Canadian Acoustics / Acoustique Canadienne 18(4) 3-10 (1990) ***Research article / Article de recherche***

SCALE AND SERIAL ORDER INFORMATION IN MELODIC PERCEPTION: INDEPENDENCE OR INTERDEPENDENCE?

Annabel J. Cohen and Bradley Frankland
Department of Psychology, Dalhousie University
Halifax, Nova Scotia

ABSTRACT

Listeners indicated the temporal order of 8 tones presented in 36 different sequences representing all combinations of four 8-note scales or alphabets (chromatic, major, minor, and augmented) and nine serial orders (of varying complexity). Accuracy, as measured by ordinal position errors and by contour errors, was poorest for the chromatic alphabet and for the most irregular sequential structures. Alphabetic effects were partially accounted for by familiarity and discriminability and serial order effects were partially accounted for by rule regularity. Effects of alphabet structure and sequential structure were, however, not independent: certain combinations of alphabets and sequential orders led to error patterns that were not consistent with the main effects. The question remains whether such deviations reflect interdependent scale and order processes.

SOMMAIRE

Les auditeurs indiquerènt l'ordre temporel de 8 notes dans chaqu'un des 36 mélodies différentes qui réprèsenterent toutes les combinaisons de quatre gammes (chromatique, majeur, mineur, et augmente) et neuf ordres temporels (de compléxites divers). Le choix de la gamme et de l'ordre influa la précision de la réponse des auditeurs, mesurée par leurs erreurs d'interpretation de l'ordre des notes et du contour de la mélodie, avec le moins de précision dans le cas de la gamme chromatique et l'ordre le plus irrégulier. On explique en partie l'influence de la gamme par sa familiarité et sa facilité de discrimination, et l'influence de l'ordre par sa régularité de structure. Cependant, les influences de la gamme et de l'ordre n'étaient pas indépendantes: certaines combinaisons de gamme et d'ordre produisirent des erreurs qui n'étaient pas d'accord avec les premiers effects. Il reste de déterminer si ces écarts indiquaient une interdépendance entre la gamme et l'ordre.

1. INTRODUCTION

Recent research in music perception has revealed that sets of tones, or scales, used in Western-European music are mentally represented not as unrelated individual tones but rather as a hierarchical structure (Krumhansl, 1983; Cuddy & Badertscher, 1987). Similarly, the representation of sequential orders of tones in music is sensitive to the presence of patterns (Martin, 1972). Thus, the weight given to the representation of a particular tone of a sequence depends upon the role of that tone in both the tonal hierarchy of the scale and the sequential pattern context.

Independence of the representations of scale and sequential rules in music is implicit in many theoretical accounts of melody perception (Deutsch & Feroe, 1981; Restle, 1971; Simon & Sumner, 1968). Tests of independence of musical dimensions have been carried out by Pitt and Monahan (1987). They observed that similarity ratings of three polyrhythms were influenced by the choice of tones comprising the polyrhythms. "Pitch information had a uniform effect on polyrhythm similarity, systematically increasing or decreasing the similarity among rhythms by roughly the same amount (p. 534)." In other words, the perceived similarity of two sequences was based on additive components of the similarity of their scales and the similarity of their rhythms.

A study by Warren and Byrnes (1975) on the other hand demonstrated an interdependence of pitch alphabets and sequential information. Their stimuli were four sets of four equidistant tones separated by .3, 1, 4, or 9 semitones recycled in one of six possible orders (i.e., 1234, 1243, 1324, 1342, 1432, 1423). Naming the correct order was affected by the alphabet; in general, the smallest intervals between tones led to poorest performance. In addition, ascending (e.g., 1234) and descending (e.g., 1432) contours were less affected than complex patterns having more than one contour (i.e., up and down) change (e.g., 1324). This observation contrasts with the demonstration of independence of pitch and rhythmic information of Pitt and Monahan (1987), in the sense that both contour and rhythm are sequential variables which were applied to different tone sets or scales. The present paper addresses the fundamental question of the independence of scale and temporal order using a new melodic tracking task, by systematically varying the stimulus structure of alphabet and serial order.

