

The processing of structured and unstructured tonal sequences

DIANA DEUTSCH

*Center for Human Information Processing
University of California at San Diego, La Jolla, California 92093*

The recall of hierarchically organized tonal sequences was investigated in two experiments. An adaptation of the technique of melodic dictation was employed, in which musically trained listeners notated each sequence after it was presented. Strong effects of sequence structure were obtained. Sequences whose tonal structure could be parsimoniously encoded in hierarchical fashion were recalled with a high level of accuracy. Sequences that could not be parsimoniously encoded produced substantially more errors in recall. Temporal segmentation was found to have a substantial effect on performance, which reflected grouping by temporal proximity regardless of tonal structure. The results provide evidence for the hypothesis that we encode tonal materials by inferring sequence structures and alphabets at different hierarchical levels, together with their rules of combination.

It is generally agreed that we can process serial information of considerable complexity when it is systematically organized and its organization is understood by the observer. Two related aspects of such processing have been emphasized. One is that we tend to group serial patterns into subsequences or "chunks" that are retained in unitary fashion. The other is that we tend to organize such "chunks" as hierarchies when given the opportunity to do so.

Much work on this issue has been performed with the use of verbal materials. It is clear from general experience that verbal items that form meaningful units are processed more easily than those that do not. For example, strings of words that form meaningful sentences are more easily perceived and remembered than the same words ordered in haphazard fashion. Similarly, strings of letters that form meaningful words are processed more readily than meaningless strings (Miller, 1956). Discussions concerning the hierarchical organization of sentence structure are to be found in Chomsky (1963), Chomsky and Miller (1963), Miller and Chomsky (1963), and Yngve (1960). Meaning and grammatical structure as chunking devices have been studied by Bower and Springston (1970), Johnson (1965, 1968), Laughery and Pinkus (1968), and Martin (1967), among others. Second, when sequences of letters and numbers form serial patterns, observers are able to infer the rules underlying these patterns and so encode them parsimoniously (Bjork, 1968; Fritzen & Johnson, 1969;

Kotovsky & Simon, 1973; Simon & Kotovsky, 1963; Vitz & Todd, 1967, 1969).

Other work on the subjective organization of serial patterns has employed nonverbal materials as stimuli. Restle (1970) and Restle and Brown (1970) have shown that we readily acquire serial patterns as hierarchies that reflect pattern structure. In their experiments, subjects were presented with a row of six lights that turned on and off in repetitive sequence, and their task was to predict which light would come on next. The sequences were structured as hierarchies of operators. For example, given the basic subsequence $X = (12)$, then the operation M ("mirror image of X ") produces the sequence 1 2 6 5, the operation R ("repeat of X ") produces 1 2 1 2, and the operation T ("transposition +1 of X ") produces 1 2 2 3. By recursive application of such operations, long sequences can be generated that have compact structural descriptions. For example, the sequence 1 2 1 2 2 3 2 3 6 5 6 5 5 4 5 4 can be described as $M(T(R(T(1))))$. By analyzing the processing of such sequences, it was demonstrated that the observer organizes such serial information in accordance with its structure. Further work has elaborated on these findings (Jones & O'Hara, 1973; Restle, 1973).

The organization of sequences as hierarchically structured subsequences clearly occurs in music. Such organization has long been recognized by music theorists (Meyer, 1956, 1973; Salzer, 1962; Schenker, 1956, 1973) and has also been the subject of speculation among psychologists (Jones, 1974, 1978; Restle, 1970; Restle & Brown, 1970; Simon & Sumner, 1968). The present study investigated recall of tonal sequences that were systematically organized in hierarchical fashion. These were constructed in accordance with

This work was supported by U.S. Public Health Service Grant MH-21001. I am grateful to Diane Williams for statistical analysis of the data. Requests for reprints should be sent to Diana Deutsch, Department of Psychology, C-009, University of California at San Diego, La Jolla, California 92093.

a formulation developed by Deutsch and Feroe (Note 1) which builds on those of other investigators, notably Restle (1970), Simon (1972), and Simon and Sumner (1968), although departing somewhat from them. In this formulation, a melodic sequence is represented as a hierarchy of subsequences, each of which is described with respect to a particular alphabet. For example, Sequence 1 shown in Figure 1 can be represented as the elements of the G major triad (G, , , D, , , B, , , G, ,) each of which is followed by a step down and then up the tone chromatic scale.

Our notation is given by the following rules:

(1) A *structure* is notated by $(A_1, A_2, \dots, A_{i-2}, A_{i-1}, *, A_{i+1}, A_{i+2}, \dots, A_n)$, where A_j is one of the operators n, p, s, n^i , or p^i .

