# Rhythm Categorization in Context

Edward W. Large

Center for Complex Systems
Florida Atlantic University
777 Glades Road, P.O. Box 3091
Boca Raton, FL 33431-0991
USA

large@walt.ccs.fau.edu

Keywords: Rhythm, meter, categorical perception, dynamics, context

The perception of rhythmic patterns exhibits certain features of categorical perception, including abrupt category boundaries and nonmonotonic discrimination functions (Clarke, 1987; Schulze, 1989). Other attributes of rhythm categorization, however, such as good within-category discrimination and strong dependence on context have special implications for the perception of musical rhythm. It has been suggested that two processes operate in rhythm perception, one assigns rhythms to categories depending on metrical context, while another interprets category deviations as expressive (Clarke, 1987). This interpretation possesses a certain circularity, however. Perceived rhythmic patterns influence the perception of metrical structure, while metrical structure influences the perception of rhythmic patterns. How does a temporal sequence give rise to a structural interpretation? How does metrical structure subserve both categorization and discrimination? Which temporal fluctuations are expressive, and which force structural reinterpretation? The current study aims to address these issues by investigating the role of rhythmic context in the categorization of temporal patterns.

## Background

In a pioneering study, Clarke (1987) demonstrated that the categorization of rhythmic patterns was sensitive to metrical context. Music students listened to short musical sequences in which the durations of the final two time intervals were varied systematically to create an interval ratio between 1:1 and 2:1, inclusive. Musicians were asked to categorize the ratio of the final two intervals as either 1:1 and 2:1. The musical sequences were presented in two blocks, one providing the context of duple meter, the other a context of triple meter. Clarke found that the position of the category boundary shifted according to metrical context: Ambiguous ratios (between 1:1 and 2:1) were more likely to be categorized as 2:1 in the context of triple meter, whereas these same ratios were more likely to be categorized as 1:1 in the context of duple meter. Moreover, in a discrimination task Clarke discovered a nonmonotonic discrimination function with a single peak at the category boundary, providing evidence for categorical perception.

Schulze (1989) criticized Clarke's findings on two grounds. First, he argued, because listeners were forced to choose between two categories, the category boundary shift might not be perceptual; it might simply reflect a shift in response criterion. Second, because tempo was held constant, listeners might not be performing rhythm discrimination task at all, but rather a time discrimination task. Thus, the evidence for categorical perception of rhythmic patterns might be suspect as well. To control for these factors Schulze (1989) asked two musicians to learn numerical category tags for prototypical rhythmic patterns. Then, during a response phase, the tempo of the patterns was roved randomly from trial to trial, and listeners' rated rhythms according to the degree to which they were perceived as realizations of the prototypical patterns. When the ratings were used to derive a measure of discriminability Schulze found nonmonotonic discrimination functions. But these were not the single-peaked functions of classic categorical perception (Liberman, et. al., 1957); these discrimination functions contained multiple peaks. These results suggest that rhythmic patterns, heard out of context, are not perceived categorically.

The specific question of whether or not rhythmic patterns are perceived categorically may be beside the point, however. First, a large branch of research into categorical perception has called the entire categorical/continuous distinction into question (Macmillan, 1987). Second, Clarke reported within-category discrimination that was much better than is typical of other categorical judgements. In support of this observation, Jones & Yee (1993) have reported that time discrimination is better in metrical than nonmetrical contexts. Finally, it is quite clear that musicians (at least) categorize rhythmic patterns all the time, hence the ability to notate musical performances. Thus, the more relevant questions would seem to be: What is the role of context in the categorization of rhythmic patterns?, and Can this phenomenon tell us something about how people perceive metrical structure?

If Clarke's (1987) interpretation is correct, that meter provides the categories available for rhythmic pattern classification, then his finding may indeed inform us as to the nature of meter perception. In dynamical systems terms, Clarke's data provide evidence of hysteresis in meter perception, the persistence of a percept (e.g. a duple meter) despite a change in the stimulus that favors an alternative pattern (e.g. a triple meter). For example, understanding the influence of context in rhythm perception could help to address the basic issue of whether meter perception is best described as a linear (e.g. Scheirer, 1998; Todd, et. al. 1999) or a nonlinear (e.g. Large & Kolen, 1994; Large, 2000) dynamical system, because nonlinear dynamical systems exhibit hysteresis, whereas linear systems do not.

## A Categorization Experiment

The initial objective of this study was to establish baseline observations regarding the role of metrical context in rhythm perception for both musician and non-musician listeners. An additional objective was to establish an experimental methodology for studying context effects within a framework that supports interpretation from a dynamical systems point of view. The main requirement for assessing the existence of hysteresis is the systematic variation of a single stimulus parameter. Thus, in the case of rhythm perception, the stimulus should be gradually changed from one rhythmic figure to another. However, this raises the difficult problem of distinguishing between hysteresis in the listener's perception and hysteresis in the response. Listeners responding to a gradually changing stimulus parameter may persevere in their responses even after their percept has changed. There are other interpretive problems as well. For example, are observed hysteresis effects truly perceptual, or do listeners persist in an earlier decision while the stimulus parameter passes through values for which they are uncertain about what they are hearing? Hock, Kelso, & Schöner (1993) developed a methodology for studying perceptual hysteresis in apparent motion patterns. It allows the study of perceptual changes resulting from varying the value one stimulus parameter using a simple modification of the psychophysical method of limits. The modified method of limits procedure minimizes the potential for confounding perceptual hysteresis with response hysteresis by requiring a single response only after an entire sequence has been heard. Here, the modified method of limits procedure is applied to the categorization of rhythmic patterns in an attempt to deal with interpretive problems.

