Quarterly Journal of Experimental Psychology (1976) 28, 255–270

RECOGNITION OF TRANSPOSED MELODIC SEQUENCES

LOLA L. CUDDY AND ANNABEL J. COHEN

Queen's University, Kingston, Ontario, Canada

Accuracy of recognition for short (three-note) transposed melodic sequences was measured and compared with accuracy predicted by three models of recognition each of which described a different degree of abstraction and synthesis of the musical intervals contained in the sequence. For subjects with musical training, recognition was best described by a model that assumed abstraction and synthesis of the musical intervals between both adjacent and non-adjacent tones of the sequence. For subjects without musical training, recognition was much less accurate but there was some evidence that intervals between adjacent tones were abstracted. Of major theoretical interest, however, was the finding that none of the models provided a comprehensive account of the data. Not merely the size of the intervals contained in a sequence determines accuracy of recognition of the sequence, but also the order or configuration of the intervals. It is suggested that particular interval configurations facilitate the abstraction of tonal structure.

Introduction

An interesting problem of musical perception is that of transposition—that is, the invariance of melodic form when the melody itself is presented in different keys. When a melody is transposed, the absolute frequencies of the tones of the melody are changed; thus an account of transposition must consider how relationships among frequencies are perceived. More specifically, the question is, how may a perceptual invariant be abstracted from different temporal patterns of tones in which the frequency ratio between successive tones is the only frequency information preserved from pattern to pattern?

Deutsch (1969) has suggested that a melodic percept results from the abstraction and synthesis of musical intervals between successive tones, and has proposed a hierarchical neuronal-network model for the abstraction of frequency ratio. The model is an auditory analogue of the model of visual processing developed by Hubel and Wiesel (1962) to account for visual stimulus equivalence. In a manner similar to their representation of a visual pattern as a composition of lines in particular orientations, Deutsch regards a melody as a sequence of successive musical intervals.

A melody or tune is shown in musical notation in Figure 1, with the musical interval between adjacent tones given in semits below the notation.* (A semit,

*Readers who do not feel comfortable with musical notation may follow the descriptions of the auditory patterns simply by remembering the definitions of semit and interval given in the Introduction.

the unit of measurement in Western equal temperament, is a ratio of two fundamental frequencies equal to 1·059.) In Figure 1, the + sign represents an interval ascending in pitch and the − sign represents an interval descending in pitch. The melody in Figure 1(a) is the key of C; it is transposed in Figure 1(b) to the key of B flat, and, although the absolute frequencies are changed, the semit distance between adjacent tones is maintained.

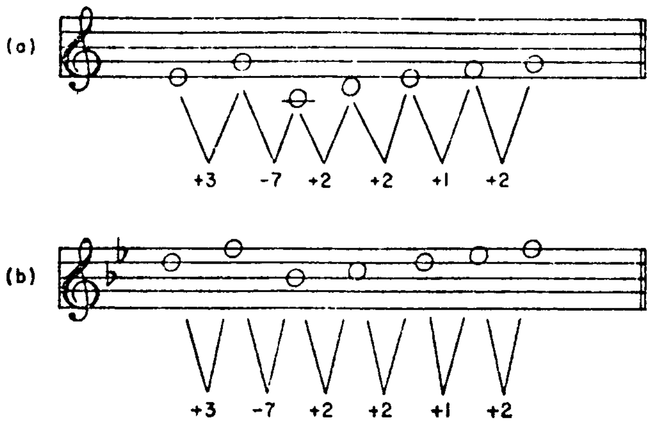

FIGURE 1. (a) Musical notation of a short melody. Size (in semits) and direction (ascending or descending) of intervals between adjacent tones are shown beneath musical notation.
(b) The melody transposed.

If intervals between non-adjacent tones are also considered, the description of a melody becomes more complicated. A simple three-note sequence generates three possible interval abstractions—the intervals between the first and second, the second and third, and the first and third tones of the sequence. Figure 2 represents these three intervals, again denoted in musical notation and in semits, for a standard sequence [Fig. 2(a)], a correct transposition of the sequence [Fig. 2(b)], and an incorrect transposition [Fig. 2(c)]. The incorrect transposition was obtained by lowering the relative position of the third tone by one semit; this lowering results in a change of two of the component intervals. The correct detection of either change, or both, would allow a listener to judge the melody of Figure 2(c) an incorrect transposition of the melody of Figure 2(a). Furthermore, if the accuracy with which the listener was able to detect a change in the intervals was known, it should be possible to predict with what accuracy the incorrect transposition would be identified.

