in pitch, the disruptive effect of the critical interpolated tone did not vary as a function of its serial position. However, the error rates at each serial position still differed significantly from baseline ($p < .01$, two-tailed, on Wilcoxon tests, for both comparisons). This difference in serial position effect points to a separate source of memory disruption.

A further study was performed to evaluate the hypothesis that errors in pitch comparison may be based on a deterioration of temporal or order information. Here the comparison tone pitch was varied systematically in terms of its relationship to the standard tone pitch. Within any given experimental session the standard and comparison tones were either identical in pitch, or they were separated by a constant amount, i.e., by either ⅓ tone, ½ tone, or ⅔ tone; and subjects received these three conditions in strictly counterbalanced order. Errors were plotted as a function of the pitch of a critical interpolated tone, placed in the second serial position of the intervening sequence, whose relationship to the standard tone pitch varied systematically. Whenever the standard and comparison tones differed, the critical interpolated tone was placed on the same side of the pitch continuum relative to the standard tone, as was the comparison tone; and it was therefore identical in pitch with the comparison tone whenever it bore the same relationship to the standard tone.

The temporal parameters were the same as in the previous experiments. Eighteen subjects were here selected, on the basis of obtaining a score of at least 95 per cent correct on a small tape in which the standard and comparison tones were separated by a semitone, and no tone that was less than 1½ tones removed from the standard tone was included in the intervening sequence.

It was found that, both where the standard and comparison tones were separated by ⅓ tone, and also where they were separated by ⅔ tone, errors peaked when the critical interpolated tone was identical in pitch to the comparison tone (Fig. 2). Indeed, for both the ½ tone and ⅔ tone plots, the error rate was significantly greater where the standard tone and the critical interpolated tone were separated by ⅔ tone than when they were separated by ⅓ tone. And in contrast, for the ⅓ tone plot, the error rate was significantly greater where the standard tone and the critical interpolated tone were separated by ⅓ tone than where they were separated by ⅔ tone
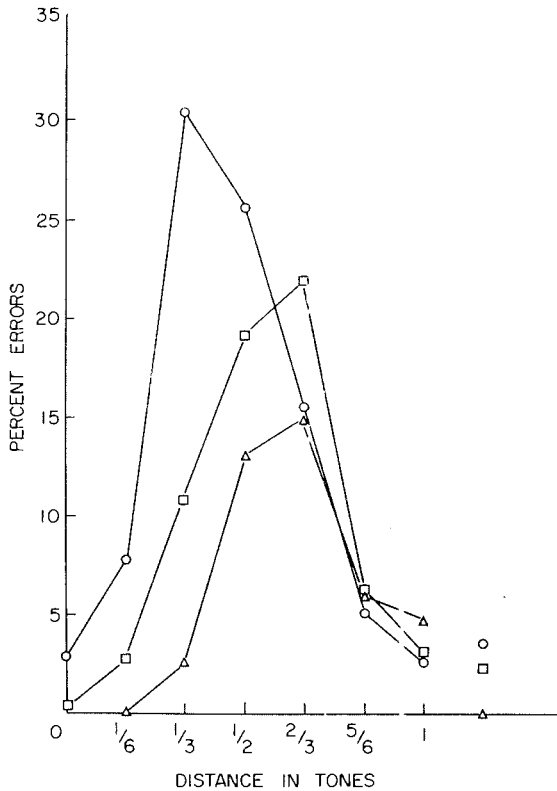
FIGURE 2. Percentage errors in pitch comparisons in
sequences where the standard and comparison tones
differed in pitch, plotted as a function of the pitch
separation between the standard and comparison tones,
and also between the standard and critical interpolated
tones. Circles plot error rates where the standard and
comparison tones were separated by ⅓ tone; squares
where they were separated by ½ tone; and triangles
where they were separated by ⅔ tone. (The separate
circle, square and triangle at right display error rates in
baseline conditions where no tone was interpolated in
the critical range under study.)

($p < .01$, two-tailed, on Wilcoxon tests, in all cases). This is in accordance
with the hypothesis that errors here were based on the correct recognition
that a tone of identical or similar pitch to the comparison tone had been
presented, but the incorrect assumption that this had been the standard
tone.