(2) Each structure $(A_1, A_2, \dots, *, \dots, A_n)$ has associated with it an alphabet, α . We call the combination of a structure and an alphabet a *sequence* (or *subsequence*). This, together with the reference element R , produces a *sequence of notes*.

(3) The effect of each operator in a sequence is

determined by that of the operator closest to it but on the same side as the asterisk, $*$. The operator n refers to traversing one step up the alphabet associated with the sequence. Similarly, the operator p refers to traversing one step down this alphabet. The operator s refers to remaining in the same position. The two further operators n^i and p^i refer to traversing up or down i steps along the alphabet, respectively. A string of length k of an operator A is abbreviated kA .

(4) The values of the sequence of notes $(A_1, A_2, \dots, *, \dots, A_n)$, α, R , where α is the alphabet and R the reference element, can be obtained by taking the value of the asterisk to be that of R .

(5) To produce another sequence from the two sequences $A = (A_1, A_2, \dots, *, \dots, A_m), \alpha_1$, and $B = (B_1, B_2, \dots, *, \dots, B_n), \alpha_2$, where α_1 and α_2 are two alphabets, we define a symbol $[pr]$ (prime). $A [pr] B; R$, where R is the reference element, refers to assigning values to the notes produced from $(B_1, B_2, \dots, *, \dots, B_n)$, such that the value of $*$ is the same as the value of A_1 , when the sequence A is applied to the reference element R . Values are then assigned to the notes produced from $(B_1, B_2, \dots, *, \dots, B_n)$, such that the value of $*$ is the same as the value of A_2 , and so on. This gives a sequence of length $m \times n$.

To give an example, Sequence 1 of Figure 1 is notated as:

$$A = (*, 3p) G_{tr}$$

$$B = (*, p, n) Cr$$

$$S = A [pr] B, G^5,$$

where G_{tr} indicates the G major triad, Cr the chromatic scale, and G^5 the reference element. Similarly, Sequence 2 of Figure 1 is notated as:

$$A = (*, 3p) G_{tr}$$

$$B = (2n, *) G_{tr}$$

$$S = A [pr] B, G^4.$$

One of the purposes of the present study was to compare the processing of sequences that can be parsimoniously represented according to the above rules with those that cannot. The sequences in Figure 1A, for instance, can be described in terms of a single higher order sequence of four elements that acts on a single lower order sequence of three elements.



Figure 1. Tonal sequences employed in Experiment 1.

The sequences shown in Figure 1B cannot be represented so parsimoniously. If we assume that the listener retains sequence structures as chunks, then clearly those sequences with parsimonious representations should impose a lighter memory load and so should be better recalled.

A second purpose of the present study was to investigate the role of temporal patterning in the processing of such sequences. As argued by Neisser (1967), such patterning can play a useful role in processing a structured series of elements, by serving to demarcate this structure. Studies employing strings of verbal materials have shown that we tend to recall sequences in accordance with their temporal grouping (Bower & Winzencz, 1969; McLean & Gregg, 1967; Mueller & Schumann, 1894). This effect can be so strong as to mask grouping by meaning, and so to obliterate the advantage incurred by such grouping (Bower & Springston, 1970).

Analogous results have been obtained with nonverbal materials. Restle (1972), using the light-switching task described above, found that inserting pauses between temporal groups either improved or disrupted performance depending on whether the pauses agreed or conflicted with pattern structure. Handel (1973) studied identification of repeating auditory patterns consisting of dichotomous elements differing in pitch. These were segmented by pauses, and it was found that compatible segmentation (e.g., an eight-element pattern segmented by two) resulted in excellent performance, but incompatible segmentation (e.g., an eight-element pattern segmented by three) led to poor performance. Dowling (1973) presented five-note phrases in lists of four phrases separated by pauses. Recognition of a single five-note phrase was superior when its components had been presented in the same temporal segment to recognition when they had not.

Given the above findings, it was predicted that temporal grouping would have a strong effect on processing the present sequences also. Grouping in accordance with sequence structure was expected to result in somewhat enhanced performance, and grouping in conflict with sequence structure to result in performance decrements.

EXPERIMENT 1

Method

Procedure. On each trial, subjects were presented with a sequence of 12 tones, which they recalled in musical notation. All sequences were in the key of G, and each sequence was preceded by G⁵, which served both as a cue and also as an anchor tone. Before the experiment began, the subjects were presented with three practice sequences for notation, each in one of the three temporal configurations to be employed.

