Perception & Psychophysics
1975, Vol. 18 (4), 273-280

Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners

RICHARD M. WARREN and DENNIS L. BYRNES
University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

Sets of recycled sequences of four successive tones were presented in all six possible orders to untrained listeners. For pitches within the musical range, recognition (as measured by matching of any unknown order with an array of permuted orders of the same tones) could be accomplished as readily for tonal durations and frequency separations outside the limits employed for melodic construction as inside these limits. Identifying or naming of relative pitches of successive tones was considerably more difficult than matching for these tonal sequences, and appeared to follow different rules based upon duration and upon frequency separation. Use of frequencies above the pitch limits for music (4,500 Hz and above) resulted in poor performance both for matching and naming of order. Introduction of short silent intervals between items was without effect for both tasks. Naming of order and pattern recognition appear to reflect different basic processes, in agreement with earlier formulations based on experiments with phonemic sequences of speech and sequences of unrelated sounds (hisses, tones, buzzes). Special characteristics of tonal sequences are discussed, and some speculations concerning music are offered.

Recent work has indicated that there are two distinct types of temporal order perception both for speech and for sequences of arbitrarily selected unrelated sounds such as hisses, tones, and buzzes (see Warren, 1974a, Note 1). The present study is designed to test whether such separate perceptual categories exist for sequences of tones as well.

Type I discrimination involves direct naming of the order of components. With extended sequences (such as those produced by recycling a sequence of three or four successive sounds over and over without pause), the lower limit for such direct naming is generally about 150 to 600 msec/item, depending upon the sounds and the experimental procedure (Warren & Obusek, 1972). The rate-limiting stage for extended sequences seems to be the time required for verbal identification or naming of the on-going sound, which must be completed before the onset of the next item. Type II discrimination involves holistic recognition of the auditory patterns, and can operate for item durations ranging from a few milliseconds to at least a few hundred milliseconds (the upper portion of the range overlapping with Type I discrimination). Type II judgments permit matching of sequences (as in same-different judgments), and also recognition of a particular pattern heard earlier when presented alone. If item durations are too short for Type I judgments, familiarization with making Type II judgments involving a particular sequence will not by itself lead to the ability to name the order of components. However, with appropriate training, Type II discrimination can allow "indirect" naming of the order of components down to 5 or 10 msec/item through recognition of an overall pattern followed by recall of a learned verbal label (which could be a list of component items in proper order). Such learning can occur inadvertently during the course of a study employing more than one condition for stimulus presentation (see Warren, 1974a, b), and it was in part for that reason that separate groups of untrained subjects were used for each experimental condition in the present investigation.

If Type I and Type II categories each apply for tonal sequences, we would expect to find characteristic differences in accuracy of performance for direct naming of order and for matching (identifying identical arrangements within an array of permuted orders). Further, conditions which enhanced accuracy of response for one type of judgment should not necessarily produce comparable effects with the other type. The present study investigated both direct naming and matching of temporal order using sets of four successive tones of different frequencies repeated over and over in a fixed order. Most experimental groups heard tones lasting 200 msec, a duration within the range of 150 to 900 msec cited by Fraisse (1963, p. 89) for durations of notes forming melodic themes in music. Two

The research was supported by grants both from the National Science Foundation (GB-26459 and GB-36986X) and the University of Wisconsin-Milwaukee Graduate School to the first author, and by a University of Wisconsin-Milwaukee postdoctoral fellowship to the second author. We wish to thank Leong F. Tai for his help in programming the stimulus sequences. Requests for reprints should be sent to Richard M. Warren, Department of Psychology, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201. D. L. Byrnes is now at Psychology Department 1, University of Massachusetts-Boston, Boston, Massachusetts 02116.

hundred milliseconds is quite sufficient for recognition of individual sounds, but is generally just below the threshold for direct naming of order, at least for extended sequences of unrelated sounds heard under conditions comparable to those of the present experiment (Warren & Obusek, 1972). Experiment I examined the accuracy of direct naming of order and of matching permuted orders of tones for a number of frequency separations lying within the pitch range employed in music [according to Wood (1962, p. 54), the range of fundamental frequencies useful in music is from 40 to 4,000 Hz]. There is evidence that judgments of relative pitch are much more difficult for frequencies above the musical limits (Bachem, 1948), and Experiment II explored the accuracy of direct naming and matching with tonal frequencies of 4,500 Hz and above. It has been reported that short silent intervals separating component items facilitates naming of order with sequences of nonrelated sounds (Warren, 1972), and Experiment III explored the effect of such silent intervals on both direct naming and matching for one of the frequency sets used in Experiment I. Sequences of nonrelated sounds can be matched at component durations too brief to permit identification of order (Warren, Note 1), and Experiment IV investigated matching at brief (50 msec) durations (using the frequency set employed in both Experiments I and III).