In the following experiment, accuracy of recognition for three-note sequences was measured and compared with accuracy predicted from a separate test of interval recognition. Three models were postulated to predict the three-note recognition accuracy. Each model assumes that accuracy of melody recognition is dependent upon accuracy of abstraction of the component intervals, that such abstraction may be quantified by measuring accuracy of recognition of the component intervals, and that abstracted intervals are independent sources of

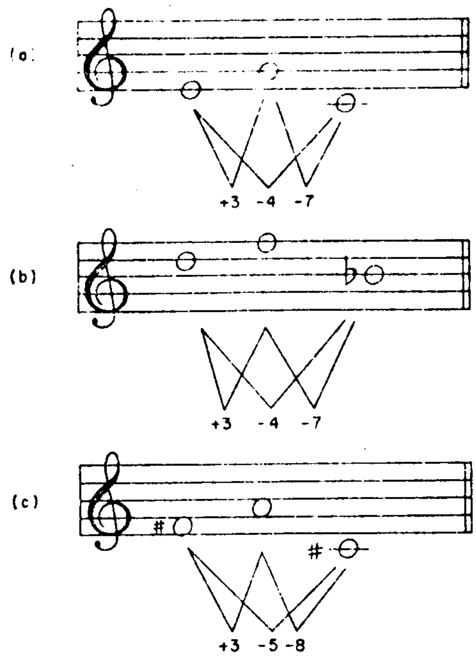

FIGURE 2. (a) Three-note sequence. (b) Correct transposition.
(c) Incorrect transposition. Size (in semits) and direction (ascending or descending) of the two intervals between adjacent tones and size and direction of the interval between non-adjacent tones are shown beneath the musical notation for each sequence.

information. The models differ in the degree to which the listener is assumed to be capable of interval abstraction and synthesis. The assumption of model MAX (Maximum Interval Information) is that all available interval information is abstracted and independently combined—that is, intervals between both adjacent and non-adjacent tones are abstracted and synthesized. The assumption of model ADJ (Adjacent) is that only intervals between adjacent tones are abstracted and synthesized. The assumption of the third model, model R.S. (Random Sampling) is that intervals between adjacent tones and non-adjacent tones may be abstracted but that for any given sequence judgment is dependent upon the ability to detect one of the changed intervals, the changed interval being randomly selected. Model R.S. implies that the listener searches for one changed interval either between adjacent or non-adjacent tones but does not or cannot appreciate the fact that two intervals are changed and that this information can be combined in a probabilistic sense to improve recognition performance.

The models refer specifically to musical perception and, for this reason, musical sequences and intervals were employed. In addition, since accuracy and structure of musical judgment have been shown to depend upon practice or to be related to musical training (Cuddy, 1968, 1970, 1971), the applicability of each model was examined at three levels of musical experience.

Method

Listeners

Forty-two volunteers from the student population at Queen's University, ranging in age from 18 to 40 years, constituted three groups differing in level of musical experience. Fourteen subjects reporting neither musical training nor participation in musical activities were placed in the "untrained" (U) group. A musical history was taken for each of 28 volunteers who reported some musical background. Subjects listed music degrees, instruments played, years of training, and extra-curricular musical activities. Five judges with advanced training in both music and psychology ranked subjects in order of musical experience. The 14 subjects receiving the highest rank, averaged across judges, were placed in the "highly trained" (H) group and the remaining 14 subjects were placed in the "trained" (T) group. Median number of years of musical training was 10 for group H and 5 for group T. Typical members of group H played two or more musical instruments; most members of group T played only one.

Three-note test

A test trial for the three-note test consisted of a standard sequence followed by two comparison sequences, one a correct and the other an incorrect transposition of the standard sequence. The paradigm was thus two-alternative forced choice, with assignment of the correct transposition to the first or second comparison sequence randomly determined and the subject was required to indicate which comparison sequence was the correct transposition.

Each standard sequence was one of the six possible sequential patterns of the major triad (doh-me-sol). Figure 3 shows these six patterns with pattern label given above the line of musical notation and semit values between intervals given below the line of musical notation. An incorrect transposition for each pattern was formed by raising or lowering by one semit the relative position of the first, second, or third tone of the pattern. Note that an alteration of the first or third tone alters the intervals between adjacent and non-adjacent tones;

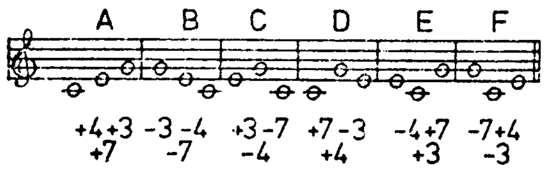

FIGURE 3. The six sequential patterns (A–F) of the major triad.

alteration of the second tone alters intervals between adjacent tones only. There were thus six types of incorrect transposition for each pattern or a total of 36 three-note test conditions.

Three 36-trial blocks were independently composed with each condition represented once per block. The trials were randomly ordered with the restriction that the same pattern or incorrect transposition did not follow on consecutive trials. Each of the three sequences per trial—the standard, the correct and the incorrect comparison—began on a tone randomly chosen from the range of 25 musical notes from D_3 to D_5 (156 Hz to 622 Hz), approximately one octave above and below middle C. Octaves between the first tone of the standard sequence and comparison sequences were avoided.

