

# Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians

CAROLYN DRAKE

*University of Paris V, CNRS URA 316, Paris, France*

Many sequential events, musical rhythms in particular, can be described by a hierarchical structure, with lower order events recursively combining to form higher levels. This study investigated factors influencing the ease of reproduction of short musical rhythms that reflect various organizational principles. For adults and children, reproduction was better for rhythms with the following characteristics: (1) binary rather than ternary subdivision, (2) two rather than three different durations, (3) the ability to be segmented into two shorter rhythms of identical duration, and (4) intensity accents on important hierarchical positions. These findings suggest that a prototypical temporal structure—that is, a regular beat with binary subdivisions—is functional in childhood. The ability to process complex hierarchical structure appeared to be influenced more by musical training than by passive acculturation.

Many temporal organizations of events in the environment can be described parsimoniously by a hierarchical structure with lower order events recursively combining to form higher levels. Musical rhythms are a typical example of this type of organization. It is not, however, entirely clear whether listeners apprehend some or all of this complex structure when listening to such sequences, that is, whether they are sensitive to the temporal hierarchical structure. This paper examines some characteristics of the reproduction of this hierarchical structure and investigates factors influencing the development of their representation and reproduction.

## Structure of Musical Rhythms

The term “musical rhythm” refers to rhythms usually used in Western tonal music, such as those described in formal music theories (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Yeston, 1976). They constitute a subset of all possible rhythms and can be characterized by several typical rules; three of the most important rules are examined here. First, these rhythms are constructed around a regular beat, with the majority of notes falling on beats rather than between beats. Second, note durations tend to fall into a few distinct categories, with longer durations an integer multiple of shorter durations. Fraisse (1982) counted the frequency of different note durations in a repertoire of classical music and found that the majority of pieces contained only two note durations in

a ratio of 1:2. Although the binary ratio of 1:2 is frequently considered to be the most dominant (Lerdahl & Jackendoff, 1983; Stoffer, 1985), other ratios, in particular 1:3, are commonly found in all kinds of Western tonal music. Third, these distinct categories of note durations do not follow each other in a random fashion but fit into a hierarchical structure characterized by relative timing, in which each event is situated in relation to all other events in the sequence, adjacent or not (Povel, 1981). These three characteristics are illustrated in Figure 1: the musical rhythm is organized around a regular beat, it contains three different durations, each twice as long as the one of shorter duration, and the note durations fit into a binary hierarchical structure.

## Mental Representations of Musical Rhythms

To what extent do our mental representations for musical rhythms coincide with this formal description? Again, three points arise. First, all of us have felt our feet beating in time to a favorite piece of music: This regular beat corresponds to a psychological concept called a pulse. The perception of temporal regularities in the environment has a privileged psychological status over the perception of irregular sequences: Sequences composed of events organized around a regular beat are easier to perceive and reproduce than are sequences without such an organization (Essens & Povel, 1985; Povel & Essens, 1985). This psychological pulse usually corresponds to one particular hierarchical level of the theoretical description of a rhythm. For instance, listeners usually hear the pulse depicted in Figure 1 at either Levels 2 or 3, depending on the tempo at which the rhythm is played. The pulse is important from an early age: By the age of 5 years, children already rely heavily on a pulse in the organization of spontaneous songs and in reproduction tasks

---

I am grateful to Claire Gérard for her guidance and to Mari Riess Jones and Caroline Palmer for their helpful comments on an earlier version of this paper. Correspondence should be addressed to C. Drake, Laboratoire de Psychologie Expérimentale, 28, rue Serpente, 75006 Paris, France.

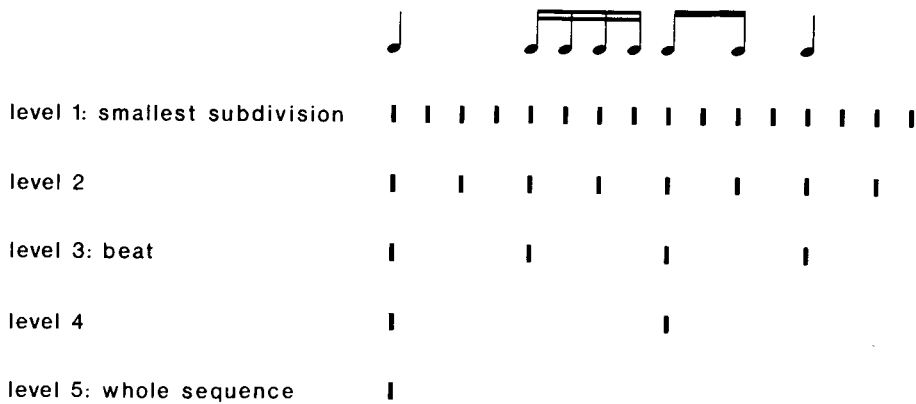


Figure 1. An example of temporal hierarchical structure associated with a musical rhythm, showing five levels ranging in length from the smallest temporal subdivision (level 1), through the beat (level 3) to the whole sequence (level 5).

(Dowling, 1984; Dowling & Harwood, 1986; Drake & Gérard, 1989).