Conditions. Four structured sequences were employed in the experiment, and these are shown in Figure 1A. It can be seen that each consisted of a higher level subsequence of four elements that acted on a lower level subsequence of three elements. From each

of these sequences, another sequence was constructed, which consisted of the identical set of tones, but arranged in haphazard fashion. The average interval size between adjacent tones in the unstructured sequences taken overall was nearly identical to the average interval size between adjacent tones in the structured sequences. The four unstructured sequences are shown in Figure 1B.

The eight sequences were all presented in each of three temporal configurations. In the first configuration, the tones were spaced at equal intervals; in the second, they occurred in four groups of three; and in the third, they occurred in three groups of four. The eight sequences were presented three times in succession in different random orders, and the three temporal configurations for each sequence were also ordered at random.

There were, therefore, six conditions in the experiment. In Conditions 0S, 3S, and 4S, the sequences were structured. In Condition 0S, they were presented with no temporal segmentation. In Condition 3S, they were temporally segmented in groups of three so that segmentation was in accordance with sequence structure. In Condition 4S, they were temporally segmented in groups of four, so that segmentation was in conflict with sequence structure. In Conditions 0U, 3U, and 4U, the sequences were unstructured. In Condition 0U, they were presented with no temporal segmentation; in Condition 3U, they were segmented in groups of three; and in Condition 4U, they were segmented in groups of four.

Stimulus parameters. All tones were taken from the equal-tempered scale (International pitch; A = 435 Hz). The frequencies employed (in hertz) were G = 194, B = 244, D = 290, F# = 366, G = 388, A# = 461, B = 488, C = 517, C# = 548, D = 581, E = 652, F# = 732, and G = 775. All tones were 300 msec in duration. For sequences with no temporal segmentation (Conditions 0S and 0U), all tones were followed by 300-msec pauses. For sequences segmented in four groups of three (Conditions 3S and 3U), pauses between tones within a group were 150 msec, and pauses following groups were 600 msec. For sequences segmented in three groups of four, pauses between tones within a group were 100 msec, and pauses following groups were 900 msec. These temporal parameters were chosen so that the sequences produced well-formed rhythmic patterns and so that the duration of each sequence was the same for each temporal configuration. A 1-sec tone of 775 Hz preceded each sequence by 5 sec. All tones were at equal amplitude.

Apparatus. Tones were produced as sine waves by a Wavetek function generator controlled by a PDP 11/03 computer and were recorded on tape. The tape was played to subjects on a high-quality tape recorder through loudspeakers.

Subjects. Twelve students at the University of California, San Diego, served as subjects for the experiment and were paid for their services. All the subjects had had at least 8 years of musical training and claimed to be reasonably adept at melodic dictation. Apart from this, there was no selection procedure. The subjects were naive concerning the purposes of the experiment.

Results

Table 1 shows the percentages of tones correctly recalled in their correct serial positions in the different conditions of the experiment. It can be seen that large effects of both sequence structure and temporal segmentation were obtained. For structured sequences that were segmented in accordance with sequence structure (Condition 3S), the performance level was extremely high. For structured sequences with no temporal segmentation (Condition 0S), the performance level again was very high, although slightly lower. However, for sequences that were segmented in conflict with sequence structure, the performance

Table 1
Percent Correct Recall (PC) of Tones in Correct
Serial Positions in Experiment 1

Condition	PC
Sequences Structured in Groups of Three	
0S. Not temporally segmented	93.5
3S. Temporally segmented in groups of three	99.3
4S. Temporally segmented in groups of four	69.2
Sequences Unstructured	
0U. Not temporally segmented	52.0
3U. Temporally segmented in groups of three	63.2
4U. Temporally segmented in groups of four	62.3

level was considerably reduced. For unstructured sequences (Conditions 3U, 0U, and 4U), performance levels were considerably lower than for structured sequences that were either not segmented or were segmented in accordance with sequence structure.

A three-way analysis of variance was performed, with structure and temporal segmentation as fixed effects and subjects as a random effect. The effect of structure was highly significant [$F(1,11) = 92.71$, $p < .01$], as was the effect of temporal segmentation [$F(2,22) = 11.93$, $p < .01$]. The effect of subjects was also significant [$F(11,216) = 4.89$, $p < .01$]. An issue of major importance was the interaction between structure and temporal segmentation. This was found to be highly significant [$F(2,22) = 25.06$, $p < .01$], reflecting the deleterious effect of incompatible segmentation of the structured sequences. No other interactions were significant.