Tones were produced at equal loudnesses by an experienced pianist on a Heintzman grand piano recently tuned to $A_4 = 440$ Hz in the recording studio of Queen's Radio Station CFRC, and were recorded on a Magnecord tape recorder using Scotch 202 magnetic tape.* Timing was determined by a metronome set at 90 beats per min so that the duration between onset of successive tones was 0·67 s. Total duration of a three-note sequence was 2 s, duration between standard and first comparison sequence, 2 s, between the two comparison sequences, 1·33 s, and between trials, 2·67 s. Total test time for three blocks of 36 trials per block was approximately 24 min.

There were five practice trials, identical in timing to the trials of the main test. Patterns other than the major triad were used, and in the first two practice trials the incorrect transposition contained a shift of more than 1 semit; in the remaining trials the shift was 1 semit.

Interval test

The two-alternative forced choice paradigm was also used for the interval test. Each trial of the interval test consisted of a standard sequence followed by two comparison sequences with each sequence now composed of only two tones presented successively rather than three. The two tones of all sequences formed a musical interval but one of the intervals presented as a comparison sequence was an incorrect transposition of the standard sequence.

The interval test trials were generated directly from the three-note test trials. Because each three-note test trial involved the alteration, in the incorrect comparison sequence, of two of the component intervals of the standard sequence (cf. Fig. 2) each three-note test trial generated two interval trials—two different standard intervals each with its own incorrect comparison interval. Each block of 36 three-note test trials thus produced a block of 72 interval test trials.

The 72 interval test trials so formed represented 12 different test conditions each occurring six times randomly throughout the block. The 12 test conditions represented the combination of six different standard intervals generated by the patterns of the three-note test (3, 4 or 7 semits in ascending or descending direction) with two sizes of incorrect transposition generated by the incorrect transposition of the patterns (one semit larger or one semit smaller than the standard interval).

Recording procedure was identical to that for the three-note test. The duration between the onset of the two successive tones of each interval was 0·67 s, so that the duration of each sequence, or interval, was 1·33 s. The duration between the standard interval and the first comparison interval was 1·33 s, between the two comparison intervals, 0·67 s, and between trials, 2 s. The interval test maintained the same relative timing between duration of sequences and duration of silent intervals as the three-note test. Total test time for three blocks of 72 trials was approximately 35 min.

Five practice trials were recorded, with the incorrect comparison differing from the standard by more than 1 semit for the first two trials, and 1 semit for the remaining trials.

Procedure

Subjects were tested individually or in pairs in sound-isolated chambers. Tones were reproduced by a Crown tape recorder Model CI 822 and were presented monaurally through MBK 600 earphones at a loudness level judged comfortable by listeners. Each subject received both tests with the order of presentation of the tests counterbalanced across subjects. The three blocks of each test were presented in random order. The instructions preceding each test explained the construction of the trials and the nature of the two-

* Piano tones were used at this stage of the research in an attempt to insure that a processing system for musical perception would be tested. Research in progress suggests that the use of a musically-familiar timbre greatly reduces variability attributable to practice effects, or familiarization with the task.

Earlier, a mechanical device for striking piano keys with equal force was constructed and evaluated. It was found far inferior to the experienced pianist for equalizing perceived loudness of tones.

Recorded sequences were examined with a Tektronix oscilloscope Type 422. The estimate of the variance of the duration between two successive tone onsets was less than 20 ms.

alternative forced-choice paradigm; the practice trials were then presented and subjects given further opportunity to ask questions about the procedure. None reported difficulty with the concept of transposition defined as "the same tune, melody or note-pattern in another key".

Scoring

For each subject, proportion of correct responses was calculated for each of the 36 three-note test conditions and each of the 12 interval test conditions. These proportions were treated as the observed probability of a correct response. Next, predicted probabilities of a correct response were obtained for each of the three-note test conditions. For each subject for each condition, the proportion correct for the two interval conditions generated by that three-note test condition were inserted into three equations corresponding to the three models of interval abstraction and synthesis as follows:

Model MAX

$$P(3N_i) = P(I_{i1}) + P(I_{i2}) - P(I_{i1})P(I_{i2}).$$

Model ADJ

$P(3N_i) = P(I_{i1})$ or coincident with model MAX (see below).

Model R.S.

$$P(3N_i) = \frac{P(I_{i1}) + P(I_{i2}).}{2}$$

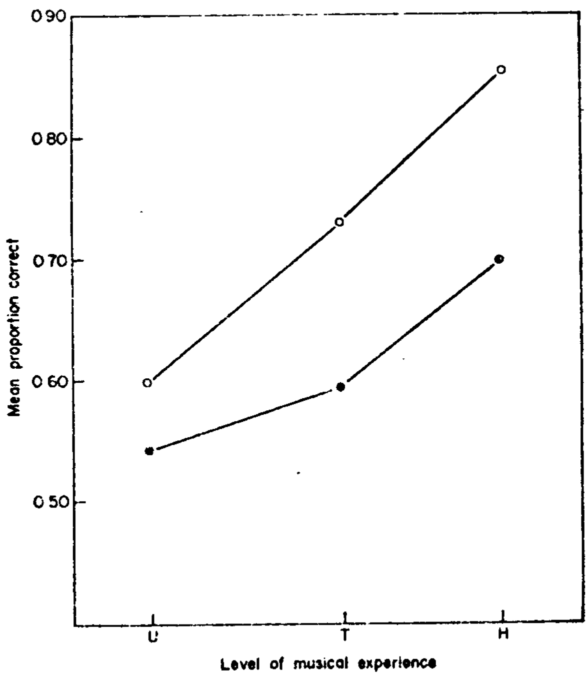

FIGURE 4. Mean proportion correct for interval and three-note tests for three levels of musical experience. Three-note, ○——○; interval, ●——●.