GENERAL METHOD

Stimuli

The four tones in each sequence were produced by separate beat-frequency oscillators adjusted to within 0.05% of the specified frequency and stable within 0.10%. Sequencing and timing of items was accomplished through Grason-Stadler Series 1200 programming equipment with rise decay of the electronic switches set for 5 msec. The sequential orders of tones were verified visually with oscilloscope tracings. Tonal frequencies and item durations were calibrated using a Hewlett-Packard Model 5302A quartz crystal time base.

The four successive tones in each stimulus sequence were repeated over and over in the same order, with no extra separation between restatements of the four items, until after the subject had completed his response. The sequence then continued to recycle (although inaudible to the subject) during the intertrial interval. Both this interval and the point in the sequence at which a particular trial was terminated were allowed to vary freely across trials. As a consequence, the item occurring at the beginning of a trial can be considered as varying randomly.

Five sets of four tones each were used. Every set of tones was presented in each of the six possible arrangements of the four items, providing six recycled sequences differing only in order. The frequencies used were: Set I (0.3 semitones)—1,000, 1,017, 1,035, 1,053 Hz; Set II (1 semitone)—1,000, 1,059, 1,122, 1,189 Hz; Set III (3 semitones)—1,000, 1,189, 1,414, 1,682 Hz; Set IV (3 semitones)—4,500, 5,350, 6,363, 7,569 Hz; Set V (9 semitones)—500, 841, 1,414, 2,378 Hz.

All tones were presented at a level of 80 ± 0.2 dB C re 20 $\mu N/m^2$ as measured by a Brüel and Kjaer Model 2204 sound-level meter with a 6-cc earphone coupler. Sound pressure level was adjusted with a McIntosh amplifier, Model MA-5100, and the sequences were delivered through matched TDH 49 headphones.

Subjects

Volunteers were recruited from introductory psychology courses and received $2 and/or course credit for their participation. None had any known hearing disability, and there was no selection made on the basis of special musical training or skills. Each of the 234 subjects served in only one experiment and heard only a single group of six stimuli differing only in the order of their tonal components.

Procedure for Matching

Subjects were tested individually in an IAC Model 1204A audiometric room. Before their matching session, they received the following information and instructions: When they placed the control switch in the position marked "Standard," they would hear four tones recycled in a particular order selected by the experimenter. When they switched to the "Comparison" position, they could choose the sequence heard by pressing one of the six buttons arranged in a row and labeled A through F. Each of the six comparison sequences had a different order of the same four items, and one of these orders was the same as the standard's. Their task was to call out and inform the experimenter which comparison was identical to the standard sequence, but they did not have to concern themselves with the order of the component tones. They could switch back and forth from the standard to the comparison series as often as they wished, without any time limit for response.

Every subject received six separate trials, each corresponding to one of the sequential arrangements of the four items. At the beginning of each trial, the experimenter depressed a button on his standard control panel which determined the fixed standard sequence heard by the subject. After the subject called out the letter corresponding to his or her response, the experimenter terminated the trial, and then presented a different permuted order for matching. The order in which the six standards was presented was determined by a Latin square, such that each permutation occurred an equal number of times as first judgment, second judgment, etc., within each subgroup of 18 subjects.

Procedure for Direct Naming of Order

The procedure was similar in general respects to that used for matching, except that subjects did not listen to comparison stimuli, and heard only the particular standard sequence presented by the experimenter without exercising any control over the sequences presented. During the single session devoted to naming, each listener received one presentation of each of the six possible arrangements of the four recycled items, the order of presentation of the sequences being determined by a separate Latin square for the naming judgments. Before the session started, they received the following information and instructions: They would hear a train of four different tones presented in a certain order and repeated over and over. Each of the four tones would have a separate pitch, and would be described in relation to the pitches of the other tones by one of the four cards lying on the table. They could listen for as long as needed to recognize the four tones and their relative pitches, and should arrange the four cards so that they would describe the order of occurrence of the tonal pitches within the sequence. When they finished their arrangement, they were to call out, and the experimenter would note the position of the cards. The four cards used for responding described the relative pitches of the components and were labeled: Highest, Second Highest, Second Lowest, Lowest. Numbers in the corner of each card also corresponded to the pitch relations, with the numeral 1 assigned to "Lowest" and 4 to "Highest." As discussed earlier (Warren & Obusek, 1972), card-ordering is much more accurate than oral responding, probably due to the ease of using a strategy consisting of successive partial responses by placing one card in position at a time.