2. METHOD

2.1 Task

It was the task of the subject to represent the order of an 8-tone melody on time-frequency coordinates with the x-axis representing 8 equal time units and the y-axis representing 8 positions for frequency (See Figure 1). The 8-tone melody had 8 different frequencies, that is, each of 8 frequencies was used only once. Each sequence began on the lowest frequency of its scale. Following five consecutive repetitions, the subject was to fill in the time-frequency graph represented as an 8 x 8 grid. Subjects knew *a priori* that the first note would be the lowest tone; that is, the lower left cell (1,1) was to be filled in automatically, leaving 7 additional judgments to be made.

2.2 Stimuli

Scales

As shown in Figure 2, the scales chosen were the major (successive intervals 2212221), minor (2122122), and chromatic (1111111), as well as a scale of successive intervals of four semitones (4444444). They are subsequently referred to as M, m, N and W representing the Major, minor, Narrow and Wide spacing patterns. In musical terminology, the latter scale could be described as augmented triadic. The augmented triad is not found in the diatonic (major/minor) scales, is unstable (Roberts & Shaw, 1984) and is difficult to remember as compared to a diatonic sequence (Cohen, Trehub, & Thorpe, 1989). The diatonic scales (M, m) on the other hand are characteristic of most music. Scales of equal intervals such as the chromatic (N) and augmented (W) are less familiar than the diatonic scale. In addition, their higher uncertainty might lead to poorer information processing. We speak of uncertainty here in the context of classical information

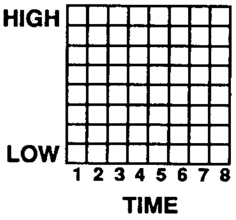

Figure 1. Matrix which was completed by the subject for each sequence.

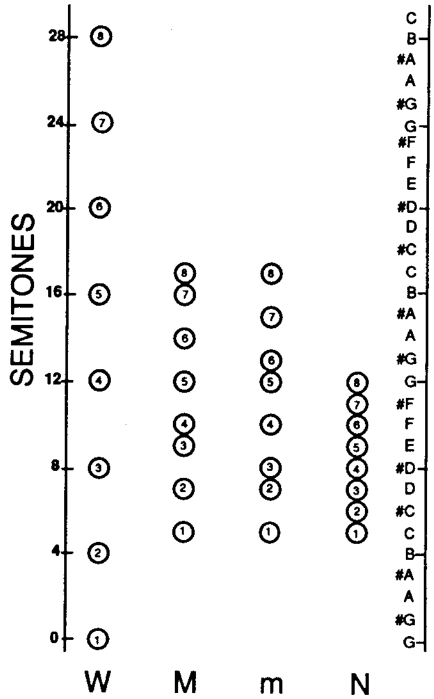

Figure 2.. Spacing of the four scales in semitone units (W=wide, M=Major, m=minor and N=Narrow)

theory where highest uncertainty is associated with equiprobability of the events. In the chromatic scale, intervals of one semitone are equiprobable but in the diatonic they are not; thus, in objective terms, the diatonic scale is less uncertain than the chromatic. Reduction in objective uncertainty has been shown to increase processing efficiency in a variety of tasks (cf., Garner, 1962; 1974). Both Helmholtz (1885/1954) and Dowling (1978) noted that there are no naturally occurring musical scales of equal intervals. This rejection of equal-interval scales in music crossculturally is consistent with the processing difficulties that such scales might create as a result of their uncertainty. In the present experiment, the two interval widths (1 and 4 semitones) allowed for comparison of effects of size of the unit in the equal interval scales. The larger size would afford greater resolution, but this advantage might be counteracted by the wider frequency band over which attention must be directed.

