Recognition of inversions of melodies and melodic contours*

W. J. DOWLING
University of California, Los Angeles, California 90024

Inversions of brief melodies are more difficult to recognize than are transposed repetitions of those melodies. Distinguishing between transposed repetitions and repetitions in which only the melodic contour (the pattern of ups and downs) is repeated is very difficult, as is distinguishing between exact inversions and inversions of the melodic contour.

Many problems in visual pattern recognition find direct analogues in auditory pattern recognition. Two such issues are: What distortions of detail leave the pattern recognizable, and what are the effects of various changes of orientation of shape on recognition (see, e.g., Neisser, 1967)? A convenient medium for exploring these cognitive problems is found in the melodic material of music. The present recognition memory experiment deals with one type of distortion of detail (changes in pitch interval sizes between notes) and one type of change of orientation (melodic inversion, or turning the pattern upside down).

Recognition of melodic material in which the contour (the pattern of ups and downs) is preserved but in which the exact interval sizes are changed is an important listening skill for both western written music (e.g., Bach fugues) and the unwritten folk music of numerous cultures (Nettl, 1956). Inversion of melodic material is also an important formal device in many cultures, especially Western European and Indonesian (Harwood & Dowling, 1970). Figure 1C shows a contour preserving distortion of the pattern in Fig. 1A, and Fig. 1E shows an inversion of that same pattern. Melodic inversion in musical practice often involves inversion of melodic contour without preservation of exact interval sizes. In melodic contour inversions the direction of each interval is reversed while the interval sizes are changed. Figure 1F shows a contour-preserving inversion of the pattern of Fig. 1A. Note that the contour (the pattern of interval directions) of Fig. 1F is the same as that of Fig. 1E, but that the interval sizes between notes are different.

I would claim that the listener's understanding of melodic inversion requires him to store melodic patterns in memory and, later in the piece, recognize new melodic patterns as inversions of the patterns initially stored. The present experiment is a simplification and abstraction of this actual situation. As such, it constitutes a test of whether melodic inversion can function as a formal device understood by the listener or must be viewed as an empty formalism.

*Supported by a faculty summer research grant from UCLA. Computing time provided by the Department of Psychology and the Campus Computing Network, UCLA. I thank E. C. Carterette, D. Norman, and R. Greer for helpful comments.

METHOD

Procedure

On each trial of the experiment, the listener first heard a standard melody presented at a rate of 5 tones/sec. This standard began on middle C (262 Hz), was five notes long, and was different on every trial. The standard was followed after a 2-sec pause by a comparison melody which always began on a different note from the standard. Assuming that Fig. 1A represents the standard, the various possible comparison melodies are shown in Fig. 1G. In each of the six sessions of the experiment, only two types of comparison were used. One of these comparisons was more like the standard than the other in retaining the contour and, in some cases, the exact interval sizes of the standard. Listeners were told to respond "same" to this more similar comparison and "different" to the less similar comparison. Listeners used a four-category confidence-level scale with categories labeled "sure same," "same," "different," and "sure different." There were 60 trials in each session, with 30 of each of the two comparison types. After each trial, listeners had 5 sec to respond. A warning tone (4,250 Hz) preceded the onset of the following trial by 2 sec.

The six sessions were arranged in a 2 by 3 factorial design. The comparison melodies were either repetitions or inversions of the standard. Within each of these two conditions there were three tasks. Task EC tested the listener's ability to distinguish exact interval-size-preserving comparisons (whether repetitions or inversions) (called "same") from merely contour-preserving comparisons (called "different"). Task CD required the listener to distinguish between contour-preserving comparisons (called "same") and completely different randomly selected comparisons (called "different"). Task ED required a distinction between exact interval-preserving comparisons (called "same") and completely different comparisons (called "different"). Each session was preceded by a thorough explanation of the condition and task, with three examples of each trial type.

Subjects

Fourteen UCLA undergraduates served in all six group sessions for course credit. They were well practiced in tasks similar to the present ones. The sessions were held on 3 days, spaced 1 week apart, with two sessions on each day and with a 15-min break between the two sessions on the same day. Previous studies led me to expect the inversion conditions and the EC task to be the most difficult, so these were placed last in the order of sessions so that practice effects, if any, would tend to overcome rather than enhance task differences in the results. Listeners had a mean of 2.43 years of musical training, including lessons on an instrument or voice and playing in an ensemble but not including singing in choirs or music appreciation classes. Median years of musical training was approximately 1.0.

