

Disinhibition in pitch memory

DIANA DEUTSCH

Center for Human Information Processing, University of California, San Diego, California 92037

and

JOHN FEROE

Department of Mathematics, University of California, San Diego, California 92037

Recognition of the pitch of a tone is disrupted by the interpolation of other tones during the retention interval. The disruptive effect of an interpolated tone varies systematically as a function of its pitch relationship to the tone to be remembered, and is maximal at a $\frac{1}{2}$ -tone separation. When such a tone is interpolated, the interpolation in addition of a further tone that is $\frac{1}{2}$ tone removed from this disruptive tone (and $\frac{1}{2}$ tone removed from the tone to be remembered) causes recognition of the first tone substantially to return. When recognition performance is plotted as a function of the pitch relationship between these two interpolated tones, the results accord well with a model assuming mutual inhibitory interactions between pitch memory elements.

Models of human memory generally assume that information is systematically organized in storage only after verbal encoding. Factors influencing the retention of unlabeled information are presumed nonspecific in nature, such as decay (Broadbent, 1958), displacement on a simple first-in, first-out basis (Crowder & Morton, 1969), or channel capacity limitation (Posner, 1967). However, information concerning tonal pitch has been shown to be subject to precise and systematic influences in storage. In one study (Deutsch, 1972), two temporally separated tones were compared for pitch and a series of tones was interpolated during the retention interval. The pitch relationship between one of these tones and the tone to be remembered was varied on a logarithmic continuum, and the error rate was found to vary systematically as a function of this relationship. The pitch memory system was thus shown to be organized in a specific and orderly fashion; however, the basis for this organization was left undetermined. In the present article, we provide evidence for the existence of mutual inhibitory interactions between elements of the system.

In the experiment of Deutsch (1972), subjects were required to compare the pitch of two tones which were separated by a retention interval of 5 sec duration, in which six extra tones were interpolated. The two test tones were either identical in pitch or they differed by a semitone. There was always placed in the second serial position of the intervening sequence a tone whose pitch bore a critical relationship to the pitch of the first test tone. This relationship varied $\frac{1}{6}$ 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 $1\frac{1}{2}$ tones removed from the first test tone. It was found that when the first test tone and the critical intervening tone were identical in pitch, memory facilitation was produced. Errors rose with increasing pitch separation between these two tones, peaked at a separation of $\frac{1}{2}$ tone, and returned roughly to baseline at a whole-tone separation.

This function may be explained by assuming that the pitch memory system is based on an array in which individual elements are activated by tones of specific pitch. These elements are tonotopically organized on a log frequency continuum, so that elements activated by adjacent pitches lie adjacent to each other. Inhibitory interactions take place along this array which are a function of distance between the elements involved, in a fashion analogous to lateral inhibitory networks in sensory systems. Recognition judgments are made between elements situated along this array; and thus the elements activated by the test tones are subject to the inhibitory influences of elements activated by the interpolated tones. It is assumed that when these memory elements are inhibited, they emit weaker signals, and so errors in recognition judgments increase.

Assume that a set of n tones activates a corresponding set of n elements on this array. Then let e be the strength of signal emitted by an element when its corresponding tone is at least a whole tone removed from the other tones in the set. Let k_s be the coefficient of inhibition

between elements underlying two tones separated by a distance of s tones; and let r_p be the strength of signal emitted by the element underlying the p^{th} tone. Then, for $n = 2$, the following pair of simultaneous linear equations describes the mutual inhibition between r_1 and r_2 ,

$$r_1 = e - k_s r_2$$

$$r_2 = e - k_s r_1$$

Solving for k_s in terms of r_1 and e , we obtain

$$k_s = \frac{e - r_1}{r_1}$$

Using the coefficients k_s , the interaction between the n elements is then described by the n simultaneous linear equations

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

where $s(p, q)$ is the distance between the p^{th} and q^{th} tones. The solution of these equations then predicts the net influence on the signal strength of the element underlying a test tone exerted by the elements underlying the other tones in the set.