Figure 2 displays the serial position curves for the different conditions of the experiment. Typical bow-shaped curves are apparent, and, in addition, discontinuities occur at boundaries between temporal groups. This type of configuration, which is very similar to that obtained by Bower and Winzencz (1969) with the use of verbal materials, implies that temporal groups tend to be coded as units or chunks and to be retained or lost independently. A further measure of interitem association is the transition shift probability (TSP), defined as the joint probability of either a correct response following an error on the previous item or of an error following a correct response on the previous item (Bower & Springston, 1970). If groups of elements tend to be retained or lost as chunks, then the TSP values should be smaller for transitions within a chunk and larger for the transition into the first element of a chunk. Figures 3 and 4 display the TSP values for sequences segmented in temporal groups of three (Conditions 3S and 3U) and temporal groups of four (Conditions 4S and 4U), respectively. The TSP after each pause is shown by shading. It can be seen that the TSPs are larger on the first element of each temporal group than on the other elements. This is expected on the assumption that pauses serve to define subjective chunks that tend to be retained or lost independently of each

other, and the results are again very similar to those obtained for verbal materials (Bower & Springston, 1970).

A further point of interest concerns the subjects' sensitivity to the musical alphabets employed in each sequence. As shown in Figure 1, four sequences employed a triadic alphabet exclusively (the three notes G, B, and D), two employed a major diatonic alphabet (the seven notes G, A, B, C, D, E, and F#), and two included other notes from the 12-tone chromatic scale. Of the 12 subjects in the experiment, six stayed entirely within the alphabet of the particular sequence they were notating. Thus, for example, in notating a sequence based on the triadic alphabet, all their incorrect responses were within that alphabet. Five more subjects produced between them a total of 15 responses that deviated from the alphabet of the sequence they were notating. The remaining subject made 26 such errors. Thus, although substantial individual differences were apparent, most of the subjects showed a remarkable sensitivity to alphabet in their responses.

EXPERIMENT 2

In Experiment 1, the structured sequences all consisted of a higher level subsequence of four elements

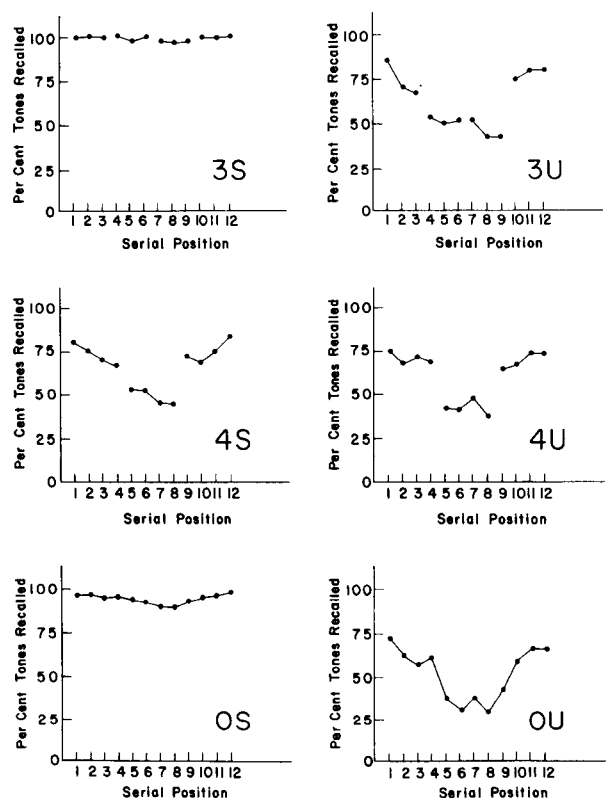


Figure 2. Percentage of tones correctly recalled at each serial position in the different conditions of Experiment 1.

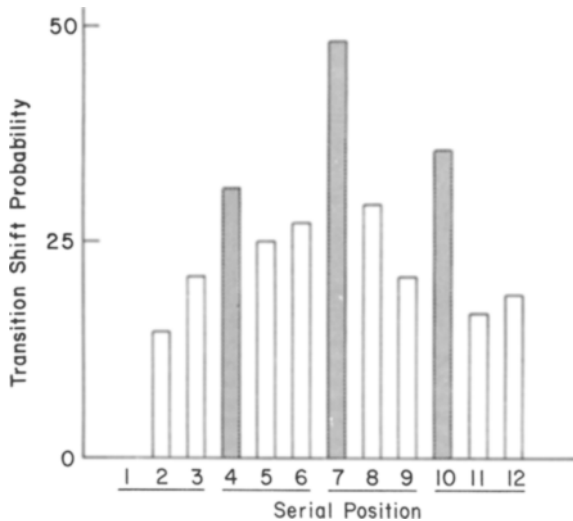


Figure 3. Transition shift probabilities for sequences segmented in temporal groups of three in Experiment 1.