The second point is that listeners tend to categorize durations into distinct categories that are frequently in ratios of 1:2, 1:3, or 1:4 (Schulze, 1989). Evidence from tasks involving motor skills confirms the tendency to use distinct duration categories: When asked to invent a rhythm, subjects produced rhythms composed of only two different durations—a long and a short duration—in a ratio of about 1:2 (Fraisse, 1956, 1982). Also, the duration of events in sequences to be reproduced are distorted toward these categories (Drake & Gérard, 1989; Povel, 1981).

A beat may be subdivided in many ways, but researchers frequently consider binary subdivisions to be more perceptually salient than ternary subdivisions (Jones, 1987, 1990). For instance, the detection of pitch alterations in melodic sequences was faster for sequences with a 2/4 and 4/4 meter than for ones with a 3/4 meter (Smith & Cuddy, 1989). Is this true for production tasks that involve a motor dimension? In sequences with a complex temporal structure, adults found it easier to synchronize with and to reproduce rhythms with binary subdivisions than they did rhythms with ternary subdivisions (Povel, 1981); however, this advantage disappeared when a metrical framework was provided (Povel, 1981).

Similar results were found with children: In 7- to 12-year-old children, Uptis (1987) found that binary rhythms were easier to describe and draw than were ternary ones. However, in very simple temporal organizations, the superiority of binary over ternary subdivisions was found in neither a perceptual nor a motor task (Drake, 1990; Drake, Dowling, & Palmer, 1991; Gérard & Drake, 1990): 5- and 7-year-old children were asked to discriminate and reproduce very simple isochronous sequences in which the beat interval (indicated by intensity accents) was subdivided into equal intervals in a ratio of either 1:2 or 1:3. In the perception task, they had to say whether or not two sequences with binary and ternary organizations were the

same. In the reproduction task, they had to tap louder in one of every two or three events than they did in the others. Both the timing and intensity dimensions of the ternary sequences were discriminated and reproduced as well as those of the binary ones, indicating that the superiority of binary over ternary subdivisions is not valid in sequences with a very simple temporal structure.

Third, musical rhythms are characterized by a hierarchical organization, but to what extent are listeners able to perceive more than just a pulse and binary or ternary subdivisions of this pulse? Are they able to detect complex hierarchical relationships between events? Jones and Boltz (1989) created versions of folk tunes in which the length of phrases (melodic and temporal) was manipulated while the overall duration of the tune and the individual note durations were kept constant. Thus, the hierarchical structure was manipulated without affecting the timing of individual notes. Moderately skilled musicians were asked to indicate which of two versions of a tune appeared longer. In a separate experiment, the last three notes were missing, and the subjects were required to press a button when they felt that the tune should end. The results of both tasks, one perceptual and one with a motor component, indicated that the manipulation of the temporal hierarchical structure of the tune influenced its perceived duration. Additional evidence for the perception of a multileveled hierarchical structure has been provided by Palmer and Krumhansl (1990a, 1990b), who asked subjects to indicate the goodness of fit (and detection) of events presented within a metrical framework. Performance was proportionally higher for events in more important hierarchical positions, but this effect was stronger for musicians than for nonmusicians. This was also evidenced in a motor task in which adult and child musicians and nonmusicians reproduced simple musical rhythms (Smith, 1983): The musicians organized their reproductions around a temporal hierarchical structure, whereas the nonmusicians organized their reproductions around more surface features.

What cues may a listener use to detect a temporal hierarchy? Two cues are of interest here: the particular pattern of durations and the regular occurrence of accented events. In musical rhythms, events of varying durations occur in an ordered fashion such that events at one hierarchical level combine to create units of equal duration at the next level. If listeners are able to detect this regular temporal pattern at more than one level, they will be able to extrapolate additional information about the hierarchical structure. Another cue to the hierarchy is the presence of regularly occurring accents that stand out perceptually from other events. A regular pattern of accented events is known as an accent structure (Jones & Boltz, 1989). If these accent structures coincide with the hierarchical temporal structure, they may aid listeners in focusing their attention on this organizational principle (Martin, 1972). Once listeners have identified a hierarchical structure, they acquire information about the relative importance of events, because events of higher levels are considered to be structurally more important than events of lower levels. It is therefore possible to preferentially focus more attention on relatively more important events and less attention on events of lesser importance. This selective attention should lead to more efficient processing, and perception of the sequence as a whole should be facilitated.

### Aims of This Study

Thus, music is composed of a particular subset of possible rhythmic structures characterized by an organization around a regular beat, with simple integer subdivisions of this beat, and with the beats fitting into a regular temporal hierarchy. To some extent, we appear to be well equipped to perceive such rhythms: Rhythms organized around a regular beat are easier to perceive than those that are not, simple integers (1:2 and 1:3) are easier to perceive than more complex ones, and at least some dimensions of the temporal hierarchy are apprehended by listeners. However, many questions remain unanswered: What is the relative importance of binary over ternary subdivisions? Are rhythms containing only two different durations easier to remember than rhythms containing more? How much of the temporal hierarchy is used by listeners when listening to simple musical rhythms? How do these factors vary with age and musical experience? These questions were examined by using a reproduction task in which subjects heard a short musical rhythm and were asked to immediately reproduce it by tapping on a drum.