Serial Orders

Just as sets of frequencies can differ structurally, so sequential orders can vary in degree of organization and processing difficulty (e.g., Restle, 1971; Warren & Byrnes, 1975). For example, a simple ascending sequence, such as 12345678 should be more easily remembered than an irregular sequence, such as 17264538. A set of nine different sequences were chosen to represent a range of complexity (see figure 3). This choice was made on the basis of intuition guided by past research on subjective measures of complexity for simple auditory and visual patterns (cf., Garner, 1974). Some of the patterns were easily described by a rule (e.g., order 2, start at the bottom go up 4, start at the top go down 4) but others such as order 9 were not. The description of complexity was not straightforward because of the necessity of considering more than one hierarchical level, or macrocontour (Cohen, et al., 1989). In other words, complexity was not simply a matter of counting the number of contour changes, because, in some cases the pattern of contour led to another higher level of organization. For example, the repetitious up-down pattern of order 4 produces an overall V-shape (90 degree rotation), whereas in sequence 9, the same number of contour changes produces no such obvious higher order structure. It was difficult to rank the serial rule complexity *a priori*, it was nevertheless possible to identify the simplest (1), the next simplest (2) and the most complex pattern (9), with the remainder falling between.

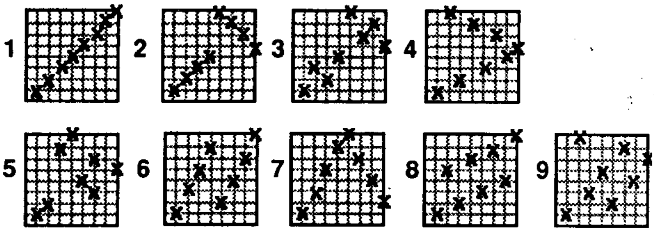

Figure 3. Representation of 9 sequential orders of 8 tones. Each matrix is a reduction of that shown in Figure 1.

2.3 Stimulus Generation

The 36 sequences produced by combining the four scales and nine serial orders were represented by sine tones generated digitally with 12-bit digital-to-analogue conversion. The 36 different melodies were recorded in random order on low noise tape using a Revox A77 reel-to-reel tape recorder. Two practice trial sequences were also recorded. Tone duration was 450 ms, intertone interval was 70 ms, and intertrial interval was 12 s. Thus, the stimulus tape was 21 min. The lowest frequency was 262 Hz (C_4) for all scales but the augmented for which the lowest frequency was 196 Hz (G_3).

2.4 Subjects

Nineteen Introductory Psychology students enrolled in a music perception seminar were tested in two groups with 11 and 8 people per group.

3. RESULTS

Two dependent measures of accuracy were considered in detail. The first, Absolute Ordinal Position Error, represented the total absolute magnitude of the ordinal errors for all 8 serial positions of the sequence. For example, if the correct order was 12345678 and the response was 14623578, the total error score would be the sum of the absolute discrepancies (i.e., 0+2 +3+2+2+1+0+0=10). The second error, Simple Contour, measured incorrect melodic contour, where contour refers to the pattern of up and down in the melody. Each incorrect change in direction within a sequence was scored 1. The correct contour always resulted in a score of 0.

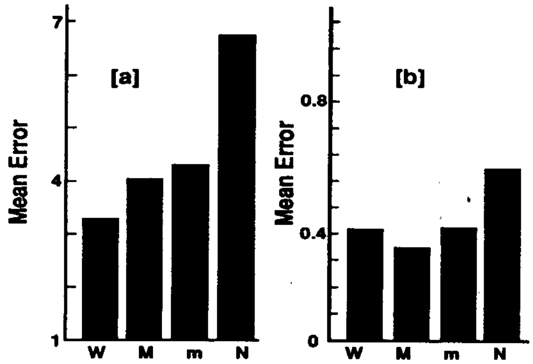

Figure 4. Mean errors as a function of scale type (a) Absolute Ordinal Position Errors (b) Simple Contour Errors

Each time a subject chose the incorrect direction, a score of 1 was added. In the previous example, the correct contour was +++++++ and the response was ++-++++ resulting in a score of 1. Since subjects were permitted to fill in the same ordinal position more than once, a contour error could be potentially created by an immediate repetition of the preceding pitch. Thus, the maximum error score per sequence was 7. Contour coding has been found to be more resilient than absolute interval information (Dowling, 1978).