Stimuli

A Hewlett-Packard 2116B computer generated the stimuli, which consisted of sawtooth waves. Stimuli were recorded on tape and played to the listeners over high-quality reproducing equipment. On each trial, the computer generated a new standard stimulus by starting on middle C (262 Hz) and selecting succeeding notes, such that intervals between successive notes occurred with the following probabilities: P(±1 semitone) = .5, P(±2 semitones) = P(±3 semitones) = .25. The comparison started on a different note selected at random from the 14 notes of a chromatic scale, 1-7 semitones higher or lower than middle C. Interval-size-

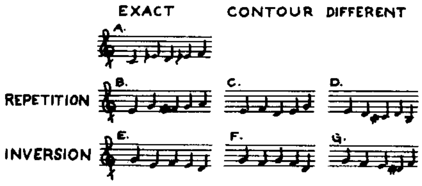

Fig. 1. Exact and contour-preserving repetitions (B and C) and inversions (E and F) of a standard stimulus (A) with randomly selected different comparisons (D and G).

Table 1
Mean Areas Under MOC (N = 14)

	Task		
Condition	EC	CD	ED
Repetition	.58	.85	.89
Inversion	.57	.73	.73

preserving comparisons were exact transpositions or inversions beginning on the new note. Contour-preserving comparisons contained newly selected intervals, different from the corresponding intervals in the standard but with the same contour, inverted or not. Different comparisons were newly selected, just as were the standards, with the stipulation that the contour be different in at least one interval direction from the standard contour or its inversion, depending on condition.

Listeners responded by marking IBM cards which were scored by a computer. Areas under the memory operating characteristic (MOC) were computed for each S in each session (see Norman & Wickelgren, 1965, for details of a similar procedure). The area under the MOC was taken as an estimate of the probability of correct response in a forced-choice procedure with a chance level of .50.

RESULTS AND DISCUSSION

Table 1 shows mean areas under the MOC for the six sessions. The difference between conditions (rows) is significant [$F(1,13) = 19.87$, $p < .01$], with inversions harder to recognize than repetitions. Recognition of inversions is clearly better than chance, however, supporting the notion that inversion is an actually perceivable compositional device. The differences among tasks (columns) are significant ($F = 72.31$, $p < .001$), with the EC task being harder than the other two tasks. The interaction of Condition by Task is significant [$F(2,26) = 5.05$, $p < .02$], with Task EC about equally difficult in the two conditions and Tasks CD and ED more difficult in the inversion condition.

I interpret these results as indicating that the melodic contour and the set of interval sizes in a melody are separable features or dimensions of the melodic pattern. Contour and interval size are handled in different and largely independent ways in cognitive processing. Recognition of exact repetitions and inversions seems to be mainly on the basis of contour. Having the exact interval sizes present (Task ED) does not improve performance appreciably over having just the contour (Task CD) in the repetition condition and does not improve performance at all in the inversion condition. The decrement in performance with inversion would seem to be due to the listener's losing his memory representation of the contour (in whole or in part) in the process of inverting it. Where the listener is sure that the comparison melody will preserve the contour of the standard, either as a repetition or an inversion (Task EC), he must solve the task by recognizing exact interval sizes. Recognition of interval sizes is apparently about equally effective whether or not the intervals have been inverted, which suggests that intervals of the same size can be processed as equivalent, regardless of direction. The fact that contour recognition is adversely affected by inversion and interval recognition is not leads me to argue that melodic contour and interval sizes are being handled differently by the listener.

Table 2 shows correlation coefficients between areas under the MOC and years of musical training in the six sessions. With two exceptions, these correlations lie in the range .55-.59. The two exceptions are Task CD in the repetition condition ($r = .31$) and Task EC in the inversion condition ($r = .24$). In the former case it may be that the musicians' training leads them to regard as different anything except an exact transposition of the standard, and the resulting confusion about the task led to poorer performance. In the latter case the opposite kind of effect may have occurred. Inversions, when they occur in western music, are more typically contour-preserving rather than exact inversions. Musical training would not lead the listener to be better at distinguishing between the two types.

Table 2
Correlation Coefficients Between Areas Under MOC and Years of Musical Training (N = 14)

	Task		
Condition	EC	CD	ED
Repetition	.59	.31	.56
Inversion	.24	.59	.55

These results can be generalized only with caution. In particular, the limited range of interval sizes used in constructing the stimuli makes interval size a less salient dimension (relative to contour) than it would be in actual music. The result obtained, distinguishing as clearly as it does between these attributes of interval size and contour, probably depends for its clarity on the particular range of intervals used.

The present experiment demonstrates an analysis of the features of auditory melodic patterns in a manner analogous to the kind of analysis typically applied to the features of visual patterns. By using a distortion of the melodic pattern which changed interval sizes but left melodic contour intact and a change in orientation which reversed melodic contour, I showed that interval sizes and contour could be made to function independently of each other in the pattern-recognition process.

REFERENCES

HARWOOD, D., & DOWLING, W. J. Musical structure and emotional response in several cultures. Paper presented to the Western Psychological Association, April 1970.

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

NETTL, B. Unifying factors in folk and primitive music. Journal of the American Musicological Society, 1956, 9, 196-201.

NORMAN, D., & WICKELGREN, W. A. Short term recognition memory for digits and pairs of digits. Journal of Experimental Psychology, 1965, 70, 479-489.

(Accepted for publication July 27, 1970.)