## Methods

Stimuli. Stimulus sequences were constructed using sine tones of 70ms total duration (20ms linear onset ramp, 30ms steady state, and 20ms linear offset ramp) at 262 Hz (middle C). Each rhythmic pattern consisted of three equal-amplitude tones. In one condition the inter-onset interval (IOI) between the first and third tone, termed the *base interval*, was 600*ms*, in another condition it was 300*ms*. An intervening tone partitioned the base interval into two sub-intervals, forming a ratio of 1:1, a ratio of 2:1, or one of nine intermediate ratios equally spaced between 1:1 and 2:1. Sequences were then created by increasing or decreasing the time interval ratio on consecutive cycles within the same sequence. The interval between cycles was always equal to the base interval (either 600*ms* or 300*ms*), creating a rhythmic sequence with a strong beat. Ascending and descending sequences were presented alternately within randomized blocks of trials.

Ascending sequences all began with a 1:1 ratio, and the ratio increased at a rate of one step per cycle for a variable number of cycles between one and eleven. Likewise, descending sequences all began with a 2:1 ratio, which decreased by one step per cycle, for between one and eleven cycles. Thus, both ascending and descending sequences varied with regard to how deeply they probed into the range of ratios that would lead to a transition from the perception of one pattern to the other. Subjects made one response to an entire ascending or descending sequence, they did not execute a response at each step in the gradually changing sequence, as in the standard method of limits procedure. Because there was no relationship between when the listeners heard a change and when the reported their response, observed effects of parameter change involved perceptual hysteresis uncontaminated by response hysteresis.

One methodological problem not directly addressed by the modified method of limits procedure concerns the possibility that hysteresis can be due to listeners persisting in the same decision while the stimulus parameter passes through values for which they are unsure of what they are hearing (e.g. Schulze, 1989). Because this presented a realistic possibility for these stimuli, subjects were asked two types of questions. One question was designed to ascertain when subjects began to hear a rhythm pattern prototypical of the alternative category. The other was designed to determine when the rhythmic pattern ceased to be heard as a pattern of the original type. Thus it was possible to determine when listeners heard patterns as neither duple nor triple.

Design. There were two blocks of 132 sequences in each of four daily sessions. Each block was further subdivided into six sets of 22 sequences, randomized for sequence length, and alternating between ascending and descending sequences. In the first and third sessions, listeners heard sequences based on a 600*ms* base interval. In the second and forth sessions, sequences were based on a 300*ms* interval. In the first two sessions, for the ascending sequences (starting with a duple), subjects responded *yes* if at any time during the sequence they clearly heard a triple pattern, otherwise they responded *no*. Conversely, for the descending sequences (starting with a triple), subjects responded *yes* if at any time during the sequence they clearly heard a duple pattern, otherwise they responded *no*. To disambiguate, subjects were further instructed that if unsure of whether they heard a duple or triple, they were to respond *no*. In the last two sessions, subjects were asked a different type of question. For the ascending sequences (starting with a duple), subjects responded *yes* if at any time during the sequence they heard a pattern that was clearly not a duple, otherwise they responded *no*. For the descending sequences (starting with a triple), subjects responded *yes* if at any time during the sequence they heard a pattern that was clearly not a triple, otherwise they responded *no*. To disambiguate, subjects were further instructed that if unsure of whether they heard a duple or triple, they were to respond *yes*.

Apparatus. Stimulus sequences were generated by a Max program, running on a 450 MHz Macintosh G3 computer (MacOS 9.0). Tones were generated on a Kurzweil K2500RS sampling synthesizer, controlled by the Max program via MIDI. Sequences were presented and played to listeners over Sennheiser HD250 headphones.

Subjects. Two musicians two non-musicians participated in this experiment. Only the results for musicians are reported here.

## Results

The results are illustrated in Figure 1 for Musician 1 (M1, left column), Musician 2 (M2, right column), for the 600*ms* base interval (top row), and 300*ms* base interval (bottom row). The solid lines in each graph summarize the results of the first two daily sessions. They indicate where the musicians first heard a transition to triple in the ascending condition (solid black line) and where they first heard the transition to duple in the descending condition (solid gray line). In each graph a strong hysteresis-like effect is apparent: the transition to triple in the ascending condition is to the right of the transition to duple in the descending condition.

Figure 1 goes about here.