For both measures, the mean error differed as a function of scale and serial order as shown respectively in Figures 4 and 5. For example, errors for the chromatic alphabet (N) were higher than those for the other three alphabets, and errors for orders 1 and 2 were lower than those for order 9. For all 36 sequences, the mean Absolute Ordinal Position Errors and mean Simple Contour Errors are shown in Figure 6(a) and (b) respectively. For each error type, the shapes of the histograms differ as a function of the nine serial orders suggesting interactions between scale and serial order. For each error type, the data were entered into an analysis of variance having two within-subjects factors of scale (4 levels) and order (9 levels). Orthogonal contrasts for scale compared the wide spacing against the other three types, the narrow spacing against the two diatonic spacings, and finally the major vs. the minor scale. Contrasts were also carried out for order pitting particular orders against each other. These sets of contrasts for main effects led to 24 contrasts for the interaction.

For each error type, the two main effects and their interaction were significant. For Absolute Ordinal Position Error for the main effect of scale, $F(3,54)=25.75; p<.001$, orthogonal comparisons revealed that wide spacing significantly produced the highest performance and narrow spacing the lowest. The main effect of order was also significant, $F(8,144)=11.3; p<.001$. Among a number of significant differences, the most complex pattern (9) was significantly more difficult than the simplest patterns (1 and 2) and pattern 2 was significantly more difficult than pattern 1. The interaction between scale and order was significant, $F(24,432)=2.42; p<.001$; four of the orthogonal comparisons for the interaction reached significance although no particular interaction component was pronounced.

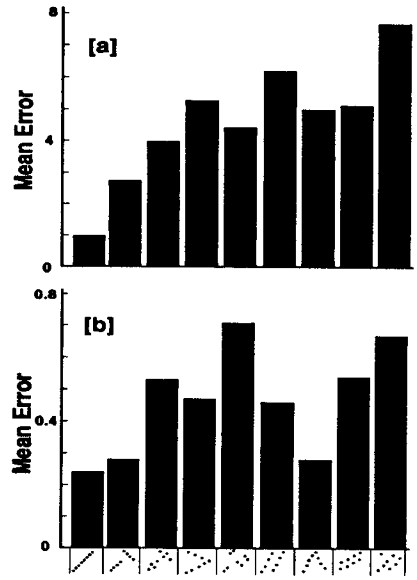

Figure 5. Mean errors as a function of sequential pattern (a) Absolute Ordinal Position Errors (b) Simple Contour Errors

For Simple Contour Error, the main effect of scale, $F(3,54)=4.07; p<.05$, was attributable to a difference between the narrow and diatonic spacings. For the main effect of order, $F(8,144)=2.15; p<.05$, orders 1 and 2 were significantly easier than order 9, and order 5 was significantly more difficult than order 7. For the interaction of scale and order, $F(24,432)=1.84; p<.01$, 5 of 24 orthogonal comparisons were significant (2 of which were common with the Absolute Ordinal Position Error interaction) but no interaction component was pronounced.

Considering the results for both Absolute Ordinal Position Error and Simple Contour Error measures described above,