**Type of subdivision.** The first question addressed in this study concerned the subdivision of the pulse level. The relative ease of reproducing binary and ternary subdivisions was examined by comparing the reproduction of rhythms of comparable complexity that contained either binary or ternary subdivisions. Although it is generally accepted that binary subdivisions are easier to perceive than are ternary subdivisions, this superiority has not been demonstrated in reproduction tasks of musical rhythms in which the metrical framework is well defined.

**Complexity.** Following the principles outlined earlier, the simplest rhythms to reproduce should be those that contain only two different durations in a ratio of 1:2 for the binary rhythms and 1:3 for the ternary rhythms. Examples of simple rhythms are shown in Figure 2. More complex rhythms were created when an additional duration ratio was introduced into the rhythms. For this study, complex binary rhythms also contain duration ratios of 1:4, and complex ternary rhythms also contain duration ratios of 2:3. Examples of complex rhythms are also shown in Figure 2. It is expected that the complex rhythms should be harder to reproduce than the simple rhythms because of the greater number of different duration ratios in the former (simple = 2 different durations, complex = 3 different durations). The number of hierarchical levels represented can also be examined by comparing the ease of reproduction of different sets of rhythms. The simplest set used in this study (binary with two different durations) are organized around only two hierarchical levels, whereas another set (binary with three different durations) are equivalent in all other dimensions except that they are organized around three hierarchical levels. If subjects are able to correctly reproduce the binary rhythms with three different durations, that would be an indication that the additional hierarchical level is represented.

**Pattern of durations.** Two rhythms may be constructed around the same type of subdivision, with the same number of different durations and events and with the same total duration, differing only in the pattern of successive durations. For instance, two rhythms may contain two short durations (s) and two long durations (l), but they may be ordered *ssll*, *lssl*, or *llss*. Does this pattern of durations influence the difficulty of reproducing a rhythm? In each set of rhythms, the ease of reproduction of rhythmic sequences that only varied in the pattern of durations was examined to investigate whether this factor influences their difficulty and, if so, to see which other factors affect this difficulty. Particular pattern durations that emphasize a higher hierarchical level are expected to be easier than those that do not.

**Role of accents.** The next question considers whether the reproduction of a rhythm may be enhanced by means of regularly occurring intensity accents that highlight the temporal hierarchical structure. The same rhythms were presented either with all events played at the same intensity level (without accents) or with events falling on beats played louder (with accents), as seen in Figure 2. If accents do facilitate the perception of the hierarchical structure, performance should be better for sequences containing intensity accents than for sequences without accents.

**Role of acculturation and training.** The fifth issue considers the influence of acculturation and formal music training on the ability to reproduce rhythmic sequences. Acculturation refers to the combined influences of maturation and passive exposure on the development of mental representations for this music (Zenatti, 1981). Music training involves an explicit and time-consuming learn-

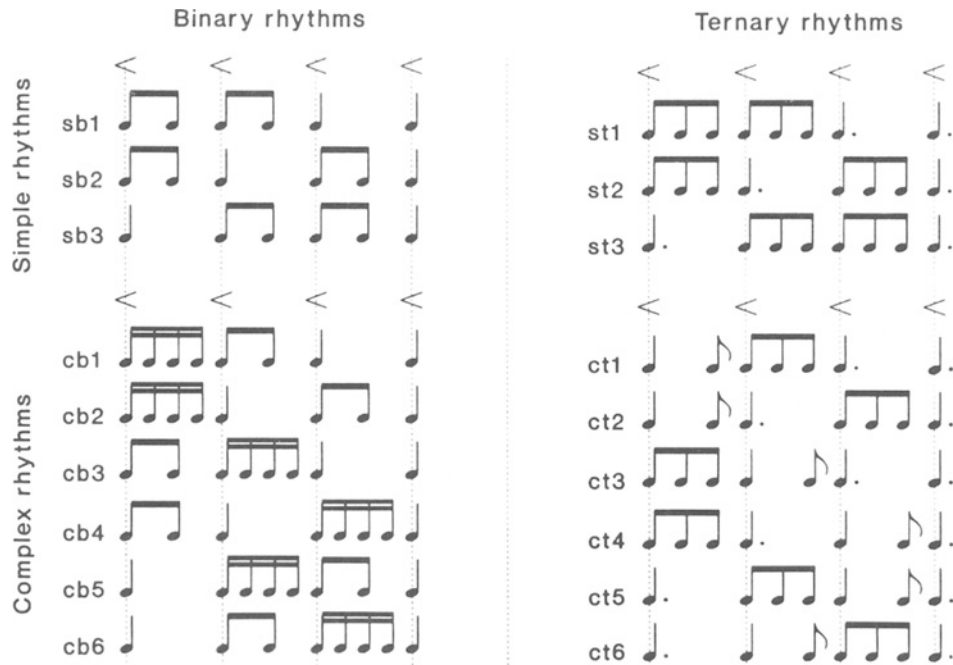


Figure 2. Four rhythm types used as stimuli (sb = simple binary, st = simple ternary, cb = complex binary, and ct = complex ternary), showing the position of accents in the with-accents condition.

ing process in which pupils are familiarized with both the practical and theoretical aspects of music. In this research, the ability of 5- and 7-year-old children to reproduce rhythms is compared with those of both adult nonmusicians and adult musicians to determine the relative importance of acculturation and training. Children and adults differ in their degree of acculturation, and adult musicians differ from the others in their formal music training. The ability to perceive a temporal hierarchical structure may be (1) functional at an early age (therefore behavior should be similar in all four subject groups), (2) acquired through maturation and passive exposure to music (hence improvements should be seen between the 5- and 7-year-old children, between the children and the adult nonmusicians, and between the adult nonmusicians and musicians), or (3) acquired only through specialized music training (therefore this behavior should be observed only in the adult musicians). Of course, any combination of these three possibilities is also conceivable.