EXPERIMENT I MATCHING AND NAMING OF ORDER: EFFECT OF FREQUENCY SEPARATION OF COMPONENTS

Procedure

Four separate groups of 36 subjects were used. Each group heard only one set of frequencies for both matching and direct naming. The four sets employed were (the frequencies within each set have been listed above): Set I (0.3 semitone); Set II (1 semitone); Set III (3 semitones); Set V (9 semitones). All stimulus frequencies were within the range of fundamental frequencies produced by orchestral instruments. The durations of the tonal components of the sequences were always 200 msec. Half the subjects of each group performed the matching task in the first session and the naming of order at a second session a week later. The other 18 subjects reversed the order of the tasks.

Results

The accuracy of responses is summarized as percent correct responses in Table 1 for matching and Table 2 for direct naming. Figure 1 summarizes the mean values from Tables 1 and 2 in graphical form. The tonal patterns of the stimuli are indicated in the tables by the arrangement of numbers 1 through 4: Number 1 represents the lowest frequency, and Number 4 the highest in the sequence (thus 4321 represents a recycled sequence in which tones are presented in order of decreasing frequency, returning to the highest tone immediately after the lowest).

As might be anticipated, the regular ascending order (1234) and descending order (4321) generally are matched more accurately (Table 1) and named more accurately (Table 2) than the other tonal permutations. Matching was found to be considerably more accurate than direct naming of order (see Figure 1).

It is of interest that the average matching accuracy improved regularly with increasing frequency separation from 0.3 through 9 semitones (Figure 1). For direct naming, on the other hand, while accuracy was poorest for 0.3-semitone separation, it remained approximately constant from 1 through 9 semitones.

The data from the four groups of subjects participating in Experiment I were submitted to an analysis of variance for a repeated measures design (Tasks by Groups by Order of Tasks). The total number of the subjects' correct responses for a task (matching or naming) was assigned as the score for that task (values could be any integer from 0 through 6). The range of tones used and the order in which the subjects performed the tasks were between-subject variables, while matching vs. naming was a within-subject variable.

Table 1
Matching of Temporal Patterns Within Repeated Tonal Sequences

Experiment No.	Item Duration (msec)	Frequency Separation (Semitones)	Tonal Pattern (Tones Rank-Ordered 1-4 by Frequency): Glissandi 4321	Glissandi 1234	Glissandi Mean	Irregular Progressions 1243	Irregular Progressions 1423	Irregular Progressions 1324	Irregular Progressions 1342	Irregular Progressions Mean	Overall Mean
I	200	.3*	64	58	61	50	72	58	47	57	58
I	200	1*	83	75	79	61	72	50	50	58	65
I	200	3*	75	78	76.5	72	67	53	61	63	68
I	200	9*	86	86	86	72	61	61	75	67	73.5
II	200	3**	33	36	34.5	36	36	33	28	33	34
III	150†	3*	89	67	78	72	78	39	67	64	69
IV	50	3*	72	78	75	61	72	56	61	62.5	67

Note. Mean scores from separate groups of subjects are expressed as percentage correct matches (17% correct corresponds to the chance score).
**Frequencies within musical range* ***Frequencies above musical range* *†50-msec silence between tones*

Table 2
Direct Naming of Order of Components Within Repeated Tonal Sequences

Experiment No.	Item Duration (msec)	Frequency Separation (Semitones)	Tonal Pattern (Tones Rank-Ordered 1-4 by Frequency): Glissandi 4321	Glissandi 1234	Glissandi Mean	Irregular Progressions 1243	Irregular Progressions 1423	Irregular Progressions 1324	Irregular Progressions 1342	Irregular Progressions Mean	Overall Mean
I	200	.3*	68	33	50.5	11	28	25	17	20	30
I	200	1*	64	47	55.5	33	33	17	42	31	39
I	200	3*	64	61	62.5	22	22	17	44	26	38
I	200	9*	67	61	64	30	31	22	28	28	40
II	200	3**	56	36	46	17	17	14	19	33.5	26.5
III	150†	3*	56	61	58.5	11	33	33	11	22	34

Note. Mean scores from separate groups of subjects are expressed as percentage correct identification (17% correct corresponds to the chance score).
**Frequencies within musical range* ***Frequencies above musical range* *†50-msec silence between tones*