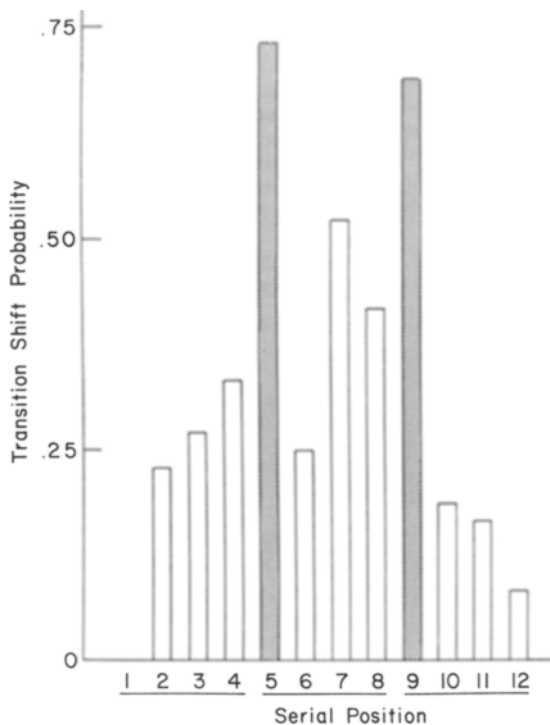


Figure 4. Transition shift probabilities for sequences segmented in temporal groups of four in Experiment 1.

that acted on a lower level subsequence of three elements. Thus, appropriate segmentation was always in groups of three and inappropriate segmentation in groups of four. One might, therefore, argue that the superior performance found for structured sequences was due simply to an advantage conferred by the size of temporal group. Experiment 2 was designed to control for this possibility. Two types

of sequence structure were here employed. In the first, a higher level subsequence of four elements acted on a lower level subsequence of three elements; in the second, a higher level subsequence of three elements acted on a lower level subsequence of four elements. Superior recall was expected for sequences in which the number of elements in the lower level subsequence corresponded to the number within a temporal group, compared with sequences in which these numbers did not correspond. In addition to temporal segmentation in groups of three and four, segmentation in groups of two was examined. It was expected that for sequences in which the lower level subsequence consisted of three elements, such segmentation would result in considerable performance decrements, since it would conflict with sequence structure. However, for sequences in which the lower level subsequence consisted of four elements, such segmentation would be less disruptive, since pauses would still be placed between structural groups.

Method

Procedure. The procedure was identical to that in Experiment 1.

Conditions. Eight sequences were employed in the experiment, and these are shown in Figure 5. Four of these, shown in Figure 5A, consisted of a higher level subsequence of four elements that acted on a lower level subsequence of three elements. The other four, shown in Figure 5B, consisted of a higher level subsequence of three elements that acted on a lower level subsequence of four elements. These eight sequences were each presented in three temporal configurations. In the first, the tones occurred in six groups of two; in the second, they occurred in four groups of three; and in the third, they occurred in three groups of four. The eight sequences were presented three times in succession in different random orders, and the three temporal configurations for each sequence were also ordered at random.

There were, therefore, six conditions in the experiment. In Conditions 3-2, 3-3, and 3-4, the sequences were structured so that the lower level subsequences consisted of three elements, and they were temporally segmented in groups of two, three, and four, respectively. In Conditions 4-2, 4-3, and 4-4, the sequences were structured so that the lower level subsequences consisted of four elements, and they were temporally segmented in groups of two, three, and four, respectively.

Stimulus parameters. All tones were taken from the equal-tempered scale (International pitch; $A = 435$ Hz). The frequencies employed (in hertz) were: $B = 244$, $D = 290$, $F\# = 366$, $G = 388$, $A = 435$, $A\# = 461$, $B = 488$, $C = 517$, $C\# = 548$, $D = 581$, $E = 652$, $F\# = 732$, and $G = 775$. All tones were 300 msec in duration. For sequences segmented in groups of two (Conditions 3-2 and 4-2), pauses between tones within a group were 100 msec and pauses following groups were 500 msec. For sequences segmented in groups of three (Conditions 3-3 and 4-3), pauses between tones within a group were 150 msec and pauses following groups were 600 msec. For sequences segmented in groups of four (Conditions 3-4 and 4-4), pauses between tones within a group were 100 msec and pauses following groups were 900 msec. These temporal parameters were chosen so that all sequences produced well-formed rhythmic patterns, and the duration of each sequence was the same for each temporal configuration. As in Experiment 1, a 1-sec tone of 775 Hz preceded each sequence by 5 sec. All tones were at equal amplitude.