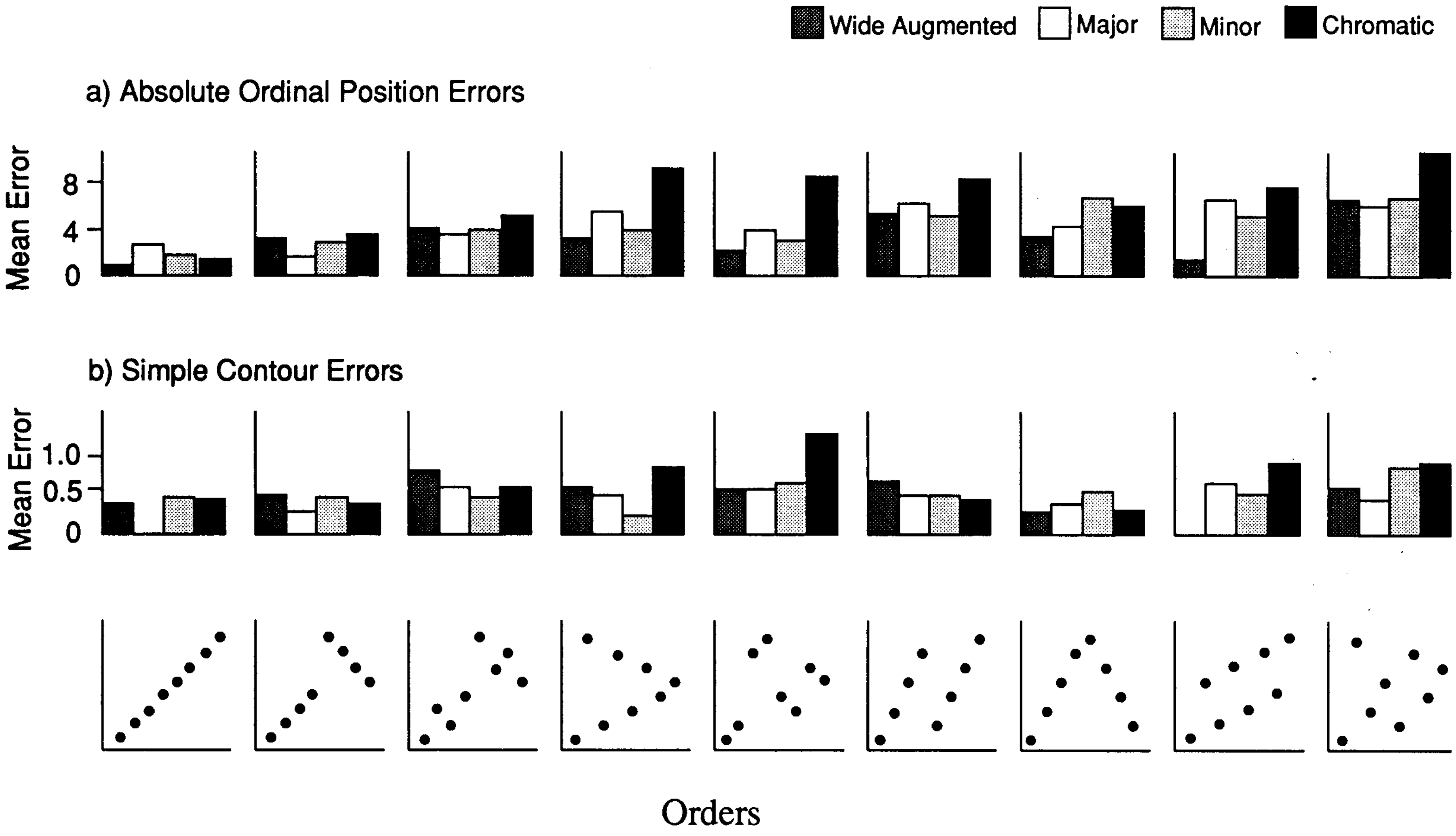

Figure 6. Mean Errors as a function of scale and order a) Absolute Ordinal Position Errors b) Simple Contour Errors

an interaction between scale and serial order clearly has been demonstrated. In order to determine whether experience in the task improved performance and therefore might account for the scale x order interaction, a correlation was computed on the 36 mean scores and their order of presentation in the session. The resulting correlation coefficients were -.25 for Absolute Ordinal Position Error and -.10 for Simple Contour Error, neither of which were significantly different from zero in a t-test. Therefore experience in the task did not account for the scale and order interaction.

Four other measures of error were also examined. The first (Binary Overall) was a binary score assigning the value of 1 for perfect performance on the sequence and 0 if any error were made. The second (Binary Serial Position) was a binary score for each serial position of the sequence; thus, a score of 1 was assigned to each position incorrect for a maximum of 8. The third measure (Complex Contour) was a variation on the previous Simple Contour error measure. In this case, the direction was measured with respect to the previous adjacent tone but the magnitude of the change was also considered. This method thus penalized the subject who lost footing at the contour change. For the final measure, the correlation was obtained between the 8 responses and the 8 correct values. Perfect performance produced a value of 1.0 and the lowest possible value was -1.0. These measures were obtained for all subjects and the means for each sequence were entered into a 6 (measures) x 36 (sequences) Spearman Rank intercorrelation matrix, the results of which are shown in Table 1.