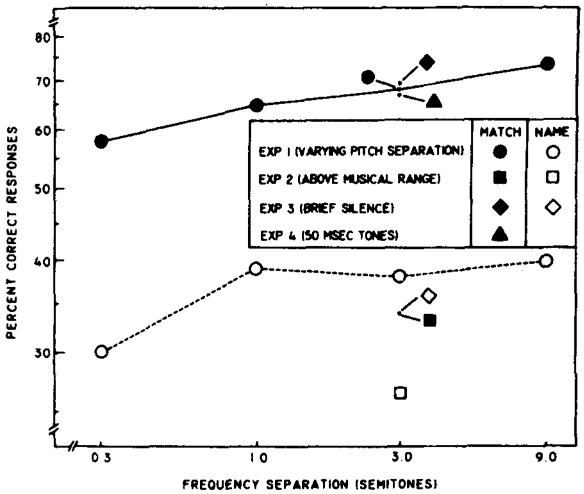

Figure 1. Accuracy of matching and of identification (naming) of order with recycled sequences containing four tones.

As indicated by the overall analysis of variance ($F = 2.80$, $df = 3/136$, $p < .05$) and by subsequent Duncan's multiple range comparisons ($df = 136$, $p < .05$), matching performance improved significantly across the extremes of frequency separation. There was no overlap between the means of the scores for matching and naming ($F = 132.80$, $df = 1/136$, $p < .001$). There was no main effect of task order, and none of the interaction terms in the analysis reached the criterion of significance.

A Friedman two-way analysis by ranks showed that for naming of order, as indicated by examination of Table 2, different tonal arrangements were not of equal difficulty ($\chi_r^2 = 20.66$, $df = 5$, $p < .001$). There was good agreement across groups as to the order of difficulty. Thus, all four frequency separations in Experiment I had the greatest number of correct naming responses to the regularly descending series. Three of the four groups had the fewest correct responses to sequence 1324. For the matching task, agreement in relative difficulty of orders across groups (see Table 1) was somewhat less pronounced than for naming, but here again the difference in difficulty across sequences is greater than would be expected by chance ($\chi_r^2 = 12.71$, $df = 5$, $p < .05$). The descending series and the ascending series were the easiest to match for all but the smallest frequency separation.

EXPERIMENT II
MATCHING AND DIRECT NAMING OF ORDER: EFFECT OF EMPLOYING FREQUENCIES OUTSIDE THE MUSICAL RANGE

Procedure

One group of 36 subjects was used. They heard Stimulus Set IV, having a frequency separation of three semitones, as in Set III. But, unlike Set III in Experiment I, the frequencies in Set IV were above the limits of the musical pitch range. The duration of the component tones was 200 msec. Half the subjects performed the matching task in the first session and the direct naming task at a second session a week later. The other half of the subjects reversed the order of the tasks.

Results

The performance of subjects in this experiment is summarized in Tables 1 and 2 and shown in Figure 1. It can be seen that the accuracy of subjects was very poor indeed, both for matching and naming of order. This group, hearing pitches above the musical range, differed reliably from performance of the group receiving the same frequency separation in semitones for pitches within the musical range in Experiment I ($F = 26.2$, $df = 1/68$, $p < .001$). In addition, there was a significant interaction of task and frequency ($F = 7.56$, $df = 1/68$, $p < .01$) due to the superiority ($t = 5.21$, $df = 68$, $p < .001$) of matching over naming only for pitches within the musical range. No other interaction terms in the analysis reached the criterion of significance.

EXPERIMENT III
MATCHING AND DIRECT NAMING OF ORDER: EFFECT OF BRIEF SILENCE BETWEEN COMPONENTS

Procedure

Two separate groups of 18 subjects were used. One group served in a session devoted to direct naming of order, the other in a matching session. Both groups heard frequencies corresponding to Set III (3-semitone separation), as in Experiment I, but in this experiment 50 msec of silence separated each tone which was reduced in duration from 200 msec to 150 msec.

Results

The results obtained in this experiment are presented in Tables 1 and 2 and Figure 1. Inspection shows that performance for both matching and direct naming were almost the same as for the corresponding groups in Experiment I hearing the same frequencies without the 50-msec silence between items. Analysis of variance was performed comparing Experiment III with the 18 first-session scores for matching and for naming with Set III in Experiment I. As in Experiment I, matching was much more accurate than direct naming ($F = 34.75$, $df = 1/68$, $p < .001$). The pause between components, however, did not affect the accuracy of performance ($F < 1$, $df = 1/68$). There was no significant interaction between the type of task (matching vs. naming) and the presence or absence of the silent intervals ($F < 1$, $df = 1/68$).