To summarize, subjects heard simple musical rhythms that varied along several dimensions—type of subdivision (binary/ternary), complexity (simple/complex), pattern of durations, and accentuation (absent and present)—which allowed for an examination of the relative importance of these organizational principles. The ease of reproduction was compared for four groups of subjects (adult musicians, adult nonmusicians, and 5- and 7-year-old children), which made possible an investigation of the factors influencing their development.

## METHOD

### Subjects

Four groups of 12 subjects each participated in this experiment. The adult musicians had received at least 5 years of formal musical training and played their instruments almost every day (mean number of years = 9). The adult nonmusicians had received less than 1 year of formal musical training. There were two groups of children, 5 years old (mean = 5.5 years) and 7 years old (mean = 7.7 years), who were pupils at a small private primary school in England, at which they received an introduction to music through singing and playing simple instruments.

### Materials

Eighteen short musical rhythms were constructed around four beats of 900 msec each—thus, all were of the same total duration. Half of the rhythms contained binary (b) subdivisions; the other half contained ternary (t) subdivisions. The rhythms were either simple (s = only two different durations with ratios of 1:2 [binary] or 1:3 [ternary]) or complex (c = three different durations with additional ratios of 1:4 [binary] and 2:3 [ternary]). Thus, four types of rhythms (sb, st, cb, and ct) were created, as shown in Figure 2. The pattern of note durations was systematically manipulated for each rhythm type to give all possible combinations, leading to three patterns of durations for the simple stimuli and six patterns of durations for the complex stimuli. There were three restrictions on the manipulations: the last interval was always long, two short intervals always occurred together in the binary stimuli, and a short interval was always associated with a long interval in the ternary stimuli. Previous studies have indicated that the number of beats (regular time markers around which a rhythm is structured) rather than the number of events in a rhythm is the most important factor influencing its difficulty (Drake, 1990; Drake & Gérard, 1989).

For this reason, the number of beats rather than the number of events was kept constant. Even so, although the number of events was not identical in the four types of rhythms (sb = 6, st = 8, cb = 8, ct = 7), they are very close for the overall comparison between simple and complex rhythms (weighted average for simple = 7, for complex = 7.5) and identical for comparisons between binary and ternary rhythms (average weighted for the number of sequences in each group = 7.3 for both binary and ternary). The stimuli were presented either without intensity accents (all the notes were played at 80 dB SPL) or with intensity accents (all of the notes falling on the beat were played at 90 dB SPL, with the others at 80 dB SPL).

### Apparatus

The rhythms were played on a Roland TX 505 rhythm box by selecting a bass drum sound (which resembles the sound produced by the drum used by the subjects). Interonset durations were manipulated, but the onset-to-offset duration of events was constant (60 msec). The subjects reproduced each rhythm by banging a wooden stick on a drum. The reproductions were recorded on audio tape and decoded through an analog/digital converter on a BBC computer (sampling rate = 2 msec) that was programmed to calculate both the duration between the beginning of successive taps (interonset durations) and the maximum intensity of each tap (scale of 0-127).

### Procedure

The subjects heard each rhythm once and were asked to reproduce it on a drum. No mention was made about whether they had to reproduce both timing and intensity variations, and none of the subjects asked for details. There was a pause of about 5 sec between the end of a reproduction and the presentation of the next rhythm. Within a session, the stimuli were presented in a counterbalanced order, with each rhythm type (sb, st, cb, and ct) blocked together. Half of the subjects in each group were randomly assigned to each of the two orders (with/without, without/with accents). The subjects were tested individually in a quiet room in the school or in the laboratory. The experimental session lasted about 30 min, and some adult subjects received course credit for their participation.

### Dependent Measure

A reproduction was considered correct if all of the durations in the reproduction corresponded to those in the model presented to the subjects ( $\pm 10\%$ ) and/or all of the duration ratios of successive events corresponded to those in the model ( $\pm 10\%$ ). The first category means that the reproduction respected both the pattern of durations and the tempo of the model (absolute); the second category means that the reproduction respected the pattern of durations but not the tempo of the model (relative). The percentage of correct reproductions (absolute or relative) is taken as the dependent variable in the following analyses.

## RESULTS

An analysis of variance (ANOVA) on the percentage of correct scores by complexity (simple or complex), type of subdivision (binary or ternary), accentuation (present or absent), and subject group (musicians, nonmusicians, and 7- and 5-year-olds) was performed. Statistically reliable main effects were observed for all the factors and all the two-way interactions. However, there were no statistically reliable three-way interactions. The influence of each main factor and its interactions with other factors will be examined separately.