where $P(3N_i)$ is the predicted probability of a correct response for a three-note test condition, $i = 1$ to 36, and $P(I_{ij})$ is the obtained proportion of a correct response to an interval test condition derived from the ith three-note test condition, $j = 1,2$. The value of j is defined in terms of the three-note test condition. For three-note test conditions that involve alteration of the two intervals between adjacent tones, $j = 1$ is the interval between the first and second tone, and $j = 2$ is the interval between the second and third tone. In these conditions the predictions of model MAX and model ADJ by definition coincide; they both predict that probability of correct recognition for a three-note test condition is the combined probability of correct recognition for either of the two changed intervals. For three-note test conditions that involved alteration of an interval between adjacent tones and an interval between non-adjacent tones, $j = 1$ is the interval between adjacent tones and $j = 2$ is the interval between non-adjacent tones. Here model MAX still predicts that probability of correct recognition for the three-note condition is the combined probability of the correct recognition of either of the two changed intervals, but model ADJ predicts that correct recognition is completely determined by the accuracy of recognition for the changed interval occurring between adjacent tones. In either case, model R.S. simply predicts that probability of correct three-note recognition is the average probability of correct recognition for the altered intervals.

Results

The first analysis compared over-all performance on the interval test with performance on the three-note test for each level of musical experience. Proportion of correct responses is shown in Figure 4. It was found that the interval

TABLE I
Mean proportion correct for three-note tests

	Incorrect transposition						
	LL	RL	LM	RM	LH	RH	Mean
Group U							
Pattern							
A	0·64	0·71	0·57	0·69	0·69	0·52	0·64
B	0·55	0·62	0·55	0·81	0·74	0·57	0·64
C	0·55	0·74	0·62	0·62	0·52	0·52	0·60
D	0·50	0·52	0·38	0·76	0·75	0·48	0·56
E	0·36	0·67	0·64	0·52	0·62	0·64	0·58
F	0·64	0·67	0·59	0·71	0·50	0·55	0·61
Group T							
Pattern							
A	0·71	0·88	0·71	0·90	0·86	0·86	0·81
B	0·76	0·71	0·64	0·93	0·81	0·69	0·76
C	0·67	0·74	0·57	0·76	0·76	0·81	0·72
D	0·69	0·78	0·52	0·88	0·75	0·62	0·71
E	0·64	0·79	0·64	0·74	0·67	0·81	0·71
F	0·57	0·67	0·57	0·76	0·52	0·76	0·64
Group H							
Pattern							
A	0·93	0·86	0·95	0·98	0·98	0·90	0·93
B	0·88	0·90	0·83	0·95	1·00	0·81	0·90
C	0·86	0·83	0·86	0·95	0·90	0·98	0·90
D	0·74	0·86	0·76	0·93	0·93	0·79	0·83
E	0·88	0·74	0·79	0·83	0·81	0·90	0·83
F	0·90	0·76	0·76	0·79	0·88	0·88	0·83

test was more difficult than the three-note test, that performance improved with increasing level of musical experience, and that the difference between tests increased with increasing level of musical experience. All effects were statistically significant beyond the 0·01 level (for tests, $F = 91{\cdot}34$, $df = 1, 39$; for level of experience, $F = 18{\cdot}14$, $df = 2, 39$; and for test × level of experience interaction, $F = 6{\cdot}80$, $df = 2, 39$).

The data for the three-note test and the interval test were then examined in detail and were separately submitted to analysis of variance. Mean proportion correct judgments for each level of experience for each condition within the three-note test is given in Table I. The rows of Table I refer to pattern label (cf. Fig. 3) and the columns refer to type of incorrect transposition. The first symbol of a column heading, R or L, indicates whether the tone altered in transposition was raised or lowered one semit, and the second symbol, H, M or L indicates whether the tone altered was the highest, middle or lowest pitch of the pattern. The column LM, for example, refers to instances where the middle pitch of the major triad pattern is lowered, thus creating an incorrect transposition based on

TABLE II

Mean proportion correct for interval test

	Incorrect transposition		
	Smaller (− 1 semit)	Larger (+ 1 semit)	Mean
Group U			
Interval			
+3	0·60	0·55	0·58
−3	0·69	0·48	0·59
+4	0·51	0·51	0·51
−4	0·54	0·55	0·54
+7	0·54	0·55	0·54
−7	0·48	0·57	0·52
Group T			
Interval			
+3	0·69	0·53	0·61
−3	0·78	0·46	0·62
+4	0·55	0·59	0·57
−4	0·53	0·64	0·58
+7	0·63	0·57	0·60
−7	0·59	0·60	0·60
Group H			
Interval			
+3	0·82	0·60	0·71
−3	0·92	0·60	0·76
+4	0·50	0·58	0·54
−4	0·52	0·76	0·64
+7	0·72	0·81	0·77
−7	0·78	0·77	0·77