EXPERIMENT IV
MATCHING: EFFECT OF SHORTENING ITEMS BELOW DURATIONS PERMITTING DIRECT NAMING

Procedure

One group of 18 subjects participated in a single matching session, using the standard experimental testing procedure. The stimuli had the same frequencies as in Experiment III and in the 3-semitone group in Experiment I. However, the duration of items was reduced to 50 msec so that there were 20 items/sec, a rate considerably above that permitting direct naming of order, and also above that used for melody in music.

Results

The great decrease in item duration did not produce an appreciable change in matching accuracy ($t = .71$, $df = 34$, $p < .20$) as can be seen in Table 1 and Figure 1. This independence from effects of item durations is in marked contrast with performance involving the direct naming of order (see Warren & Obusek, 1972).

DISCUSSION

It has been assumed quite generally that the capacity to detect (and name) the order of the successive components of a sequence is a fundamental ability which permits recognition of sequences and discrimination between permuted orders of the same components. Accepting this common-sense view, studies dealing with auditory temporal resolution usually have measured accuracy of naming the order of components. Most studies have reported values for auditory acuity (thresholds for naming of order) of about 15 to 60 msec, depending upon procedure and practice (for review, see Fay, 1966). It was with considerable surprise that we observed that the order of recycled sequences containing three or four arbitrarily selected sounds (hisses, tones, buzzes) could not be named at 200 msec/item, even through each of the sounds could be heard clearly (Warren, 1968; Warren, Obusek, Farmer, & Warren, 1969). Simplifying strategies did not help under these conditions, so that it was not possible to choose one of the sounds and detect which of the other sounds either preceded or followed it. Follow-up work indicated that the tacit assumption that the earlier studies tested the ability to perceive directly the order of components was not valid. It was suggested that fine temporal acuity (which was called Type II discrimination) involves recognition of overall temporal patterns without any need for the capacity to recognize the order of items within these patterns (Warren, 1972, 1974a, b). Indeed, it appears that Type II discrimination can occur even without the ability to recognize the nature or number of component sounds within a sequence. Much of what has been accepted as threshold measurement for naming of order at durations below 200 msec appears to involve two stages: (a) recognition of discrete auditory patterns (Type II discrimination), followed by (b) recall of a learned verbal label consisting of the names of components in the proper order. Direct naming of the order of items within an extended sequence (Type I discrimination) appears to be limited by the time required to attach a verbal label to on-going stimuli, so that the naming of each of these items must be completed before the onset of the next. The last item (and to a lesser extent, the initial item) of a sequence can be named with special ease. Also, short sequences consisting of single presentations of up to four items can be held briefly in short-term auditory storage until after the sequence ends, permitting direct naming of order at item durations below the limit for extended sequences (Warren, 1972).

This model for auditory temporal discrimination was developed mainly from work with sequences of unrelated sounds. However, there is evidence from work with phonemic restorations (Warren, 1970; Warren & Obusek, 1971) and identification time for targets of different phonetic complexity (Savin & Bever, 1970; Warren, 1971) that the fine temporal discrimination characteristic of speech perception is based upon Type II discrimination, and that both phonemic identification and the naming of the order of individual phonemes require prior identification of clusters or syllables.

The present experiment tests whether the model developed from observations employing arbitrary sequences of sounds and speech also applies to tonal sequences.

Recognition of tonal sequences is, of course, related to basic aspects of music. The selection of tonal ranges and frequency separations in this study were influenced by musical considerations. Thus, the frequency separations used for the recycled sequences of tones were not only within the range used in melodic themes, but above (27 semitones spanned in three steps) and below (0.9 semitones spanned in three steps).

Some of the questions posed by the experimental design were: Is matching and naming of temporal order better for tonal sequences consisting of pitch changes within the limits employed in melodic sequences? Can matching and naming be accomplished for pitches above the musical range? Can recognition of permuted orders as measured by matching be achieved at durations below the limits for naming of order and shorter than those employed for construction of melodies in music? Can short silent intervals between items enhance matching and naming as found with other types of sequences? Is accuracy for the different types of responses measured in this study (matching and naming) influenced in similar fashions by changes in experimental conditions? In answering these questions, it is desirable to consider the two types of response separately.