### Rhythm Type

**Type of subdivision and complexity.** There was a main effect of type of subdivision [ $F(1,88) = 339.7, p < .01$ ], with binary subdivisions reproduced correctly more frequently than those with ternary subdivisions. There was also a main effect of complexity [ $F(1,88) = 450.8, p < .01$ ], with the simple rhythms reproduced correctly more frequently than the complex rhythms (see Table 1). A statistically reliable interaction between complexity and type of subdivision [ $F(1,88) = 485.3, p < .01$ ] showed that complexity had a greater effect for the stimuli with binary subdivisions than for those with ternary subdivisions (see Table 1), with rhythms containing binary subdivisions and only two different durations proving easier to reproduce than other rhythms.

**Pattern of durations.** The percentage of correct reproductions for the 18 stimuli is presented in Figure 3. Remember that within a set (sb, cb, st, and ct), the rhythms only differ in the pattern of the durations: They are all of the same total duration, and all contain the same number and type of intervals. The ANOVA revealed four statistically reliable differences within three of the four rhythm types [for sb,  $F(2, 176) = 4.2, p < .01$ ; for st,  $F(1,88) = 6.1, p < .01$ ; for cb,  $F(5,440) = 21.8, p < .01$ ; and for ct,  $F(5,440) = 2.6, n.s.$ ]. Post hoc comparisons (Tukey) confirmed that for the simple binary rhythms, sb2 was reproduced better than sb1 and sb3 [ $F(2, 176) = 4.2, p < .01$ ], and for the simple ternary rhythms, st2 was reproduced better than st1 and st3 [ $F(1,88) = 6.1, p < .01$ ]. For the complex binary rhythms, cb3 and cb6 were reproduced less well than were the others [ $F(5,440) = 21.75, p < .01$ ]. No statistically reliable differences were seen within the complex ternary rhythms. This pattern of results was observed for all subject groups. Can these easier duration patterns be distinguished from the others? All of the rhythms that were reproduced correctly more frequently are characterized by a long duration in the middle, which separates them into two subunits of equal duration.

### Role of Accents

The same ANOVA that was used in the previous section also revealed a statistically reliable effect of accentuation [ $F(1,88) = 42.6, p < .01$ ], with the rhythms containing accents reproduced correctly more frequently than those without accents (without accents = 46%, with accents = 57%). There was a statistically reliable inter-

**Table 1**  
Percentage of Correct Reproductions for the Four Rhythm Types Averaged Over the Four Subject Groups and Weighted for the Number of Rhythms in Each Rhythm Type

	Binary	Ternary	<i>M</i>
Simple	90	47	69
Complex	52	35	44
<i>M</i>	65	39	

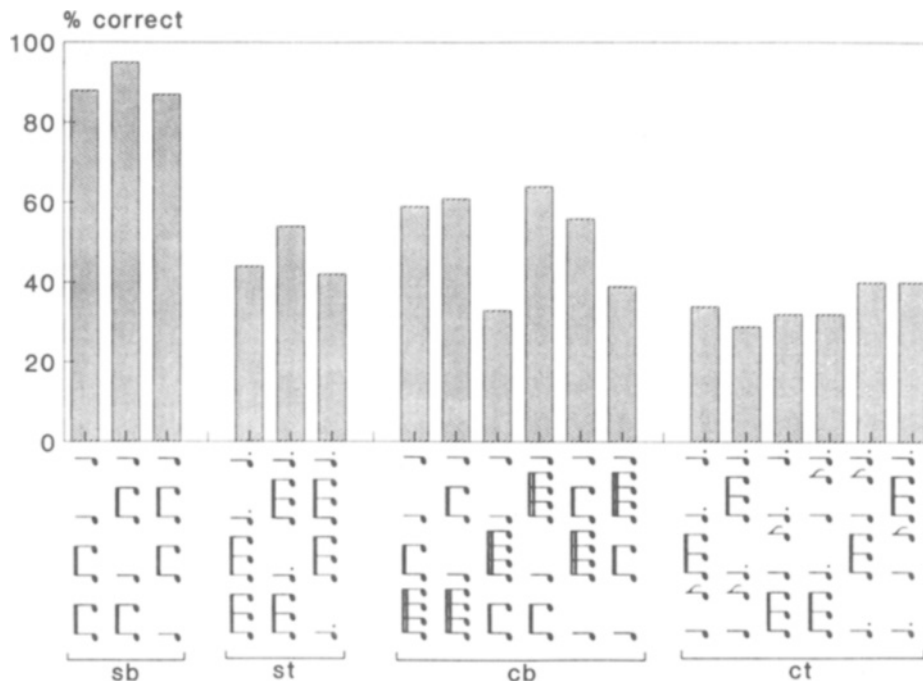


Figure 3. Percentage of correct reproductions of the four rhythm types (sb = simple binary, st = simple ternary, cb = complex binary, and ct = complex ternary) averaged over the four subject groups (adult musicians, adult nonmusicians, and 5- and 7-year-old children).

action between type of subdivision and accentuation [ $F(1,88) = 5.8, p < .05$ ] but not between complexity and accentuation: greater improvements were seen for the ternary (without accents = 34%, with accents = 47%) than for the binary (without accents = 68%, with accents = 74%) rhythms. In brief, the rhythms containing intensity accents were easier to reproduce than were those without intensity accents. The greatest improvements were seen for the hardest rhythms, which may be related to a ceiling effect whereby there was less room for improvement in the easiest rhythms.