The mathematical model proposed here is based upon models describing recurrent lateral inhibitory interactions in systems processing incoming sensory information (Ratliff, 1965). However, we are not suggesting that a sensory memory system is identical to the system which processes information at the perceptual level. The two systems must by their very functions differ in important respects. For instance, a perceptual system cannot continue to respond to a stimulus long after its termination, or we should soon cease to have discriminable images. In contrast, a memory system exists precisely to retain information in the absence of the original stimulus. The present model simply proposes the existence of a network of memory elements whose interactions can be described by a set of equations analogous to those describing certain interactions at the perceptual level.

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 test tones are identical in pitch, if there were interpolated two critical tones, one $\frac{1}{2}$ tone removed along the pitch continuum, the error rate should be a function of the pitch relationship between the two critical interpolated 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 $\frac{1}{2}$ tone separation, and then return to baseline. For sequences where the test tones differ in pitch, the prediction is complicated by the fact that these tones now activate two memory elements, each differentially affected by the elements underlying the interpolated tones. In such sequences, the second test tone may be placed either on the same side of the first test tone along the pitch continuum as the critical intervening tones, or on the opposite side. The simplest case theoretically is where it is placed on the opposite side; and this case is considered here. The effect on the element underlying the second test tone should here be in the opposite direction to the effect on the element underlying the first test tone; and it should be smaller in size, since it is of the third order. No significant overall effect is therefore predicted for such sequences.

The first experiment was designed to determine whether a significant disinhibition effect is indeed produced by interpolating a tone which is $\frac{1}{2}$ tone removed from the disruptive tone.

EXPERIMENT I

Method

Procedure. In all experimental conditions, a test tone was presented, which was followed by six intervening tones, and then, after a pause, by a second test tone. Subjects were instructed to remember the first test tone, to ignore the six intervening tones, and then to indicate whether the pitch of the second test tone was the same or different from the first by writing "S" (same) or "D" (different). The tones were all 200 msec in duration, and separated by 300-msec pauses, except for a 2-sec pause before the second test tone. All tones were sinusoids.

Conditions. There were eight conditions in the experiment. In all conditions there was placed in the second serial position of the intervening sequence a tone whose pitch was $\frac{1}{2}$ tone removed from the pitch of the first test tone. Further, in all conditions but the last, there was placed in the fourth serial position a tone, further removed along the pitch continuum, whose pitch bore a critical relationship to the pitch of the tone in the second serial position. A unique value of pitch separation between the two critical interpolated tones was incorporated in each of the seven conditions. These values were spaced in 1/6-tone steps from identity to a whole-tone separation on the equal-tempered scale. In the eighth condition, the tone in the fourth serial position was chosen instead, as were the other tones in the intervening sequence. This last condition thus provided a baseline error rate for sequences in which a tone that was $\frac{1}{2}$ tone removed from the first test tone was interpolated. In all conditions, the test tones were identical in pitch in half of the sequences, and they differed by a semitone in the other half.

The tape consisted of 96 sequences which were presented in eight groups of 12. Sequences within each group were separated by 10-sec pauses, and a 5-min break was incorporated between each group. The sequences were presented in random order, with no separation by condition. The subjects listened to the entire tape on two separate occasions, and their results were averaged.

Test tones. The test tone pitches were drawn from a set of 12, taken from an equal-tempered scale (International Pitch; A = 435) and ranging from Middle C to the B above. The frequencies employed (in hertz) were: C = 259, C# = 274, D = 290, D# = 308, E = 326, F = 345, F# = 366, G = 388, G# = 411, A = 435, A# = 461, B = 488. Within each condition, in sequences where the test tones differed, the first test tone was

a semitone higher than the second in half of the instances and a semitone lower in the other half. All combinations of test tone pitches were employed equally often in all conditions.