It will also be noted that in sequences where the standard and com-
parison tones were separated by ½ tone, errors peaked at a pitch separation

between the standard tone and critical intervening tone not of ½ tone, but of ⅝ tone. Though this difference is not statistically significant, it indicates that some fixed source of disruption which peaks at ⅝ tone might be superimposed on the shiftable source of disruption manifest in the other two plots.

In considering the possible basis for this fixed source of disruption, two points of interest may be noted. First, the relative frequency range over which this effect occurs corresponds well with that found for centrally acting lateral inhibition in single units of the central auditory pathway (Klinke, Boerger, & Gruber, 1969, 1970). Second, this disruptive effect cumulates when instead of one tone in the disruptive range two are presented, placed on either side of the standard tone along with the pitch continuum (Deutsch, 1973). Analogously, in lateral inhibitory networks, cumulation of inhibition occurs from stimuli placed on either side of the test stimulus (Ratliff, 1965). Psychophysical evidence for lateral inhibition has been found in the system that processes incoming acoustic information (Carterette, Friedman, & Lovell, 1969; Zwislocki, Buining, & Glantz, 1968; Houtgast, 1972). It was therefore proposed as a working hypothesis that the elements of the pitch memory system are arranged as a recurrent lateral inhibitory network, analogous to those in sensory systems. Elements of this system are activated by tones of specific pitch, and are organized tonotopically on a log frequency continuum (Deutsch & Feroe, 1975).

One prediction from the proposed model is that the pitch memory system should exhibit not only inhibition but also disinhibition. Thus in sequences where the standard and comparison tones are identical in pitch, if there were interpolated two critical tones, one ⅝ tone removed from the standard tone pitch, and the other further removed along the pitch continuum, then errors should be a function of the pitch relationship between these two tones. The error rate should be greatest when these two tones are identical in pitch, decline as the second tone moves away from the first, dip maximally at a ⅝ tone separation, and then return to baseline at around a whole tone separation.

This prediction was put to experimental test using sequences where a tone which was ⅝ tone removed from the standard tone was placed in the second serial position of the intervening sequence, and the second critical interpolated tone was placed in the fourth serial position. The temporal parameters were the same as in the previous experiments. Subjects were selected for obtaining a score at least 95 per cent correct on a tape containing sequences in which no tones were interpolated in the critical range under study. As before, the sequences were presented in random order, with no separation by condition. Twenty-one subjects were employed for this experiment, the selection ratio being about one in six.
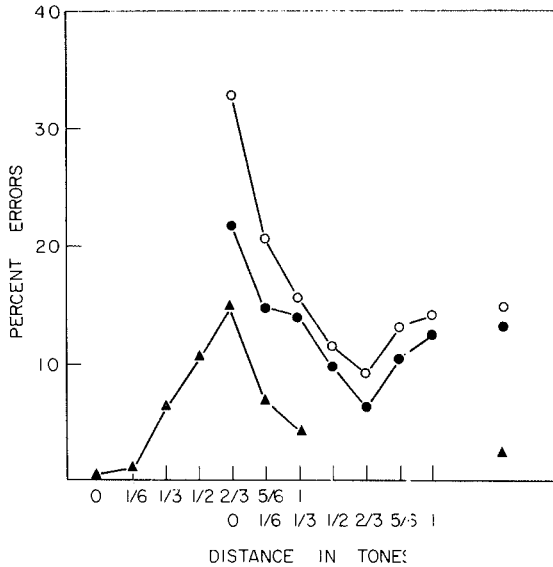
FIGURE 3. Percentage errors in pitch recognition obtained experimentally and predicted theoretically. Triangles display percentage errors in the baseline experiment which varies the pitch relationship between the standard tone and a critical interpolated tone. (Triangle at right shows percentage errors where no tones were interpolated within the critical range under study.) Filled circles display percentage errors in the experiment where a tone that is ⅔ tone removed from the standard tone is always interpolated. Errors are plotted as a function of the pitch relationship between this tone and a second critical interpolated tone which is further removed along the pitch continuum. Open circles display percentage errors for the same experimental conditions predicted theoretically from the lateral inhibition model. (Filled and open circles at right show percentage errors obtained experimentally and assumed theoretically where no further critical tone is interpolated.)