Apparatus. The apparatus was identical to that in Experiment 1.

Subjects. Ten students at the University of California, San Diego, served as subjects in the experiment and were paid for their services. These had not participated in Experiment 1 and were



Figure 5. Tonal sequences employed in Experiment 2.

naive concerning the purposes of the experiment. They had had at least 8 years of musical training and claimed to be reasonably adept at melodic dictation. Apart from this, there was no selection procedure.

Results

Table 2 shows the percentages of tones correctly recalled in their correct serial positions in the different conditions of the experiment. It can be seen that for both types of sequence structure, the performance level was very high when the sequences

Table 2
Percent Correct Recall (PC) of Tones in Correct
Serial Positions in Experiment 2

Condition	PC
Sequences Structured in Groups of Three	
3-2. Temporally segmented in groups of two	45.4
3-3. Temporally segmented in groups of three	93.1
3-4. Temporally segmented in groups of four	50.6
Sequences Structured in Groups of Four	
4-2. Temporally segmented in groups of two	80.8
4-3. Temporally segmented in groups of three	52.9
4-4. Temporally segmented in groups of four	85.4

were segmented in accordance with structure (Conditions 3-3 and 4-4). It can also be seen that the performance level was considerably lower when temporal segmentation was in conflict with sequence structure (Conditions 3-2, 3-4, and 4-3). When pauses were placed both between and within structural groups (Condition 4-2), the performance level was slightly lower than when the pauses were placed only between groups, but was considerably higher than when the pauses conflicted with sequence structure.

A three-way analysis of variance was performed, with size of structural unit and size of temporal unit as fixed effects and subjects as a random effect. The effect of size of structural unit was significant [$F(1,9) = 8.162$, $p < .025$]. The effect of size of temporal unit was not significant [$F(2,18) = 3.176$, $p > .05$]. The effect of subjects was significant [$F(9,180) = 9.124$, $p < .01$]. The issue of importance in the experiment was the interaction between size of structural and size of temporal unit. This was found to be very highly significant [$F(2,18) = 67.131$, $p < .01$], reflecting the deleterious effect of incompatible segmentation. No other interactions were significant.

Figure 6 displays the serial position curves for the different conditions of the experiment. It can be seen that, as in Experiment 1, discontinuities appear at temporal group boundaries, reflecting the formation of subjective chunks on the basis of temporal proximity. As a further investigation of the basis for subjective chunking, the following analysis was performed. For Conditions 3-4 and 4-3, the proportion of chunks that were notated without error was calculated both for assumed chunks of size 3 and also for assumed chunks of size 4. (The same data from both conditions were therefore used for these two calculations.) These proportions are shown in Table 3. It can be seen that more chunks of size 3 were correctly recalled in their entirety when the sequences were presented in temporal groups of three and were structured in groups of four (Condition 4-3) than when they were presented in temporal groups of four and structured in groups of three (Condition 3-4). Similarly, more chunks of size 4 were correctly recalled in their entirety when the sequences were presented in temporal groups of four but structured in groups of three (Condition 3-4) than when they were presented in temporal groups of three and structured in groups of four (Condition 4-3).

A three-way analysis of variance was performed, with type of structure and assumed chunk size as fixed effects, and subjects as a random effect. The effect of subjects was significant [$F(9,120) = 4.999$, $p < .01$]. The effect of type of structure was not significant ($F < 1$). The effect of assumed chunk size was significant [$F(1,9) = 9.113$, $p < .01$], presumably reflecting the greater probability of recalling three, rather than four, items in a row correctly. The comparison of importance here was the interaction between type of

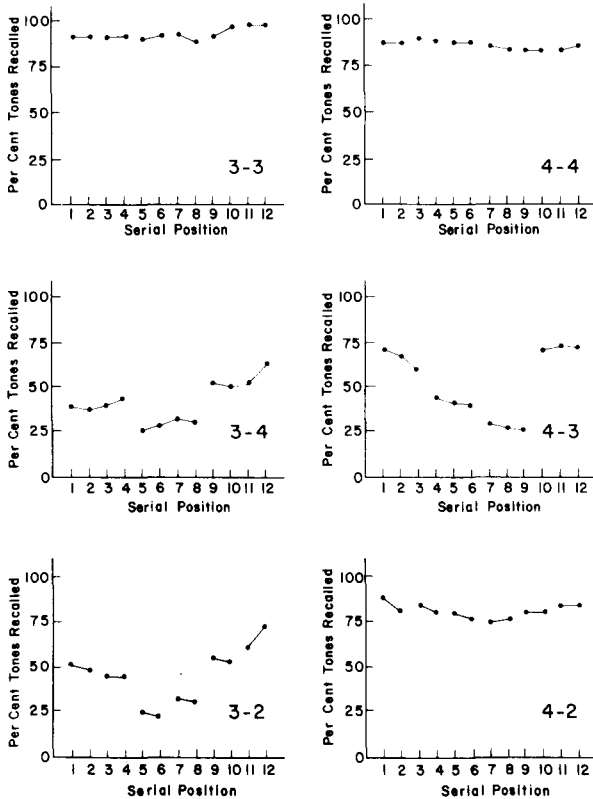


Figure 6. Percentage of tones correctly recalled at each serial position in the different conditions of Experiment 2.