Interpolated tones. The pitch relationships between the test tones and the critical interpolated tones were as described in the conditions section. Within each condition, the critical interpolated tones were higher than the first test tone in half of the sequences and lower in the other half. This was true both when the test tones were identical in pitch and also when they differed. In sequences where the test tones differed, the critical interpolated tones were always placed opposite to the second test tone along the pitch continuum. The pitches of the other interpolated tones were taken from the same scale as the test tones ($A = 435$), and the 24 semitones comprising the two-octave range from the F# below Middle C to the F an octave and a half above were all employed. The frequencies used (in hertz), therefore, were: F# = 183, G = 194, G# = 205, A = 218, A# = 230, B = 244, C = 259, C# = 274, D = 290, D# = 308, E = 326, F = 345, F# = 366, G = 388, G# = 411, A = 435, A# = 461, B = 488, C = 517, C# = 548, D = 581, D# = 615, E = 652, and F = 691. The tones were chosen randomly from this range with the following restrictions. No sequence contained repeated tones or tones separated exactly by an octave. Further, no tone was included in any sequence which was separated from the first test tone by less than five semitones on the side where the critical interpolated tones were placed or by less than three semitones on the opposite side. Tones separated by exactly an octave from tones within this critical range were also excluded. All tones were of equal loudness.

Subjects. Twenty-one undergraduates at the University of California at San Diego served as subjects for the experiment, and were paid for their services. The subjects were selected for obtaining a score of at least 95% correct on a tape containing sequences in which all the interpolated tones were chosen in the same way as the randomly interpolated tones of the experiment.

Apparatus. The tones were generated by a Wavetek oscillator controlled by a PDP-8 computer, and the output was recorded on high-fidelity tape. The tape was played to the subjects on a high-quality tape recorder through loudspeakers.

Results

The filled circles on Figure 1 plot error rates for sequences where the test tones were identical in pitch. It can be seen that the predicted disinhibition function was indeed obtained. When the second critical intervening tone was identical in pitch to the first, the error rate was significantly higher than in the baseline condition when only the first critical tone was interpolated ($p < .005$, one-tailed, on a Wilcoxon test). Errors decreased systematically as the second critical tone moved further along the pitch continuum, and dipped maximally when this tone was $\frac{2}{3}$ tone removed from the first critical tone. The error rate when the two critical tones were separated from each other by $\frac{2}{3}$ tone was significantly lower than in the baseline condition when only the first critical tone was interpolated ($p < .01$, one-tailed, on a Wilcoxon test). For sequences where the test tones differed in pitch, no significant effects were obtained, as expected on the present model.

Discussion

The finding of a significant cumulation of errors produced by repetition of the disruptive tone, together with a significant decrement in errors by interpolating a tone which is $\frac{2}{3}$ tone removed from the disruptive tone, pro-

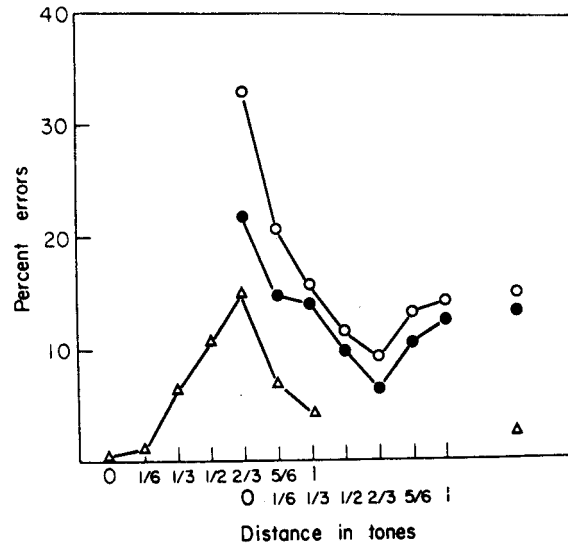


Figure 1. Percent errors in pitch recognition obtained experimentally and assumed theoretically. Filled circles display percent errors in Experiment I. Here a tone that was $\frac{2}{3}$ tone removed from the test tone was always interpolated in the intervening sequence. Errors are plotted as a function of the pitch relationship between this tone and a second critical interpolated tone which was further removed along the pitch continuum. (Filled circle at right displays percent errors where no further critical tone was interpolated.) Open triangles plot percent errors in Experiment II. Here the pitch relationship between the test tone and a critical interpolated tone was varied, and errors are plotted as a function of this relationship. (Open triangle at right displays percent errors where no tone was interpolated in the critical range under study.) Open circles plot percent errors theoretically predicted for Experiment I from the baseline function obtained empirically in Experiment II. (Open circle at right shows percent errors assumed where no further critical tone is interpolated.)

vides support for the present model. However, in the absence of a baseline inhibitory function, the parameters of this second-order effect cannot be compared quantitatively with the function predicted by the model. So, in Experiment II, a baseline curve for the original inhibitory effect was obtained, using subjects selected on the same criterion as for Experiment I. This baseline function was used to compute values of e and k_s , assuming that error rate was a function of strength of signal emitted by the element underlying the test tone. The predicted disinhibition function was then computed from these parameters, and compared with the function obtained experimentally.