Matching of Order

In our present experiments, the accuracy of matching tonal patterns was found to be quite good for Frequency Sets I, II, III, and V (stimuli from 500 to 2,378 Hz), having pitches within the limits of the musical range. Table 1 shows that glissandi (regularly increasing or decreasing order of frequencies) were generally matched somewhat more accurately than the irregular progressions of the same tones. Considering the four experimental groups in Experiment I, means for all permutations varied from 58% correct with 0.3-semitone separation to 73.5% with 9-semitone separation, with accuracy increasing regularly with increasing frequency separation. Since subjects had no practice judgments and no information concerning the accuracy of their responses, the overall performance indicates that the task was quite easy. The pitch patterns formed with 0.3-semitone separations (four frequencies from 1,000 to 1,053 Hz) consist of smaller frequency changes than used in melodic groupings in music of our culture, and the patterns corresponding to 9-semitone separations (four frequencies from 500 to 2,378 Hz) have larger frequency changes than found in melodic groupings. It is of interest that the performance of these groups was comparable to that observed for those subjects hearing either the 1-semitone or the 3-semitone separations (which corresponded more closely to the size of familiar steps between notes). Thus, it can be concluded that recognition of tonal patterns in this experiment apparently is not based simply upon experience with similar musical patterns, but reflects a more general perceptual ability. Further, it is of interest to note that melodic groupings appear to use only a small part of the range of readily discriminable temporal patterns of pitches.

Discriminable auditory patterns permitting recognition of different permuted orders have been called "temporal compounds" by Warren (1974a, Note 1) in analogy to chemical compounds. Both types of compounds have characteristics rather different from those of the component elements. Further, such compounds exist as organized entities which may not be resolvable directly into their component items without special analytical procedures.

In Experiment IV, the accurate matching of patterns of tones lasting only 50 msec (well below the threshold for direct naming of order) indicates that temporal compounds are formed readily at this duration. Fifty milliseconds is also well below the range of durations from about 150 through 900 msec used for the organization of melodic themes in music (Fraisse, 1963, p. 89). It seems to have been considered generally that, within music, use of tones as brief as 50 msec does not permit formation of discriminable permuted patterns. Thus, Winckel (1967, p. 54) has stated that with notes occurring at 50-msec intervals, metathesis occurs in musical passages, with listeners unable to detect order. Such a rapid presentation of successive notes is found occasionally in compositions by Liszt and Ravel, but Winckel stated that the effect of these very rapid sequences was only to introduce a "flickering or rustling" to the selection. Yet, in Experiment IV, subjects could readily distinguish between permuted orders at 50 msec/item, so that perceptual metathesis did not occur. Perhaps the lower limit of durations used for melodic construction does not result from inability to distinguish between orders of the components, but rather some other requirement for melodic sequences. It may be that conventional usage in music is governed in part not only by perceptual limitations but also by limits in generating sounds: aside from glissandi, sequences of successive pitches presented at the rate of 20/sec may be beyond the usual limits of motor performance for some musical instruments (and for singing), and hence generally not available for composition.

However, while we have seen that temporal compounds can be formed readily with frequency separations both greater and less than melodic usage and with durations below those employed with recognizable musical themes, there is a limiting parameter in music that does correspond to a limit we have found for matching of tonal sequences. When we used tones having frequencies above the musical range in Experiment II (4,500 to 7,569 Hz), mean correct matches dropped sharply. Even the "4321" (descending) and "1234" (ascending) glissandi could not be matched accurately. Ward (1954) observed that "pitch" seemed to disappear for trained musicians above the musical range, and it appears possible that, although our high-frequency tones had separations well above the just noticeable difference in frequency, the lack of qualitative pitch differences between tones may have inhibited temporal compound formation.

Experiment III indicated that short silent intervals do not facilitate matching of tonal sequences, although there is evidence that brief silent periods between items facilitates matching with sequences consisting of nonrelated sounds (Warren, 1974a). The lack of any appreciable effect of silence upon matching recycled sequences of tones has been found independently by other investigators. The recent results of Nickerson and Freeman (1974) are in agreement with our observations in both Experiments I and III. They trained subjects to identify each permuted order of four recycled tones with one of the numbers from 1 through 6. Although they considered the naming of the appropriate number as "identification" of order, it would be considered as "matching" in our system, since subjects recognize or match the patterns with remembered distinguishing labels rather than naming the individual components in the proper order. Nickerson and Freeman found, as we did in Experiment I, that accuracy of judgments increased with increasing frequency separation of components. They also introduced silent intervals between successive tones (keeping the tone plus silence equal to 200 msec), and found, as we did in Experiment III, that silence produced no conclusive effect upon accuracy. One part of their study dealt with performance of a single highly trained subject, and it was found that accuracy of matching each sequence with its appropriate number name was possible down to durations of only a few milliseconds.