TABLE III

*Predictions from three models for group U**

		Incorrect transposition					
Pattern	Model	LL	RL	LM	RM	LH	RH
A	MAX	0·78	0·77			0·82	0·80
		5·64	0·43			2·25	11·32†
				{0·78	{0·80		
	ADJ	0·51	0·51	6·48	2·77	0·60	0·55
		5·67	3·82			0·98	0.13
	R.S.	0·53	0·52	0·53	0·55	0·57	0·55
		4·26	4·30	0·24	4·28	2·12	0·10
B	MAX	0·80	0·76			0·85	0·77
		12·18†	3·68			1·60	6·78
				{0·76	{0·86		
	ADJ	0·55	0·5+	5·69	0·65	0·69	0·48
		0·01	1·47			0·22	1·27
	R.S.	0·56	0·51	0·51	0·62	0·59	0·53
		0·02	2·27	0·18	9·09†	2·94	0·34
C	MAX	0·80	0·76	0·79	0·82		
		16·29†	0·11	2·46	5·48		
						{0·79	{0·81
	ADJ	0·57	0·48	0·55	0·60	13·58†	8·33
		0·07	7·56	0·37	0·05		
	R.S.	0·56	0·51	0·54	0·57	0·54	0·56
		0·02	10·07†	0·50	0·27	0·05	0·13
D	MAX	0·78	0·77	0·75	0·85		
		8·38	6·60	21·95†	1·00		
						{0·86	{0·76
	ADJ	0·55	0·54	0·48	0·69	1·89	16·99†
		0·22	0·04	1·48	0·65		
	R.S.	0·53	0·52	0·50	0·60	0·62	0·52
		0·08	0·01	2·08	3·62	2·80	0·32
E	MAX			0·79	0·82	0·82	0·80
				2·61	11·75†	3·49	2·13
		{0·79	{0·79				
	ADJ	41·32†	3·59	0·54	0·55	0·54	0·55
				1·22	0·05	0·49	0·70
	R.S.	0·55	0·54	0·54	0·57	0·57	0·55
		7·65	4·42	1·19	0·31	0·23	0·73
F	MAX			0·75	0·85	0·85	0·77
				5·99	1·73	29·32†	12·12†
		{0·79	{0·74				
	ADJ	3·81	0·49	0·51	0·51	0·48	0·57
				1·03	3·77	0·06	0·09
	R.S.	0·54	0·50	0·50	0·60	0·59	0·53
		2·03	2·92	2·17	1·28	1·79	0·11

* First entry in each cell is predicted proportion correct; second entry is F-ratio for predicted vs. obtained score. Bracketed values represent coincident predictions.

† $P < 0·01$.

the minor triad. Table II presents mean proportion correct judgments for each level of experience for the interval test. The rows of Table II indicate size of the standard interval generated from the three-note test and the columns represent size of incorrect transposition (to an interval one semit smaller or one semit larger). There are (unpredicted) statistically significant differences among the cell entries of Tables I and II that are reported in detail by Cohen (1972). These differences are not critical, however, to the question of immediate interest—whether or not the interval data can predict the cell entries for the three-note data.

The next analysis examined predicted values for each cell of the 36-cell matrices of Table I. Predictions are shown in Tables III, IV and V for the three levels of experience respectively, and a summary of the predictions is given in Table VI. The construction of Tables III, IV and V is similar to that of Table I, with rows indicating pattern label and columns representing type of incorrect transposition. The matrices are now expanded to include proportion of correct judgments predicted by each model, which is the first entry in each cell, and an F-ratio, which is the second entry in each cell. The F-ratio was used to test the significance of the difference between observed and predicted mean scores; the error term for the F-ratio was the within-subject within-cell variance of the difference and F is thus distributed on 1 and 13 degrees of freedom. Bracketed values in Tables III to VI represent instances where the predictions of models MAX and ADJ by definition coincide.

The rows of the summary table, Table VI, refer to the predictions of each of the three models for each of the three levels of musical experience. The columns separate the 12 conditions where the incorrect transposition involved alteration of the two intervals between adjacent tones from the 24 conditions where the incorrect transposition involved alteration of one interval between adjacent tones and one interval between non-adjacent tones. The column heading "Rejections" refers to the number of times that the difference between observed and predicted values was significant at the 0·01 level. The column heading "Best-fit" refers to the number of times each model provided the closest prediction of obtained values, in terms of standard scores. Table VI shows the relative standing of each model, but no further significance tests are implied. In fact, as will be seen, it is probably inappropriate to attempt to derive an over-all significance level for each model.

Discussion

Figure 4 shows that, under the present conditions, increasing the length of a stimulus sequence leads not to interference with recognition performance, but to facilitation. It suggests that there is a more effective encoding process to handle three-note as opposed to two-note test sequences, and that musically-trained listeners are particularly adept at this form of encoding. However, the question whether or not the process involves the abstraction and synthesis of interval information is difficult to answer.