Table 3
Proportion of Chunks of Assumed Sizes 3 and 4 Recalled Without Error in Conditions 3-4 and 4-3 of Experiment 2

Assumed Chunk Size	Condition	
	3-4	4-3
3	36.3	41.9
4	44.0	26.7

Note—In Condition 3-4, structural groups consisted of three elements and temporal groups of four elements. In Condition 4-3, structural groups consisted of four elements and temporal groups of three elements.

structure and assumed chunk size. This was highly significant [$F(1,9) = 33.574, p < .01$], reflecting the formation of subjective chunks on the basis of temporal proximity.

Finally, the issue of sensitivity to alphabet was again examined. As can be seen from Figure 5, three of the sequences employed a triadic alphabet exclusively, three employed a diatonic alphabet, and two included tones from the chromatic scale. Of the 10 subjects in the experiment, five stayed entirely within the alphabets of the particular sequence they were notating. Two subjects produced between them a total of three responses that deviated from these alphabets, one subject made 10 such errors, and the last two made 27 and 28, respectively. Thus, as in Experiment 1, a considerable sensitivity to alphabet

was demonstrated, although marked individual differences were obtained.

DISCUSSION

The two experiments reported here lead to several conclusions. First, they demonstrate that listeners perceive hierarchical structures that are present in tonal sequences and can utilize such structures in recall. For the structured sequences employed in the study, the listener need only retain two chunks of three or four items each (together with their alphabets, the reference element, and a single rule of combination). However, for the unstructured sequences, no such parsimonious encoding was possible. The unstructured sequences, therefore, imposed a much heavier memory load, with resultant performance decrements.

Second, the experiments demonstrate that temporal segmentation has a profound effect on perceived structure, as has been noted by others with the use of different stimulus materials (Bower & Springston, 1970; Dowling, 1973; Handel, 1973; Restle, 1972). Temporal segmentation in accordance with sequence structure resulted in somewhat enhanced performance, but temporal segmentation in conflict with sequence structure led to severe performance decrements. The shapes of the serial position curves, the TSP profiles, and the proportions of temporal vs. structural chunks recalled in their entirety also demonstrated grouping in accordance with temporal proximity rather than sequence structure. According to the present formulation, when such grouping is in conflict with sequence structure, there results a less parsimonious representation. For example, Sequence 3 of Experiment 1 would be encoded in the absence of temporal segmentation as:

$$A = (*, 3n) G$$

$$B = (2p, *) G$$

$$S = A[pr] B, D^5,$$

where G indicates the G major scale and D^5 the reference element.

However, with temporal segmentation in groups of four, this sequence would be represented as:

$$A = (*, 2n^2) G$$

$$B = (2p, n, *) G$$

$$C = (p, n, p, *) G$$

$$D = (n, 2p, *) G$$

$$S = A [pr] (B, C, D) ; C^5.$$

Thus, four chunks would need to be encoded and retained, together with their rules of combination and alphabets. Other structured sequences, especially those involving the chromatic scale, would require an even more elaborate representation when temporally segmented in conflict with sequence structure.

A third point to be noted from these experiments is that listeners appear to retain information concerning alphabet independently of structure, as suggested by the present formulation. There was a very strong tendency for incorrectly notated sequences to remain within the alphabets presented.

The very high level of performance for structured sequences in this experiment stands in sharp contrast to the poor performance obtained when subjects make pitch-recognition judgments involving single tones that are separated by a sequence of interpolated tones chosen at random from the 12-tone chromatic scale (Deutsch, 1970, 1975). Differences in the subject populations employed in the two sets of experiments may well have been responsible in part, but this was probably not the major factor. It would appear that the superior performance levels obtained here were due largely to the projection of tonal information onto highly overlearned alphabets, together with the opportunity for parsimonious encoding. These results, therefore, do not primarily reflect memory for pitch (or even interval, since the lower level subsequences often involved different intervals when they were realized with respect to different elements of the higher level subsequences), but rather memory for a set of abstractions.