It is noted that all correlation coefficients are statistically significant beyond the .01 level, however, the Simple Contour measure produces the lowest correlations ranging from .61 to .75. All of the other measures produce correlations greater than .84. Indeed, a crude measure of whether judgments of the sequence was completely correct or not (Binary Overall) provides roughly the same information as the more sophisticated measures such as Absolute Ordinal Position Error. This pattern of results emphasizes the robustness of the findings and suggests that a different process is tapped by the Simple Contour measure in comparsion with the remaining five.

Table 1

Spearman Rank Correlations of 6 Error Measures

	Type of Measure					
	Absolute Ordinal	**Simple Contour**	**Binary Overall**	**Binary Ser.Pos.**	**Complex Contour**	**Corre-lation**
Absolute Ordinal	1.0	.61	.93	.96	.90	.93
Simple Contour		1.00	.63	.57	.75	.73
Binary Overall			1.00	.97	.90	.84
Binary Ser. Pos.				1.00	.88	.86
Complex Contour					1.00	.94
Corre-lation						1.00

4. DISCUSSION

The main finding of the study is that accuracy of reproduction of the sequential order of a set of tones depends upon the particular serial order and the particular scale. The scale with the widest spacing, which afforded the greatest frequency resolution, led to highest performance on the Absolute Ordinal Position Error. This is consistent with Warren and Byrnes (1975); however, all of their scales were composed of equal intervals. The familiar diatonic scales in the present study did not produce an advantage over the widely spaced equal-interval scale for Absolute Ordinal Position Error but did show a large advantage over the narrow equal-interval spacing. For Simple Contour Error, there was no difference between the wide and diatonic scales. It is, thus, likely that both familiarity and discriminability had an influence. To test this assumption in the future, the diatonic scale should be compared against equal-interval scales of two semitones as well as unfamiliar scales built from one and two semitone intervals. If familiarity plays a role, the diatonic scales should lead to higher performance than the other scales mentioned. The results for serial order are consistent with the notion that sequential rule regularity played a role: the poorest performance was associated with the most complex order and the best performance with the simplest order. In future work, the orders should be selected in terms of a rule-regularity measure objectively defined. These main effects of scale and order just described do not, however, account for all aspects of the data. In addition, there was evidence for effects of the interaction of order and scale for ordinal errors summed across serial position and for errors of contour.

Theories of melodic perception have often focused on the dichotomy of a tone alphabet to be ordered sequentially and rules for ordering tone alphabets (e.g., Cohen, 1976; Deutsch & Feroe, 1981; Restle, 1971; Simon, 1971; Simon & Sumner, 1968). These models are consistent with the significant main effects of scale structure and sequential structure in the present task but do not account for interactions between alphabet and serial order. They cannot explain, for example, that when order 8 is paired with the widely spaced alphabet, performance, as measured by absolute ordinal position errors, is superior to order 3, but when paired with any of the other alphabets, the relative level of difficulty reverses.

Our observed interaction between alphabet and serial order is, however, conceptually consistent with the views of Jones (1987) and her colleagues who have investigated the effects of combining different rhythms and contours. They observed evidence of higher accuracy in notating melodies for some rhythmic/contour combinations but not others (Boltz & Jones, 1986). This violated the notion of independence, that "the temporal grouping properties of a rhythm will confer equivalently stable accentuations on a sequence... regardless of the way in which the melodic [contour] relationships fit the temporal cues" (Jones, Summerall, & Marshburn, 1987, p. 98). Jones and her colleagues did not, however, specifically address the issue of the combination of pitch alphabet and serial rules. What is novel about the present contribution is the evidence that the processing of sequential order is not independent of alphabet structure.