A preliminary requirement for the application of the proposed models is, of course, that performance on either test is not chance or random. Tables I and II show all mean values (extreme right-hand columns) to be above the chance

TABLE IV

*Predictions from three models for group T**

Pattern	Model	Incorrect transposition					
		LL	RL	LM	RM	LH	RH
A	MAX	0·81	0·83			0·87	0·80
		1·78	1·16			0·05	0·68
				{0·79	{0·86		
	ADJ	0·59	0·55	0·85}	0·44}	0·69	0·53
		2·14	51·97†			4·90	16·20†
	R.S.	0·58	0·59	0·54	0·64	0·66	0·55
		3·02	37·34†	4·16	14·74†	8·75	18·08†
B	MAX	0·85	0·82			0·91	0·77
		1·16	2·28			2·19	1·66
				{0·76	{0·91		
	ADJ	0·64	0·53	1·12}	0·18}	0·78	0·46
		2·19	4·36			0·17	13·30†
	R.S.	0·62	0·56	0·50	0·71	0·69	0·53
		3·32	4·95	1·83	18·58†	3·22	5·83
C	MAX	0·85	0·82	0·78	0·88		
		6·24	0·76	6·82	1·67		
						{0·88	{0·82
	ADJ	0·60	0·59	0·53	0·69	3·76}	0·02}
		0·62	2·61	0·26	0·51		
	R.S.	0·62	0·56	0·53	0·67	0·64	0·57
		0·40	3·72	0·28	1·15	4·52	14·59†
D	MAX	0·81	0·83	0·76	0·90		
		2·90	1·27	10·17†	0·08		
						{0·90	{0·76
	ADJ	0·57	0·63	0·46	0·78	4·35}	2·17}
		2·23	8·83	0·65	1·90		
	R.S.	0·58	0·59	0·51	0·69	0·71	0·52
		2·06	19·89†	0·04	7·57	0·33	1·20
E	MAX			0·78	0·88	0·87	0·80
				2·94	4·53	8·89	0·02
		{0·84	{0·83				
	ADJ	12·30†}	0·57}	0·53	0·64	0·63	0·57
				1·71	2·68	0·18	17·17†
	R.S.	0·61	0·58	0·53	0·67	0·66	0·55
		0·50	13·16†	2·03	1·33	0·01	12·81†
F	MAX			0·76	0·90	0·91	0·77
				9·45†	5·15	24·12†	0·03
		{0·82	{0·82				
	ADJ	8·33}	4·05}	0·55	0·59	0·59	0·60
				0·06	5·63	0·60	5·59
	R.S.	0·60	0·57	0·51	0·69	0·69	0·53
		0·09	1·46	1·09	1·50	5·20	9·20†

* Cell entries represented as in Table III.
† $P < 0{\cdot}01$.

TABLE V

*Predictions from three models for group H**

Pattern	Model	Incorrect transposition					
		LL	RL	LM	RM	LH	RH
A	MAX	0·92 0·10	0·87 0·03			0·95 1·03	0·92 0·17
	ADJ	0·59 65·01†	0·50 26·87†	{ 0·80 11·83†	{ 0·92 11·30†	0·82 30·49†	0·60 31·85†
	R.S.	0·70 46·14†	0·61 15·75†	0·55 90·77†	0·70 89·73†	0·77 27·44†	0·70 17·23†
B	MAX	0·94 0·77	0·89 0·08			0·98 19·34†	0·91 2·35
	ADJ	0·76 2·22	0·52 28·37†	{ 0·80 0·19	{ 0·98 0·81	0·92 19.10†	0·60 6·77
	R.S.	0·76 2·51	0·65 16·48†	0·56 15·90†	0·84 13·04†	0·85 67·99†	0·69 3·09
C	MAX	0·94 2·03	0·89 0·69	0·79 1·64	0·95 0·01		
	ADJ	0·77 1·83	0·77 0·52	0·60 20·98†	0·82 4·57	{ 0·96 2·00	{ 0·91 5·14
	R.S.	0·76 2·16	0·65 7·79	0·56 33·35†	0·79 10·18†	0·80 5·65	0·68 50·20†
D	MAX	0·91 3·81	0·87 0·05	0·80 0·50	0·96 0·81		
	ADJ	0·81 0·71	0·72 3·98	0·60 4·35	0·92 0·11	{ 0·98 1·31	{ 0·93 3·83
	R.S.	0·70 0·17	0·61 25·57†	0·55 13·32†	0·75 19·22†	0·82 5·97	0·71 1·19
E	MAX			0·79 0·01	0·95 3·05	0·95 4·49	0·92 0·15
	ADJ	{ 0·95 1·38	{ 0·86 2·30	0·52 7·65	0·76 1·07	0·72 0·99	0·81 2·66
	R. S.	0·79 2·14	0·62 1·87	0·56 5·76	0·79 0·34	0·77 0·28	0·70 13·20†
F	MAX			0·80 0·25	0·96 8·14	0·98 5·02	0·91 0·29
	ADJ	{ 0·91 0·01	{ 0·88 3·69	0·50 8·03	0·58 6·46	0·78 3·24	0·77 2·68
	R.S.	0·68 21·27†	0·64 3·48	0·55 6·22	0·75 0·27	0·85 0·46	0·69 9·28†

* Cell entries represented as in Table III.
† $P < 0·01$.