EXPERIMENT II

Method

Procedure. The procedure was identical to the employed in Experiment I.

Conditions. There were eight conditions in the experiment. In every condition but the last, a tone whose pitch bore a critical relationship to

the pitch of the first test tone was placed in the second serial position of the intervening sequence. The relationship between these two tones varied from identity to a whole-tone separation on the equal-tempered scale. A unique value of pitch separation was incorporated in each of the seven conditions, these values being placed at equal intervals of $1/6$ tone within this whole tone range. In the eighth condition, the pitch of the tone in the second serial position was chosen instead, as were the other tones in the intervening sequence. This last condition therefore provided a baseline error rate for sequences in which no critical tone was interpolated.

In all other respects, the conditions were the same as in Experiment I.

Tonal stimuli. Except as specified by the experimental conditions, the tonal stimuli were chosen in the same way as in Experiment I.

Subjects. Twenty-three undergraduates at the University of California at San Diego served as subjects for this experiment, and were paid for their services. The subjects were selected on the same criterion as for Experiment I.

Apparatus. This was the same as in Experiment I.

Results

The open triangles in Figure 1 plot error rates for sequences when the test tones were identical in pitch. It can be seen that errors rose systematically with increasing pitch separation between the test tone and the critical intervening tone, peaked at a separation of $\frac{2}{3}$ tone, and returned to baseline at a whole-tone separation. The increase in errors at a separation of $\frac{2}{3}$ tone was statistically significant ($p < .005$, one-tailed, on a Wilcoxon test). This function was essentially the same as that found in Deutsch (1972), except that the overall error rate was lower, due to the more rigorous selection criterion.

The predicted disinhibition function computed from this baseline curve is plotted by the open circles in Figure 1. It can be seen that there is a close correspondence between the disinhibition functions predicted theoretically and obtained experimentally. The only quantitative discrepancy appears to be in the amount of error cumulation when a tone that is $\frac{2}{3}$ tone removed from the test tone is included twice in the intervening sequence.

Discussion

The striking correspondence between the theoretically predicted disinhibition function and the function obtained experimentally lends further support to the proposed model. Moreover, it is difficult to account for the present findings on other grounds. Certainly, these results would not be expected from memory models based on general factors such as decay or general capacity limitation. The classical concept of similarity-based interference would also not expect a cancellation of the disruptive effect.

As described above, the form of organization here proposed for pitch memory elements is essentially similar to the organization of recurrent inhibitory networks in systems processing incoming sensory information (Ratliff, 1965). Such systems have been most intensively investigated in the case of vision. However, both psychophys-

ical and physiological evidence for lateral inhibition in the auditory system have been provided.

At the psychophysical level, Carterette, Friedman, and Lovell (1969, 1970) measured the masker thresholds for pure tones in the presence of bands of masking noise with very sharp edges. They were able to demonstrate sharpening in the neighborhood of the cut-off frequencies of these bands. Houtgast (1972) presented short test-tone bursts in the gaps between repeated masker bursts that were composed of noise with a steep positive or negative gradient at a particular frequency. He also obtained edge effects with this method. Further, Zwislocki, Buing, and Gantz (1968) and Zwislocki (1970) studied the shift of thresholds for tones presented to one ear due to the simultaneous presentation of tones to the other ear. Their data show, at low masking intensities, clear relative minima on either side of the principal masking frequency.