Naming of Order

A comparison of Tables 1 and 2 indicates that, not only does performance differ greatly for matching and naming within the groups receiving identical sequences for matching and for naming, but changes in conditions across groups did not have comparable effects upon the two types of judgments. This indication of fundamental differences between processes underlying matching and naming is consistent with the dichotomous formulation of Warren (Note 1). Direct naming of order of recognizable components in extended sequences (Type I discrimination) was considered to require the completion of naming of each sound while it occurred, before stimulation by the next item to be named produced interference with verbal encoding. Garner (1951) found that the lower limit for counting identical tone bursts in extended sequences (which could be considered as attaching successive verbal labels to each item) was about 170 msec/item, and it appears that a roughly equivalent duration is required for responding with the names of sounds in proper order while a sequence is in progress. The speed of naming of sounds in order of occurrence is facilitated when items are readily discriminable and the names familiar, and is fastest when the items require an echoic response, as with a sequence of vowels (Thomas, Hill, Carroll, & Garcia, 1970; Warren, 1968).

With our recycled tonal sequences, there was a general tendency for the ascending ("1234") and descending ("4321") orders to be named more readily than the other four permuted arrangements. This identification of rising and falling pitch may have involved initial recognition of the overall pattern rather than direct naming of the individual components. Once the glissandi are recognized, the correct order of components can be inferred. There is evidence that identification of phonemic orders within words is accomplished through such initial identification of temporal compounds followed by an inferred, indirect identification of phonemes and their order (Savin & Bever, 1970; Warren, 1971). However, for the four out of six permuted tonal sequences in which frequency changes did not form a simple rising or falling pattern, the mechanisms for verbal labeling of each component would be expected to set the limit for accurate naming of order.

Experiment I indicates that order naming was quite difficult with a separation of 0.3 semitones. Although the differences between adjacent tones are below those encountered in Western music, the frequencies have separations well above the just noticeable difference, and despite the difficulty in naming, Table 1 shows that matching is accomplished readily for the sequence. Once separation was increased to 1.0 semitone, accuracy of naming order in Experiment I remained quite constant with increasing tonal separation (up to 9 semitones).

There is evidence in the literature that would lead us to anticipate that naming of order should become more difficult with greater frequency separation. Thomas and Fitzgibbons (1971) presented a paper reporting that an increase in frequency separation in recycled sequences containing four tones led to poorer performance in naming the order of the components. Separations greater than a musical fourth (6 semitones) from highest to lowest frequencies in the sequence (an average of 2 semitones between components) were found to reduce accuracy of order identification. Norman (1967) had used a somewhat different procedure, and also found increasing difficulty in naming of order of tones with increasing frequency separation. He alternated two 100-msec tones having a separation less than 2 semitones with 30-msec silence separating these tone bursts. A 30-msec probe tone was inserted into one of the intertone intervals, and Norman found that when the frequency of the probe tone was much higher or lower than the background tones, it was very difficult to tell whether the probe followed the higher or the lower note. Bregman and Campbell (1971) concluded that recycled sequences of tones which are close in frequency form a single "auditory stream" permitting identification of the order of the tones. They suggested that each item in a recycled sequence which did not form a stream with temporally adjacent items would be subject to "primary auditory stream segregation" forming a perceptual stream with its own prior and subsequent iterations, thereby preventing judgment of order relative to the other sounds in the sequence.[1] As described earlier, Nickerson and Freeman (1974) reported a contrary effect—that identification of order with a recycled sequence of tones became easier with increasing frequency separation. We have seen that their procedure of using a learned number code for identification of each of the sequences would qualify as "matching" in our system, and their results are in agreement with ours for matching. In order to deal with what they saw as a possible conflict with their results and the earlier literature, they considered that stream segregation might facilitate grouping their four tones into pairs, which could aid detection of order by breaking the task down into simpler parts. A facilitation similar to this could be attained in our study by the response procedure involving arranging of cards. Further, there might be some enhancement of perceptual pairing of neighboring pitches with greater overall frequency separation in our experiment, which could cancel any increase in overall difficulty associated with the entire pattern due to increasing frequency separation.

In Experiment II, the tones (3-semitone separation) were above the musical range (4,500 Hz, and above). Discrimination between the pitch of tones lying above the musical range, as noted in the discussion of pattern matching, is quite poor. This lack of readily discriminable differences may have led to difficulty in assigning the proper verbal labels to the individual tones, resulting in the poor performance observed for naming of order with the high-frequency tones.

The introduction of silent intervals between tones in Experiment III did not help the naming of order, although with sequences of nonrelated sounds (such as hisses, tones, and buzzes), it has been reported that short silent intervals do facilitate naming (Neisser, Note 2; Warren, Note 1).