The closed circles on Figure 3 plot error rate as a function of the pitch relationship between the two critical intervening tones, in sequences where the standard and comparison tones were identical in pitch. It can be seen that the predicted disinhibition function was indeed obtained. The error rate in sequences where the second critical interpolated tone was identical in pitch with the first was significantly higher than baseline ($p < .005$, one-tailed, on a Wilcoxon test). Further, the error rate where the two critical interpolated tones were separated by ⅔ tone was significantly lower than baseline ($p < .01$, one-tailed, on a Wilcoxon test).

In order to make quantitative comparison between the disinhibition function obtained experimentally and the function predicted on the lateral inhibition hypothesis, the following model was advanced (Deutsch & Feroe, 1975). It was proposed that a set of n tones activates a corresponding set of n elements along the memory array. Then let k_s be the inhibitory coefficient between elements underlying two tones separated by a distance of s tones; let r_p be the signal strength of the element underlying the pth tone; and let e be the signal strength of an element when its corresponding tone is at least a whole tone removed from the other tones in the set. The interaction between the n elements may then be described by the n simultaneous linear equations

$$r_p = e - \sum_{\substack{q=1 \\ p \neq q}}^{n} k_{s(p,q)} r_q \quad \text{for } p = 1, 2, \ldots, n,$$

where $s(p,q)$ is the distance between the pth and qth tones. By solving these equations we can predict the net influence on the strength of the signal emitted by the element underlying a test tone exerted by the elements underlying the rest of the tones in the set.

A baseline function for the first-order inhibitory effect was obtained using subjects selected on the same criterion as for the present study. This function was used to compute values of e and k_s, assuming that error rate was a function of the signal strength of the element underlying the test tone. These parameters were then used to compute the predicted disinhibition function. Twenty-five subjects were used to obtain this baseline curve, which is plotted by the triangles on Figure 3. The predicted disinhibition function, computed from these parameters, is plotted by the open circles. It can be seen that the theoretically derived disinhibition function corresponds well with the function obtained experimentally. This experiment strongly supports the hypothesis that the elements of the pitch memory system are arranged as a recurrent lateral inhibitory network, analogous to those found in sensory systems.

DISCUSSION

From the above experiments it seems plausible to believe that the system which retains auditory information, at least over time periods characteristic of short-term memory, is very similar in structure to the system which processes this information at the incoming level. These experiments show the presence of specific interactive effects in pitch memory which vary as a systematic function of the log frequency difference between the interacting tones. This leads to the hypothesis that pitch information is

retained on an array where elements are activated by tones of specific pitch; and that elements activated by tones separated by the same difference in log frequency units are spaced the same distance apart. Such arrays have been demonstrated by neurophysiological methods in the auditory projection areas of the central nervous system (Deutsch & Deutsch, 1973). Further, interactions appear to take place along this memory array which are remarkably similar to lateral inhibitory interactions in sensory systems, and which even occur over the same relative frequency range as centrally acting lateral inhibition in audition.

It is not here argued that there is an identity of structure between a sensory system and its memory counterpart. For instance, a memory system must retain the order in which different stimuli arrive, and this is not required of a sensory system. However, the present evidence leads to the hypothesis that the system that retains auditory information is a direct projection from the system that handles this information as it arrives. If this were so, then the concept of a fast-deteriorating preperceptual buffer which initially retains auditory information in uncategorized form (such as envisaged by Massaro, 1972) would appear redundant; we would then have to assume that information is first categorized in a particular manner by the sensory receiving areas of the nervous system, then translated into uncategorized form, and finally translated back into the form in which it had originally been categorized. How such recategorization could take place after it had been lost remains a mystery. On the other hand, the idea of a direct projection from a sensory array into a memory array, which preserves the initial form of stimulus categorization, is simple to envisage.