At the physiological level, evidence for both peripherally and centrally acting lateral inhibition has been obtained. Sachs and Kiang (1968), recording from single units in the auditory nerve of the cat, demonstrated the existence of inhibitory areas arising systematically on both sides of the characteristic frequency of the unit. The interactions producing these effects must have occurred at the periphery, since similar findings were also obtained in animals in which the olivocochlear efferents had been transected (Kiang, 1968). Indeed, Johnstone and Taylor (1970, Legoux, Remond, and Greenbaum (1973), and Rhode (1971) present evidence that such two-tone inhibition results from mechanical events in the cochlear partition. On the other hand, Klinke, Boerger, and Gruber (1969, 1970), studying single-unit responses in the cochlear nucleus of the cat, found that contralateral stimulation reduced the response of a unit to ipsilateral stimulation. The strongest inhibition arose from frequencies adjacent to the characteristic frequency of the unit, and the overall curve was one of typical lateral inhibition. In this study, the sound pressure levels were kept low enough to avoid interaural cross-talk. This experiment therefore strongly suggests the presence of centrally acting lateral inhibition in the auditory system. Klinke et al. (1969) conclude that it is here mediated by olivocochlear efferents. Centrally acting lateral inhibition has also been found in the visual system. For instance, the excitatory centers of receptive fields in the lateral geniculate have been shown in cats to derive from different retinal areas than their inhibitory surrounds (Mafei & Fiorentini, 1972).

The present findings suggest that networks similar to those underlying the processing of incoming sensory information also exist at the level of its short-term storage. Such networks would help to preserve the fineness of the memory image in the same way as they function to sharpen the sensory image at the perceptual level.

REFERENCES

- BROADBENT, D. E. *Perception and communication*. New York: Pergamon, 1958.
- CARTERETTE, E. C., FRIEDMAN, M. P., & LOVELL, J. D. Mach bands in hearing. *Journal of the Acoustical Society of America*, 1969, **45**, 986-998.
- CARTERETTE, E. C., FRIEDMAN, M. P., & LOVELL, J. D. Mach bands in auditory perception. In R. Plomp and G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.
- CROWDER, R. G., & MORTON, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, **5**, 365-373.
- DEUTSCH, D. Mapping of interactions in the pitch memory store. *Science*, 1972, **175**, 1020-1022.
- HOUTGAST, T. Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America*, 1972, **51**, 1885-1894.
- JOHNSTONE, B. M., & TAYLOR, K. Mechanical aspects of cochlear function. In R. Plomp and G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.
- KIANG, N. Y. S. A survey of recent developments in the study of auditory physiology. *Annals of Otolaryngology, Rhinology and Laryngology*, 1968, **77**, 656-676.
- KLINKE, R., BOEGER, G., & GRUBER, J. Alteration of afferent, tone-evoked activity of neurons of the cochlear nucleus following acoustic stimulation of the contralateral ear. *Journal of the Acoustical Society of America*, 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 and G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.
- LEGOUIX, J., REMOND, M., & GREENBAUM, H. B. Interference and two-tone inhibition. *Journal of the Acoustical Society of America*, 1973, **53**, 409-418.
- MAFFEI, I., & FIORENTINI, A. Retinogeniculate convergence and analysis of contrast. *Journal of Neurophysiology*, 1972, **35**, 65-72.
- POSNER, M. I. Short term memory systems in human information processing. In A. F. Sanders (Eds.), *Attention and performance*, Amsterdam: North Holland, 1967.
- RATLIFF, F. *Mach bands: Quantitative studies of neural networks in the retina*, San Francisco: Holden-Day, 1965.
- RHODE, W. S. Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique. *Journal of the Acoustical Society of America*, 1971, **49**, 1218-1223.
- SACHS, M. B., & KANG, N. Y. S. Two-tone inhibition in auditory nerve fibers. *Journal of the Acoustical Society of America*, 1968, **43**, 1120-1128.
- ZWISLOCKI, J. J., BUINING, E., & GLANTZ, J. Frequency distribution of central masking. *Journal of the Acoustical Society of America*, 1968, **43**, 1267-1271.
- ZWISLOCKI, J. J. Central masking and auditory frequency selectivity. In R. Plomp and G. F. Smoorenburg (Eds.), *Frequency analysis and periodicity detection in hearing*. Leiden: Sijthoff, 1970.

(Received for publication August 2, 1974:
revision accepted December 10, 1974.)