#### Differences Between Groups of Subjects

We will now examine whether the pattern of difficulty as a function of complexity and accentuation was consistent across the subject groups. The same ANOVA revealed a statistically reliable effect of the subject group [ $F(3,88) = 86.5, p < .01$ ], with the highest scores for the musicians (91%) and lowest scores for the other groups of subjects (nonmusicians = 41%, 7-year-olds = 41%, 5-year-olds = 31%). Planned comparisons indicated statistically reliable differences between musicians and nonmusicians [ $F(1,44) = 182.4, p < .01$ ] and between the two groups of children [ $F(1,44) = 4.7, p < .05$ ], but the difference between nonmusicians and 7-year-old children was not statistically reliable. There were also statistically reliable interactions between type of subdivision and subject group [ $F(3,88) = 10.7, p < .01$ ] and

between complexity and subject group [ $F(3,88) = 10.1, p < .01$ ], as seen in Figure 4. The general pattern of difficulty (sb, st, cb, ct, in increasing order of difficulty) was observed for all groups of subjects, except that the differences were much less marked for the musicians, who found all of the rhythms relatively easy, perhaps because of a ceiling effect. The improvement with accentuation also varied with the subject group [ $F(3,88) = 5.0, p < .01$ ]: Table 2 shows that the greatest percentage increases were seen for the nonmusicians and the 7-year-old children in the st, cb, and ct rhythms (i.e., the hardest rhythms).

To summarize, binary rhythms were reproduced correctly more frequently than were ternary rhythms, and rhythms with only two different durations were reproduced correctly more frequently than were those with three different durations. One particular set of rhythms—binary with two different durations—was frequently reproduced

Table 2  
Percentage Increases From the Absence to the Presence of Accents by Subject Group and Rhythm Type

	sb	st	cb	ct	<i>M</i>
Musicians	0	4	0	2	2
Nonmusicians	2	22	17	18	15
7-year-olds	6	20	14	14	14
5-year-olds	6	7	7	11	8
<i>M</i>	3	13	10	11	

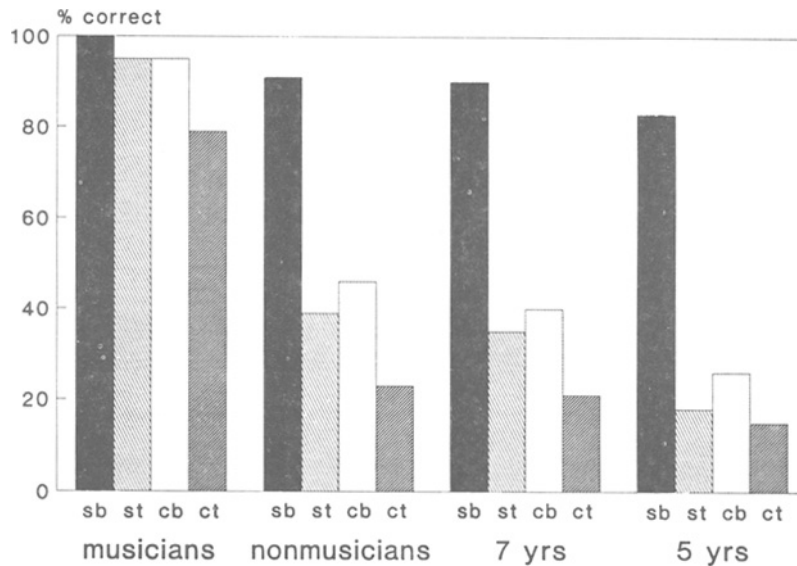


Figure 4. Percentage of correct reproductions of the four rhythm types (sb = simple binary, st = simple ternary, cb = complex binary, and ct = complex ternary) for the four subject groups (musicians, nonmusicians, and 7- and 5-year-old children).

correctly by all groups of subjects. Rhythms that could be segmented into two shorter units of equal duration were easier to represent than were other comparable rhythms that could not be segmented in this way. The placing of intensity accents on important hierarchical events led to improved performance, particularly in the hardest rhythms. This was true for all groups of subjects, but most marked for the adult nonmusicians and the 7-year-old children. Overall performance was highest for musicians, lowest for 5-year-old children, and intermediate for adult nonmusicians and 7-year-old children, whose performance did not differ significantly. The same patterns of difficulty were, however, seen in each group.

## DISCUSSION

When listening to a musical rhythm, it is theoretically possible to listen uniquely to the surface features of the rhythm, that is to say, processing each event and duration in turn in an associated chain, without attempting to identify a higher order pattern (Level 1 in Figure 1). At the other extreme, it could be possible to perceive the highest hierarchical level and thus situate each event in relation to all the others in the sequence in a complex matrix of relatively important and unimportant events (Level 5 in Figure 1). It is reasonable to assume that most people function somewhere between these two extremes, detecting more than just surface features but not necessarily able to immediately perceive complex relationships between events. The results presented here provide support for this supposition; they describe factors that influence the difficulty of a rhythm and factors involved in the development of these abilities.