REFERENCE NOTE

1. Deutsch, D., & Feroe, J. *A hierarchical model for the generation of tonal sequences*. Manuscript in preparation.

REFERENCES

- BJORK, R. A. All-or-none subprocesses in the learning of complex sequences. *Journal of Mathematical Psychology*, 1968, **5**, 182-195.
- BOWER, G. H., & SPRINGSTON, F. Pauses as recoding points in letter series. *Journal of Experimental Psychology*, 1970, **83**, 421-430.
- BOWER, G. H., & WINZENZ, D. Group structure, coding, and memory for digit series. *Journal of Experimental Psychology Monographs*, 1969, **80**, 1-17.
- CHOMSKY, N. Formal properties of grammars. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2). New York: Wiley, 1963.
- CHOMSKY, N., & MILLER, G. A. Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2). New York: Wiley, 1963.
- DEUTSCH, D. Tones and numbers: Specificity of interference in short-term memory. *Science*, 1970, **168**, 1604-1605.
- DEUTSCH, D. The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short term memory*. New York: Academic Press, 1975.
- DOWLING, W. J. Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, 1973, **4**, 37-40.
- FRITZEN, J., & JOHNSON, N. F. Definiteness of pattern ending and uniformity of pattern size: Their effect upon learning number sequences. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 575-580.
- HANDEL, S. Temporal segmentation of repeating auditory patterns. *Journal of Experimental Psychology*, 1973, **101**, 46-54.
- JOHNSON, N. F. The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 469-475.
- JOHNSON, N. F. The influence of grammatical units on learning. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 236-240.
- JONES, M. R. Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N.J.: Erlbaum, 1974.
- JONES, M. R. Auditory patterns: Studies in the perception of structure. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 8). New York: Academic Press, 1978.
- JONES, M. R., & O'HARA, J. W. Memory interference as a function of rule-governed expectancies. *American Journal of Psychology*, 1973, **86**, 523-536.
- KOTOVSKY, K., & SIMON, H. A. Empirical tests of a theory of human acquisition of concepts for sequential events. *Cognitive Psychology*, 1973, **4**, 399-424.
- LAUGHERY, K. R., & PINKUS, A. L. Recording presentation rate in short-term memory. *Journal of Experimental Psychology*, 1968, **76**, 636-641.
- MARTIN, J. G. Hesitations in the speaker's production and listener's reproduction of utterances. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 903-909.
- MCLEAN, R. S., & GREGG, L. W. Effects of induced chunking on temporal aspects of serial retention. *Journal of Experimental Psychology*, 1967, **74**, 455-459.
- MEYER, L. B. *Emotion and meaning in music*. Chicago: University of Chicago Press, 1956.
- MEYER, L. B. *Explaining music: Essays and explorations*. Berkeley: University of California Press, 1973.
- MILLER, G. A. The magical number seven, plus or minus two: Some limits in our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.
- MILLER, G. A., & CHOMSKY, N. Finitary models of language users. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2). New York: Wiley, 1963.
- MUELLER, G. E., & SCHUMANN, F. Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 1894, **6**, 81-190, 257-339.
- NEISSER, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- RESTLE, F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 1970, **77**, 481-495.
- RESTLE, F. Serial patterns: The role of phrasing. *Journal of Experimental Psychology*, 1972, **92**, 385-390.
- RESTLE, F. Serial pattern learning: Higher-order transitions. *Journal of Experimental Psychology*, 1973, **99**, 61-69.
- RESTLE, F., & BROWN, E. Organization of serial pattern learning. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 4). New York: Academic Press, 1970.
- SALZER, F. *Structural hearing*. New York: Dover, 1962.
- SCHENKER, H. *Neue Musikalische Theorien und Phantasien: Der Freie Satz*. Vienna, Austria: Universal Edition, 1956.
- SCHENKER, H. *Harmony*. O. Jones (Ed. and annotator; E. M. Borgese, trans.). Cambridge, Mass: M.I.T. Press, 1973.

- SIMON, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, **79**, 369-382.
- SIMON, H. A., & KOTOVSKY, K. Human acquisition of concepts for sequential patterns. *Psychological Review*, 1963, **70**, 534-546.
- SIMON, H. A., & SUMNER, R. K. Pattern in music. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.
- VITZ, P. C., & TODD, T. C. A model of learning for simple repeating binary patterns. *Journal of Experimental Psychology*, 1967, **75**, 108-117.
- VITZ, P. C., & TODD, T. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, 1969, **76**, 433-449.
- YNGVE, V. H. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, **104**, 444-466.

(Received for publication February 14, 1980;
revision accepted July 24, 1980.)