TABLE VI

Summary of rejected models and best-fitting models

Musical experience	Model	Adjacent tones ($n = 12$) Rejections	Adjacent tones ($n = 12$) Best-fit	Adjacent + Non-adjacent Tones ($n = 24$) Rejections	Adjacent + Non-adjacent Tones ($n = 24$) Best-fit
U	MAX	3	5	7	2
	ADJ			0	13
	R.S.	1	7	1	9
T	MAX	1	7	3	12
	ADJ			4	6
	R.S.	4	5	5	6
H	MAX	2	9	1	15
	ADJ			7	4
	R.S.	6	3	13	5

level of 0·50, but closer examination of the cell entries for group U shows two entries for the interval test below 0·50, and five entries for the three-note test below or exactly 0·50. This result, coupled with the generally low values for group U, indicates that the fit of Models ADJ (where it does not coincide with MAX) and R.S., both of which predict 0·50 for chance performance on both tests, may be spurious in many cases. Even if not so, an over-all distinction between the two models is virtually impossible. Performance for group U is disappointingly poor, especially on the interval test; other data we have since collected support the notion that interval recognition is extremely difficult for the untrained listener.

Despite this finding, two points may be added: first, model MAX, where it does not coincide with model ADJ, is clearly unable to describe performance for group U. Second, however, there is some evidence for some experimental conditions that listeners in group U attempt to abstract at least one interval between adjacent tones. Highest performance on the three-note test occurs where one component interval condition is the interval condition most easily discriminated by all listeners—3 vs. 2 semits, ascending or descending. Six of the eight three-note conditions containing the easiest interval discrimination between adjacent tones produced performance scores above 0·68, while only two of the 24 three-note conditions not containing the discrimination exceeded 0·68. The probability of obtaining this finding by chance is less than 0·01.

For the musically-trained listeners, the fit of model MAX must be seriously considered. Model MAX, or the coincidence of models MAX and ADJ, was rejected least often, and most frequently provided the best fit to the data. It is therefore tempting to conclude that, for the musically-trained listeners, recognition of a three-note transposed sequence involves a process of abstraction and synthesis of intervals occurring between adjacent and non-adjacent tones.

Other aspects of the data, however, make such conclusions tentative. One is that even the most appropriate model was rejected more often than expected under the null hypothesis. Another is that the most appropriate model tended generally to overestimate performance values for patterns D, E, and F, although it must also be noted that the three cases of rejection of models MAX or the coincidence of MAX and ADJ for group H were instances where the prediction for comparisons involving patterns A and B actually significantly underestimated performance. This leads to a further, related, difficulty—the best-fitting model for each level of musical experience failed to predict the significant effect of pattern in observed performance. The main effect of pattern ($F = 6{\cdot}59$, $df = 5, 195$, $P < 0{\cdot}001$) reflects the finding that, over all subjects, there is a decrease in accuracy of recognition across patterns A to F; in particular, patterns A and B are more easily recognized in transposition than patterns C to F. This effect did not interact with level of experience.

Possibly the fit of the most appropriate model would be improved with modification of experimental procedures. On the other hand, the failures of the most appropriate model may necessitate a re-examination of basic assumptions, in particular the assumption of independence. With respect to the first possibility, three procedural details were examined—number of trials, classification of subjects, and selection of patterns. There was no evidence of practice or order effects to suggest improvement of fit with more or fewer trials or a different experimental design. Although we have not found a more reliable system of subject classification than the "subject profile" used here, subject classification will doubtless improve when the mechanisms of auditory memory are better understood. However, since the main effect of pattern did not interact with level of experience, reclassification of subjects would not alter the basic problem of interpreting the the pattern effect.

Finally, one might ask whether or not the three-note standard of each trial ought to have been a random selection of three tones. Note, first of all, that if a model of interval abstraction is seriously to be considered for the representation of musical perception it must be properly tested within a musical framework and must be applicable to the perception of familiar musical forms. Second, the patterns and incorrect transpositions employed produced a wide range of performance values (across levels of experience from slightly below chance to perfect performance; even within group H, from about 75% correct to 100% correct). The task was by no means trivial. Third, the data show that the amount of within-subject within-cell variability of the difference between observed and predicted scores is dependent upon the particular configuration of the intervals. It is clearly necessary to examine the scores for each pattern and incorrect transposition individually, and while, of course, a wider variety of patterns should be studied, it is not appropriate to average data across different patterns for which the tones are randomly selected from trial to trial. The latter procedure might well inflate the error variance term used to test the fit of the models: the fit of all models would be improved but not, necessarily, our understanding of the basic processes.