Lower-level variables may contribute to these interactions, for example, discriminability of tones (Warren & Byrnes, 1975) and perceptual grouping on the basis of pitch proximity (Bregman, 1978; Deutsch & Feroe, 1987). Nevertheless, the question remains as to whether performance depends upon higher level congruencies based on a shared pattern of dominance between atemporal (scale) and temporal (serial order) relations. As an example from the present study, could interdomain congruence account for the fact that wide spacing led to greater accuracy for pattern 8 (15263748) than did diatonic spacing? For wide spacing, after serial position 2, every other tone is succeeded by a tone one octave below, whereas for the diatonic sets, the relation is a fifth or a tritone. Matched hierarchies across atemporal and temporal domains could account for this interaction of scale and serial order information.

Correlations of mean performance and position in the experimental session suggested little effect of learning. It is possible that effects of immediately prior sequences may have influenced certain judgements. Since all subjects received the same order of sequences, this situation could have contributed to the observed interactions. To rule out this possibility, in future studies, different random orders of the sequences should be provided for each subject.

In the present study, sequences were presented five times in succession before the subject responded. It is possible that interdependence emerges only after a number of presentations. Thus, notions of the independence of atemporal and temporal rules must be modified to include the developing interdependence based on congruencies that emerge with repetition. But the primary point remains: although the distinction between scale and serial order information is a good starting point for some theories of melodic perception, the problem of the interaction between these two domains must be specifically addressed at perceptual and cognitive levels.

Note. The research was supported by grants from the National Research Council and the Natural Sciences and Engineering Research Council to A. J. Cohen. A portion of these data were reported at the Annual Meeting of the Canadian Acoustical Association, Toronto, 1988, and appear in the *Proceedings* of that meeting. The assistance of Debora Dunphy with both graphics and formatting and the insightful comments of two anonymous reviewers are gratefully appreciated.

REFERENCES

Boltz, M. & Jones, M.R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, **18**, 389-431.

Bregman, A.J. (1978). The information of auditory streams. In J. Requin (Ed.) *Attention and performance.* VIII. Hillsdale: N.J.: Erlbaum.

Cohen, A.J. (1976). Perception of tone sequences of the Western European chromatic scale: Tonality, transposition and the pitch set. *Journal of the Acoustical Society of America, abs.* **60**, 1421.

Cohen, A.J., Trehub, S.E. & Thorpe, L.A. (1989). Effects of uncertainty on melodic information processing. *Perception & Psychophysics*, **46**, 18-28.

Cuddy, L.L. & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception and Psychophysics*, **41**, 609-620.

Deutsch, D. & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, **88**, 503-522.

Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, **85**, 341-354.

Dowling, W.J. & Fujitani, D.S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**, 524-531.

Garner, W. (1962). *Uncertainty and structure as phychological concepts.* New York: Wiley.

Garner, W. (1974). *The processing of information and structure.* New York: Wiley.

Helmholtz, H. (1885/1954). *Sensations of tone as the physiological basis of music.* A. Ellis (trans). New York: Dover.

Jones, M.R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, **41**, 621-634.

Jones, M.R., Summerall, L. & Marshburn, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology*, **39A**, 89-121.

Krumhansl, C. (1983). Perceptual structures for tonal music. *Music Perception*, **1**, 28-62.

Martin, J.G. (1972). Rhythmic (hierarchical) versus serial structure in speech and music and other behaviour. *Psychological Review*, **79**, 487-509.

Pitt, M.A. & Monahan, C.B. (1987). The perceived similarity of auditory polyrhythms. *Perception & Psychophysics*, **41**, 534-546.

Roberts, L. & Shaw, M. (1984). Perceived structure of musical triads. *Music Perception*, **2**, 95-124.

Restle, F. (1979). Serial pattern learning. *Journal of Experimental Psychology*, **83**, 120-125.

Simon, H.A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, **79**, 369-382.

Simon, H.A. & Sumner, R.K. (1968). Pattern in music. In. B. Kleinmuntz (Ed.), *Formal representations of human judgment.* New York: Wiley.

Warren, R.M. & Byrnes, D.L. (1975). Temporal discrimination of recycled tone sequences: Pattern matching and naming of order by untrained listeners. *Perception and Psychophysics*, **19**, 273-280.