It further seems plausible to suppose that the form of organization shown here in the case of memory for a first-order acoustic attribute exists in memory for higher-order attributes also. As an example, we may consider the case of memory for tonal relationships. When two tones are presented simultaneously or in succession, a musical interval is perceived; and when the waveform frequencies of two tone pairs are related by the same ratio, the pairs form intervals of the same apparent size (Deutsch, 1969). We may therefore speculate that there exists a memory continuum whose individual elements are activated by the simultaneous or successive presentation of pairs of tones. Tone pairs standing in the same ratio project onto the same point along this continuum; adjacent ratios project onto adjacent points; and so on, to form a monotonic continuum of interval size. We may further speculate that systematic interactions occur along this continuum, analogous to those found in the case of memory for absolute pitch values.

We may thus envisage that systems retaining higher-order acoustic attributes are similar in principle to systems retaining basic attributes such

as pitch. Complex acoustic stimuli would then be represented on a number of memory arrays simultaneously. However, the persistence of memory along these various arrays would be expected to differ (as is clear when we compare memory for tonal relationships with memory for absolute pitch values). This would have the consequence that when memory for a complex acoustic stimulus is measured after varying time intervals have elapsed, there would be an apparent difference in its form of encoding. A similar argument has been made in a more general setting by Craik and Lockhart (1972).

CONCLUSION

In this article, a conceptual framework for auditory memory has been suggested which is very different from that generally assumed by memory theorists. It is proposed that auditory information is first subjected to a set of analyses, which provide input to various subdivisions of the auditory memory system. Information in these subdivisions is retained in parallel; however, the time constants of storage in the different subdivisions vary considerably. Finally, it is proposed that systematic interactions take place within these subdivisions, which are similar to effects in systems handling information at the perceptual level. At present this view is speculative, but it is hoped that it will provide a useful framework for future research.

RÉSUMÉ

Etude des modèles courants de la mémoire auditive. Après une analyse critique de l'hypothèse de l'affaiblissement sensoriel, il est proposé plutôt que la mémorisation auditive est la fonction d'un système dont les diverses subdivisions diffèrent considérablement dans leur persistance à retenir l'information. Il est proposé de plus que, au lieu d'être sensible à des influences générales telles qu'une limitation de capacité, la mémoire acoustique est sujette à des effets d'interaction hautement spécifiques, qui varient systématiquement selon les rapports existant entre les éléments impliqués.

REFERENCES

ATKINSON, R.C., & SHIFFRIN, R.M. Human memory: a proposed system and its control processes. Vol. II. In K.W. SPENCE & J.T. SPENCE (eds.), *The psychology of learning and motivation: Advances in research and theory.* New York: Academic Press, 1968.

ATTNEAVE, F., & OLSON, R.K. Pitch as a medium: a new approach to psychophysical scaling. *Am. J. Psychol.*, 1971, 84, 147–165.

BROADBENT, D.E. *Perception and communication* New York: Pergamon, 1958.

BROADBENT, D.E. Flow of information within the organism. *J. verb. Learn. verb. Behav.*, 1963, 2, 34–39.

BURNS, E.M., & WARD, W.D.   Categorical perception of musical intervals. Paper presented at the 86th meeting of the Acoustical Society of America, Los Angeles, October 1973.

CARTERETTE, E.C., FRIEDMAN, M.P., & LOVELL, J.D.   Mach bands in hearing. *J. acoust. Soc. Am.*, 1969, 45, 986–998.

COLE, R., COLTHEART, M., & ALLARD, F.   Memory of a speaker's voice; reaction time to same or different-voiced letters. *Quart. J. exp. Psychol.*, 1974, 26, 1–7.

CONRAD, R., & HULL, A.J.   Input modality and the serial position curve in short-term memory. *Psychonom. Sci.*, 1968, 10, 135–136.

COOLEY, R.K., & MCNULTY, J.A.   Recall of individual ccc trigrams over short intervals of time as a function of mode of presentation. *Psychonom. Sci.*, 1967, 9, 543–544.

CORBALLIS, M.C.   Rehearsal and decay in immediate recall of visually and aurally presented items. *Canad. J. Psychol.*, 1966, 20, 43–51.