This study confirms the relative ease with which people of all age and experience levels are able to reproduce sequences with just two hierarchical levels—namely, a regular beat and binary subdivisions of that beat (simple binary). Even the youngest children (5 years) almost always successfully reproduced this type of simple rhythm, nearly as well as the older children and the two groups of adults. This type of rhythm may be a “prototypical” rhythm, with the ability to reproduce this sort of rhythm appearing at an early age. This leads to the finding of the relative dominance of binary over ternary subdivisions in the case of reproduction. Although previous studies showed that 5-year-old children find tapping to binary and ternary rhythmic subdivisions equally easy, these studies relied simply on isochronous sequences (Drake, 1990; Gérard & Drake, 1990). The present research suggests that in musical rhythms, the problems encountered may be related to incorporating these subdivisions into a more complex temporal organization. Also, in previous studies, the superiority of the reproduction of binary over ternary subdivisions disappears when a metrical framework is provided. As expected, the superiority of binary over ternary subdivisions was less marked in the with-accent condition than in the without-accent condition, in which the metric framework is less obvious.

### Influence of Hierarchical Structure

Is there evidence for a hierarchical organization of more than two levels? The design of this experiment allows three types of comparisons that shed light on this problem. First, the simple and complex binary rhythms allow a comparison between rhythms organized on two main levels (sb = pulse and binary subdivisions) and three main

levels (cb = pulse and two levels of binary subdivisions). The results indicate that the two main hierarchical levels are reproduced by all groups of subjects, but that rhythms with an additional level are only reproduced correctly by the musicians, although this ability does improve between the ages of 5 and 7.

A second source of evidence comes from the comparison between the different duration patterns within a set of rhythms. Within each of the four sets, rhythms only differed by the pattern of note durations, with the note durations, total duration, and ratio between the note durations kept constant. The rhythms that could be easily subdivided into two subunits of equal duration separated by a longer duration were reproduced correctly more frequently than were the other rhythms in the set. When a rhythm can be subdivided in this way, it means that an additional hierarchical level is highlighted. This pattern was equally marked for all subject groups.

Third, the addition of intensity accents to hierarchically important events systematically affected performance. It is worthwhile to observe that, in principle, the intensity differences represent additional information that might have disrupted the perception of the sequence, thus causing a deterioration in reproduction. Also, the subjects could have tried to reproduce the intensity variations, which would have made the task harder. On the contrary, performance improved when accents were added, indicating that the accents actually aided in the perception of the rhythms. This is probably because they serve to highlight the hierarchical structure and thus lead to a more efficient coding of the rhythm. Of great interest is the fact that improvements were seen for all groups of subjects, including the youngest children. In this study, we chose to add intensity accents to all events on the beat level, which emphasized this particular hierarchical level. An alternative strategy, whereby the intensity of the accents varies as a function of the metrical importance of the events, will be the subject of future studies.

### **Role of Acculturation and Training**

As described earlier, a process influenced by musical training may result in similar patterns of behavior for children of all ages and for adult nonmusicians, whereas a process influenced by acculturation may gradually improve with age. Also, if an ability is functional at a young age, similar behaviors will be observed for all groups of subjects. Instances of all three patterns have been found in this study.

By the age of 5 years, the children were able to successfully reproduce simple rhythms only when they were organized around two hierarchical levels (pulse and subdivisions) and when they contained binary subdivisions (the simple binary rhythms in this experiment). This ability did not improve significantly with age or experience; it was already extremely high at the younger age.

Some abilities improved with age. For instance, the ability to reproduce ternary subdivisions and the ability to incorporate a higher number of different durations into

a sequence improved considerably between the ages of 5 and 7. However, this improvement did not continue with increased acculturation; no improvements were seen between the 7-year-old children and the nonmusician adults. These abilities did, however, improve with training, since the most marked differences were seen between the musicians and all of the other subjects. In other words, some abilities required to reproduce the rhythms presented here improved both between the ages of 5 and 7 and with musical training.

Why don't adult nonmusicians do better than 7-year-old children? Can we reject all possible influence of acculturation in these abilities and conclude in favor of an influence of maturation for the children and a main influence of training for the adults? We are wary of drawing the first conclusion, because the children in this study came from a school in which music plays an important role in the curriculum. All of the children were receiving a general musical education involving singing and playing simple musical instruments. The 7-year-olds had therefore received a higher degree of music training than had the younger children. Perhaps we can conclude in favor of the unique influence of training. Also, we are wary about concluding that there was an absence of an influence of acculturation simply on the evidence that the adult musicians did not do better than the 7-year-old children. Other factors are assuredly involved. For instance, the children were eager to participate in the experiment, finding it quite natural to copy rhythms just as they did in class. However, the majority of the adult nonmusicians were reluctant about taking part in the study, repeatedly saying that they were not "musical" and did not have a good "sense of rhythm." Another possible explanation for the lack of differences between the performances of adult nonmusicians and the 7-year-old children may be related to the range of scores for these two groups: performance was nearly 100% for the simple binary rhythms and below 50% for the other sets. A ceiling effect may have occurred for both groups on the former type of rhythm, and a floor effect may have occurred on the latter types. Adults may, in fact, have the capability to represent more levels of a rhythmic hierarchy than do children. Other studies, perhaps using longer, more complex, rhythms could address this point.