It is of interest that naming of order seems to follow the rules governing use of tonal patterns in music more closely than does matching of order. Thus, a duration of 50 msec/item is too rapid for both naming of order in our experiment and also for melodic grouping in music, but matching is quite accurate at this presentation rate. Also, when frequency separation was 0.3 semitones in Experiment I (corresponding to a smaller pitch difference than that used in Western music), naming of order was impaired, but patterns could still be matched readily.

The results obtained with tonal sequences in this study are in agreement with those obtained for sequences of unrelated sounds and for the sequences of phonemes forming speech. It appears that matching (Type II discrimination) corresponds to a general perceptual ability to form recognizable groupings readily with brief successive sounds. Direct naming of order (Type I discrimination) follows different rules, and is possible generally with extended sequences only for individual item durations lasting longer than 200 msec.

REFERENCE NOTES

1. Warren, R. M. *Temporal resolution of auditory events.* Mimeographed copy of paper presented at a symposium at the American Psychological Association, Summer 1972.

2. Neisser, U. *On the perception of auditory sequence.* Mimeographed copy of paper presented at a symposium at the American Psychological Association, Summer 1972.

REFERENCES

Bachem, A. Chroma fixation at the ends of the musical frequency scale. *Journal of the Acoustical Society of America*, 1948, **20**, 704-705.

Bregman, A. S., & Campbell, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, **89**, 244-249.

Fay, W. H. *Temporal sequence in the perception of speech.* The Hague: Mouton, 1966.

Fraisse, P. *The psychology of time* (trans. J. Leith). New York: Harper & Row, 1963.

Garner, W. R. The accuracy of counting repeated short tones. *Journal of Experimental Psychology*, 1951, **41**, 310-316.

Nickerson, R. S., & Freeman, B. Discrimination of the order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. *Perception & Psychophysics*, 1974, **16**, 471-477.

Norman, D. A. Temporal confusions and limited capacity processors. *Acta Psychologica*, 1967, **27**, 293-297.

Savin, H. B., & Bever, T. G. The nonperceptual reality of the phoneme. *Journal of Verbal Learning & Verbal Behavior*, 1970, **9**, 295-302.

Thomas, I. B., & Fitzgibbons, P. J. Temporal order and perceptual classes. *Journal of the Acoustical Society of America*, 1971, **50**, 86-87. (Abstract)

Thomas, I. B., Hill, P. B., Carroll, F. S., & Garcia, G. Temporal order in the perception of vowels. *Journal of the Acoustical Society of America*, 1970, **48**, 1010-1013.

Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, **26**, 369-380.

Warren, R. M. Relation of verbal transformations to other perceptual phenomena. In *Conference Publication No. 42, IEE/NPL Conference on Pattern Recognition.* Teddington, England: Institution of Electrical Engineers, 1968 (Suppl., 8 pp.).

Warren, R. M. Perceptual restoration of missing speech sounds. *Science*, 1970, **167**, 392-393.

Warren, R. M. Identification times for phonemic components of graded complexity and for spelling of speech. *Perception & Psychophysics*, 1971, **9**, 345-349.

Warren, R. M. Perception of temporal order: Special rules for the initial and terminal sounds of sequences. *Journal of the Acoustical Society of America*, 1972, **52**, 167. (Abstract)

Warren, R. M. Auditory temporal discrimination by trained listeners. *Cognitive Psychology*, 1974, **6**, 237-256. (a)

Warren, R. M. Auditory pattern recognition by untrained listeners. *Perception & Psychophysics*, 1974, **15**, 495-500. (b)

Warren, R. M., & Obusek, C. J. Speech perception and phonemic restorations. *Perception & Psychophysics*, 1971, **9**, 358-362.

Warren, R. M., & Obusek, C. J. Identification of order within auditory sequences. *Perception & Psychophysics*, 1972, **12**, 86-90.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. Auditory sequence: Confusion of patterns other than speech and music. *Science*, 1969, **164**, 586-587.

Winckel, F. *Music, sound and sensation: A modern exposition* (trans. T. Binkley). New York: Dover, 1967.

Wood, A. *The physics of music* (6th ed.). London: Methuen, 1962.

NOTE

1. Perceptual auditory stream segregation was used also to explain the inability of subjects to name the order of recycled sequences of unrelated sounds (hisses, tones, buzzes), it being considered that these sounds would group with their own repetitions rather than adjacent sounds. This explanation in terms of noncontiguous perceptual groupings does not handle the ease of *matching* permuted orders of nonrelated sounds reported since Bregman and Campbell's formulation (Warren, 1974a).

(Received for publication April 20, 1975;
accepted June 20, 1975.)