There is then the problem of the assumption of independence—under this assumption the models ignore the forms of the patterns themselves. Patterns A and B may be considered "good" forms in a general sense; for one thing monotonically ascending and descending sequences belong to a set of few (i.e. two) alternatives (cf. Garner, 1970, 1974; Royer and Garner, 1970). They may also be the most familiar forms of the triad, at least as far as musical exposure is concerned (Dykema and Cundiff, 1955; Lieberman, 1956; Shera, 1922). At any rate, the most important point is that the facility with which a sequence of three notes of the triad is recognized depends not merely on the intervals contained in the sequence but upon the interval configuration.

The data thus provide evidence in support of a very plausible musical notion—the triad is a higher-order structure that is not merely the sum of the component intervals. Particular interval configurations may facilitate the abstraction of structure. It may be suggested further that the facility with which a structure is abstracted may depend upon the number or difficulty of temporal transformations that intervene between sensory input and the stored archetypal form of the triad. The process cannot, however, be identical to that proposed by Deutsch (1969) for chord abstraction; Deutsch's description of chord abstraction is not intended to deal with the order or configuration of the intervals comprising the chord.

A further relevant finding is that the most difficult transposition across levels of experience, and patterns, is LM, the shift to a minor triad. It may be that the similarity of the archetypal structure for the major and the minor triad is a potential source of confusion. On the other hand, this finding *is* predicted by the best-fitting interval abstraction model. The shift from major to minor triad involves the discrimination between the intervals of 3 and 4 semits (the minor and the major third, respectively). Table II shows that this discrimination is particularly difficult. (It remains to be explained, however, why this discrimination is significantly more difficult for musically-trained listeners than the discrimination of a semit difference between the larger intervals.)

The abstraction of structure may occur either in addition to, or as alternative to, interval abstraction and synthesis. Certainly, before a theory of interval abstraction and synthesis is rejected, other forms of prediction should be studied. The data suggest that it would not be profitable to pursue single-interval models such as model R.S. for musically-trained listeners. Even a model proposing that three-note recognition was based on (*a priori*) knowledge of which changed interval was most likely to be judged correctly would underestimate the responses to patterns A and B 97% of the time. More intriguing, although the design of the present experiment does not allow a precise test, is the possibility that the tendency of model MAX to overestimate performance values would be corrected by building in to the model a factor reflecting the interference created by an interpolated tone with the comparison of non-adjacent tones (Deutsch, 1970) or interference with the process of synthesis itself. In the derivation of such a correction factor, however, the main effect of pattern must also be taken into account.

In conclusion, it may be suggested that interval abstraction and synthesis play a role in melody recognition, but that there are severe limitations to interval processing. Recognition of a three-note sequence based on the triad may depend at least in part upon the facility with which a tonal pattern is abstracted; what is needed is an independent measure of the structural cues provided by a tone sequence, and research now underway is addressing this problem.

This research was supported by grants to L. L. Cuddy from the National Research Council of Canada (AO333), the Defence Research Board of Canada (9425-17) and the British Council. The data were initially presented in an M.A. thesis by A. J. Cohen, and this paper was completed while L. L. Cuddy was on sabbatical leave at the Department of Experimental Psychology, University of Oxford. We acknowledge with thanks the critical advice and assistance of Professor M. G. Wiebe of Queen's University and the helpful comments of an anonymous reviewer. A. J. Cohen is now in the Dept. of Psychology, University of Waterloo, Ontario.

References

COHEN, A. J. (1972). *Melodic Interval Abstraction and Melodic Sequence Recognition.* Unpublished M.A. thesis, Queen's University.

CUDDY, L. L. (1968). Practice effects in the absolute judgment of pitch. *Journal of the Acoustical Society of America*, **43**, 1069–76.

CUDDY, L. L. (1970). Training the absolute identification of pitch. *Perception and Psychophysics*, **8**, 265–9.

CUDDY, L. L. (1971). Absolute judgment of musically-related pure tones. *Canadian Journal of Psychology*, **25**, 42–55.

DEUTSCH, D. (1969). Music recognition. *Psychological Review*, **76**, 300–7.

DEUTSCH, D. (1970). Tones and numbers: specificity of interference in short-term memory. *Science (New York)*, **168**, 1604–05.

DYKEMA, P. W. and CUNDIFF, H. M. (1955). *School Music Handbook: A Guide for Music Educators.* Evanston: Summy-Birchard.

GARNER, W. R. (1970). Good patterns have few alternatives. *American Scientist*, **58**, 34–42.

GARNER, W. R. (1974). *The Processing of Information and Structure.* Potomac, Maryland: Erlbaum.

HUBEL, D. H. and WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, **160**, 106–54.

LIEBERMAN, M. (1956). *Ear Training and Sight-singing.* New York: Norton.

ROYER, F. L. and GARNER, W. R. (1970). Perceptual organization of nine-element auditory temporal patterns. *Perception and Psychophysics*, **7**, 115–20.

SHERA, F. H. (1922). *Musical Groundwork: A Short Course of Aural Training.* Toronto: Oxford University Press.

Revised manuscript received 23 May 1975