Let us draw a tentative picture of what is happening when subjects listen to rhythmic sequences and then reproduce them. Static coding models (e.g., Povel & Essens, 1985) can only process the rhythm once it has finished, incorporating each event into the global picture. Rhythmic sequences by definition, however, occur in time, with the listener hearing the beginning before the end. A processing model is therefore necessary to capture the exact behavior involved. On hearing the first few events in a rhythm, it is reasonable to assume that listeners will focus on local relationships between these events, concentrating on the serial relationships between successive events. Thus, the second duration may be longer, shorter, or the same as the first duration, and the exact ratio should



be defined. However, as more successive events are processed, a larger pattern may emerge, with listeners' attention directed at higher order regularities. This process would involve incorporating information about serial ratios into a hierarchical organization based on embedded ratios. As we have seen, some structural characteristics of the rhythms appear to facilitate the passage from a serial to a hierarchical organization. Two cases have been examined here: first, the patterns of durations that emphasize a higher order regularity because the rhythmic groups begin on hierarchically important positions, and second, the occurrence of intensity accents on the important hierarchical positions. In both cases, reproduction was better for rhythms that obeyed these organizational principles. Indeed, the best performances were observed for sequences that obeyed both, and the worst for those that obeyed neither. So, to summarize, listeners might first focus on the relationships between the first few events and then gradually abstract higher order relations among events. They may rely on many cues to do this, including the regular occurrence of intensity and temporal accents.

Finally, there is the question of whether the findings in this reproduction task teach us anything about the perception of musical rhythms. Obviously, no direct conclusions are warranted because a reproduction task involves a motor as well as a perceptual component. There is, however, evidence suggesting a relatively close relationship between results found with both types of tasks. For instance, Jones and Boltz (1989) found evidence reflecting a common activity in duration judgments and overt productions of musical events. Gérard & Drake (1990) also found converging evidence from both perceptual and motor tasks. In addition, a reproduction task may avoid some decisional aspects that may interfere with more purely perceptual tasks. With these reservations, it is tempting to consider that one of the main factors influencing the difficulty of reproducing a rhythm may be the difficulty of perceiving the structure of that rhythm: A rhythm that is easy to perceive will be relatively easy to reproduce. All of the conclusions drawn in this paper about the factors influencing the difficulty of reproducing a rhythm (the relative importance of binary and ternary subdivisions, the influence of the number of different durations in a rhythm, the role of accents in highlighting the hierarchical structure, and changes in abilities with age and experience) may to some extent be relevant to the perception of these dimensions. Future work will examine further the relationship between the perception and performance of musical rhythms in development.

#### REFERENCES

COOPER, G., & MEYER, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.

- DOWLING, W. J. (1984). Development of musical schemata in children's spontaneous singing. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North-Holland, 145-163.
- DOWLING, W. J., & HARWOOD, D. L. (1986). *Music cognition*. Orlando, FL: Academic Press.
- DRAKE, C. (1990). *Processus cognitifs impliqués dans l'organisation du rythme musical*. Unpublished doctoral dissertation, University of Paris V, Paris.
- DRAKE, C., DOWLING, W. J., & PALMER, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, **8**, 313-332.
- DRAKE, C., & GÉRARD, C. (1989). A psychological pulse train: How young children use this cognitive framework to structure simple rhythms. *Psychological Research*, **51**, 16-22.
- ESSENS, P. J., & POVEL, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, **37**, 1-7.
- FRAISSE, P. (1956). *Les structures rythmiques*. Louvain, Belgium: Publications Universitaires de France.
- FRAISSE, P. (1982). *Rhythm and tempo*. In D. Deutsch (Ed.), *The psychology of music* (pp. 203-254). New York: Academic Press.
- GÉRARD, C., & DRAKE, C. (1990). The inability of young children to reproduce intensity differences in musical rhythms. *Perception & Psychophysics*, **48**, 91-101.
- JONES, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, **41**, 621-634.
- JONES, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology*, **9**, 193-228.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.
- LERDAHL, F., & JACKENDOFF, R. (1983). An overview of hierarchical structure in music. *Music Perception*, **1**, 229-252.
- MARTIN, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, **79**, 487-509.
- PALMER, C., & KRUMHANSL, C. L. (1990a). Independent temporal and pitch structures in the determination of musical phrases. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 116-126.
- PALMER, C., & KRUMHANSL, C. L. (1990b). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 728-741.
- POVEL, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 3-18.
- POVEL, D.-J., & ESSENS, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-440.
- SCHULZE, H. H. (1989). Categorical perception of rhythmic patterns. *Psychological Research*, **51**, 10-15.
- SMITH, J. (1983). Reproduction and representation of musical rhythm: The effects of musical skill. In D. R. Rogers & J. A. Sloboda (Eds.), *Acquisition of symbolic skill*. New York: Plenum.
- SMITH, K. C., & CUDDY, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 457-471.
- STOFFER, T. H. (1985). Representation of phrase structure in the perception of music. *Music Perception*, **3**, 191-220.
- UPTIS, R. (1987). Children's understanding of rhythm: The relationship between development and music training. *Psychomusicology*, **7**, 1-17.
- YESTON, M. (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.
- ZENATTI, A. (1981). *L'enfant et son environnement musical*. Paris: Presses Universitaires de France.

(Manuscript received November 4, 1991;  
revision accepted for publication May 20, 1992.)