CORTEEN, R.S., & WOOD, B.   Autonomic responses to shock associated words in an unattended channel. *J. exp. Psychol.*, 1972, 3, 308–313.

CRAIK, F.I.M.   Modality effects in short-term storage. *J. verb. Learn. verb. Behav.*, 1969, 8, 658–664.

CRAIK, F.I.M., & KIRSNER, K.   The effect of the speaker's voice on word recognition. *Quart. J. exp. Psychol.*, 1974, 26, 276–284.

CRAIK, F.I.M., & LOCKHART, R.S.   Levels of processing: a framework for memory research. *J. verb. Learn. verb. Behav.*, 1972, 11, 671–684.

CROWDER, R.G.   The role of one's own voice in immediate memory. *Cognitive Psychol.*, 1970, 1, 157–178.

CROWDER, R.G.   The sound of vowels and consonants in immediate memory. *J. verb. Learn. verb. Behav.*, 1971, 10, 587–596.

CROWDER, R.G.   Visual and auditory memory. In J.F. KAVANAGH & I.G. MATTINGLY (eds.), *Language by eye and ear: the relation between speech and learning to read.* Cambridge: M.I.T. Press, 1972.

CROWDER, R.G., & MORTON, J.   Precategorical acoustic storage (PAS). *Percept. Psychophys.*, 1969, 5, 365–373.

DEUTSCH, D.   Music recognition. *Psychol. Rev.*, 1969, 76, 300–307.

DEUTSCH, D.   Tones and numbers: specificity of interference in short-term memory. *Science*, 1970, 168, 1604–1605.

DEUTSCH, D.   Mapping of interactions in the pitch memory store. *Science*, 1972, 175, 1020–1022. (*a*)

DEUTSCH, D.   Effect of repetition of standard and comparison tones on recognition memory for pitch. *J. exp. Psychol.*, 1972, 93, 156–162. (*b*)

DEUTSCH, D.   Interference in memory between tones adjacent in the musical scale. *J. exp. Psychol.*, 1973, 100, 228–231.

DEUTSCH, D., & FEROE, F.   Disinhibition in pitch memory. *Percept. Psychophys.*, 1975, 17, 320–324.

DEUTSCH, J.A., & DEUTSCH, D.   Attention: some theoretical considerations. *Psychol. Rev.*, 1963, 70, 80–90.

DEUTSCH, J.A., & DEUTSCH, D.   *Physiological psychology.* (2nd Ed.). Homewood, Ill.: Dorsey Press, 1973.

ERICKSON, C.W., & JOHNSON, H.J.   Storage and decay characteristics of non-attended auditory stimuli. *J. exp. Psychol.*, 1964, 68, 28–36.

GUTTMAN, N., & JULESZ, B.   Lower limits of auditory periodicity analysis. *J. acoust. Soc. Am.*, 1963, 35, 610.

HARRIS, J.D.   The decline of pitch discrimination with time. *J. exp. Psychol.*, 1952, 43, 96–99.

HOUTGAST, T.   Psychophysical evidence for lateral inhibition in hearing. *J. acoust. Soc. Am.*, 1972, **51**, 1885–1894.

KLINKE, R., BOERGER, G., & GRUBER, J.   Alteration of afferent, tone-evoked activity of neurons of the cochlear nucleus following acoustic stimulation of the contra-lateral ear. *J. acoust. Soc. Am.*, 1969, **45**, 788–789.

KLINKE, R., BOERGER, G., & GRUBER, J.   The influence of the frequency relation in dichotic stimulation upon the cochlear nucleus activity. In R. PLOMP & G.F. SMOORENBURG (eds.), *Frequency analysis and periodicity detection in hearing.* Leiden: Sijthoff, 1970.

LESHOWITZ, B., & CUDAHY, E.   Frequency discrimination in the presence of another tone. *J. acoust. Soc. Am.*, 1973, **54**, 882–887.

LEWIS, J.L.   Semantic processing of unattended messages using dichotic listening. *J. exp. Psychol.*, 1970, **85**, 225–228.

LIBERMAN, A.M., COOPER, F.S., SHANKWEILER, D.P., & STUDDERT-KENNEDY, M.   Per-ception of the speech code. *Psychol. Rev.*, 1967, **74**, 431–461.

MASSARO, D.W.   Preperceptual auditory images *J. exp. Psychol.*, 1970, **85**, 411–417.

MASSARO, D.W.   Preperceptual images, processing time and perceptual units in audi-tory perception. *Psychol. Rev.*, 1972, **79**, 124–145.

MORTON, J.   A functional model for memory. In D.A. NORMAN (ed.), *Models for human memory.* New York: Academic Press, 1970.

MORTON, J., CROWDER, R.G., & PRUSSIN, H.A.   Experiments with the stimulus suffix effect. *J. exp. Psychol. Monogr.*, 1971, **91**, 169-190.

MURDOCK, B.B., JR.   Auditory and visual stores in short-term memory. *Acta Psycholo-gica*, 1967, **27**, 316–324.

MURDOCK, B.B., JR.   Four-channel effects in short-term memory. *Psychonom. Sci.*, 1971, **24**, 197–198.

MURDOCK, B.B., JR.   Short-term memory. Vol. 5. In G.H. BOWER & J.T. SPENCE (eds.), *The psychology of learning and motivation: advances in research and theory.* New York: Academic Press, 1972.

MURDOCK, B.B., JR., & WALKER, K.D.   Modality effects in free recall. *J. verb. Learn. verb. Behav.*, 1969, **8**, 665–676.

MURRAY, D.J.   Vocalization-at-presentation and immediate recall, with varying recall methods. *Quart. J. exp. Psychol.*, 1966, **18**, 9–18.

NEISSER, U.   *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

PAIVIO, A.   Mental imagery in associative learning and memory. *Psychol. Rev.*, 1969, **76**, 241–263.

POLLACK, I.   Memory for auditory waveform. *J. acoust. Soc. Am.*, 1972, **51**, 1209–1215.

POSNER, M.I.   Short term memory systems in human information processing. In A.F. Sanders (ed.), *Attention and performance.* Amsterdam: North Holland Publishing Company, 1967.

RATLIFF, F.   *Mach bands: quantitative studies of neural networks in the retina.* San Francisco: Holden-Day, 1965.

SHEPARD, R.N.   Recognition memory for words, sentences and pictures. *J. verb. Learn. verb. Behav.*, 1967, **6**, 156–163.

SIEGEL, W., SIEGEL, J.A., HARRIS, G., & SOPO, R.   Categorical perception of pitch by musicians with relative and absolute pitch. *University of Western Ontario Research Bulletin*, #305, September 1974.

SPERLING, G.   A model for visual memory tasks. *Human Factors*, 1963, **5**, 19–31.

SPERLING, G.   Successive approximations to a model for short-term memory. In A.F. SANDERS (ed.). *Attention and performance.* Amsterdam: North Holland Publish-ing Company, 1967.

STUDDERT-KENNEDY, M., LIBERMAN, A.M., HARRIS, K.S., & COOPER, F.S.   Motor theory of speech perception: a reply to Lane's critical review. *Psychol. Rev.*, 1970, 77, 234–249.

TREISMAN, A.M.   Monitoring and storage of irrelevant messages in selective attention. *J. verb. Learn. verb. Behav.*, 1964, 3, 449–459. (*a*)

TREISMAN, A.M.   Selective attention in man. *Brit. med. Bull.*, 1964, 20, 12–16. (*b*)

TREISMAN, M., & ROSTRON, A.B.   Brief auditory storage: a modification of Sperling's paradigm applied to audition. *Acta Psychologica*, 1972, 36, 161–170.

ZWISLOCKI, J.J.   A theory of central auditory masking and its partial validation. *J. acoust. Soc. Am.*, 1972, 52, 644–659.

ZWISLOCKI, J.J., BUINING, E., & GLANTZ, J.   Frequency distribution of central masking. *J. acoust. Soc. Am.*, 1968, 43, 1267–1271.

Auditory Memory. Canadian Journal of Psychology, 1975, 29, 87-105