The dashed curves show the results of the second two sessions. These indicate where the musicians first heard a transition away from duple in the ascending condition (dashed black line) and where they first heard the transition away from triple in the descending condition (dashed gray line). These clearly indicate an ambiguous region. But how should it be interpreted? Compare the two gray lines (farthest left) for M2, 600*ms*. These measure boundary the same boundary between duple and not duple, but in different contexts (ascending and descending). Likewise the two black lines (farthest right) measure the boundary between triple and not triple in different contexts. This leads to two preliminary conclusions. First, there are (at least) three categories: duple, triple, and neither. Second, category boundaries shift depending on context.

With these observations in hand, context effects can be interpreted. For M1, 600*ms*, there is a slight hysteresis in the boundary between duple and non-duple (gray lines). The boundary between triple and non-triple (black lines), however, shows a strong *enhanced contrast* effect. Enhanced contrast is the opposite of hysteresis, the switch to an alternative percept before the stimulus parameter reaches a value that favors the alternative percept. (Tuller, et. al., 1994). This is a nonlinear effect that is often observed in perceptual switching studies and is discussed in more detail below. M2 displayed enhanced contrast at every boundary. Finally, in the 300*ms* condition, M1 displayed hysteresis at every boundary. Although not reported in detail here, both non-musicians also displayed greater hysteresis in the 300*ms* condition.

Finally both listeners were interviewed regarding their perception of the patterns. M1 found the 600*ms* patterns slightly ambiguous: “Sometimes the pattern just before the switch sounded like neither”, but reported little or no perceptual ambiguity in the 300*ms* patterns. M2 reported hearing several ambiguous cycles in every sequence. The results confirm the individual introspective reports. In addition, both musicians reported hearing distinct accents in the rhythmic patterns. M1 reported basing judgements mainly on the perceived patterns of accent, while M2 reported trying to focus on pattern timing, especially in the last two sessions.

## Discussion

These results extend Clarke's (1987) finding that metrical context influences the categorization of rhythmic patterns. However, the situation is more complex than the original observations indicated. In many cases, musicians perceive patterns to be neither duple nor triple. When this third possibility is accounted for, however, metrical context remains a powerful determinant of perceptual categorization. Moreover, the methodology used here minimized the possibility that observed boundary shifts reflected decision or response processes, therefore it is likely that the these effects are truly perceptual in nature. In addition, strong individual differences can be seen. These may reflect different listening strategies, differing temporal acuity, or different interpretations of what constitutes category membership. Nevertheless, both hysteresis and enhanced contrast are present in rhythm categorization, emphasizing the importance of individual differences and illustrating the dangers or averaging over the subjects. Finally, these data hint at the possibility that perceptual categorization may show different properties depending upon the absolute length of the time intervals involved, i.e. M1's stronger hysteresis in the 300*ms* condition.

The findings of hysteresis and enhanced contrast in rhythm perception represent perhaps the first direct evidence addressing the distinction between linear and nonlinear models of beat / meter induction. The current data may be taken as a strong indication of nonlinearity in categorization of temporal intervals. However, this data goes beyond any current proposals; for example, Large's (2000) model of meter perception includes hysteresis as a basic prediction, but not enhanced contrast. Enhanced contrast in perception is usually taken as evidence of adaptation at the neural level, and this is not a feature of any current model. Thus, this data may prove useful in extending models of meter perception.

There are still at least two of problems to be addressed, however. First, the meaning of the third category, 'neither duple nor triple' is unclear. From a theoretical point of view, there are several possibilities: Both categories could be activated simultaneously, or neither category may be activated, or perhaps ambiguity indexes the relaxation time of the system, the time required for one category to be activated and the other deactivated. It is also possible that there is an actual representation of a different category, based on a finer subdivision of the base interval. Further research will be required to sort out these possibilities. Second, it should be emphasized that we do not yet know whether the perceptual categorization of rhythms in metrical contexts can be taken as diagnostic of the meter perception process itself. At this point, this is a working assumption, but one that can and should be tested.

In summary, this study extends previous findings that metrical context effects the categorization of rhythm patterns. Rhythm categorization in context is not due to decision or response processes; it is perceptual in nature. There are strong individual differences among musicians and there is some evidence that perceptual categorization operates differently depending on the absolute time intervals involved. The context effects observed constitute evidence that processes responsible for rhythm categorization within a metrical context are nonlinear. Under the assumption that rhythm categorization reflects meter perception, these results suggests that perceptual stability is an important aspect of meter perception and may be taken as evidence for a nonlinear dynamical model of metrical pattern formation.

## References


Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective, In A. Gabrielsson (Ed.) *Action and Perception in Rhythm and Music*, pp 19-33. The Royal Swedish Academy of Music, 55.

Jones, M.R. & Yee, W. (1997) Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception & Performance*, *23*, 693-709.

Hock, H. S., Schöner, G., & Kelso, J. A. S. (1993). Bistability and hysteresis in the organization of apparent motion patterns. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 63-80.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*, 177 - 208.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, In press.

Liberman, A. M., Harris, K. S., Hoffman, H. S., Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, **54**, 358 - 368.

Macmillan, N. A. (1987). Beyond the categorical/continuous distinction: A psychophysical approach to processing modes. In S. Harnad (Ed.) *Categorical Perception: The Groundwork of Cognition*, pp 53-85. Cambridge: Cambridge University Press.

Scheirer, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, 103, 588-600.

Schulze, H. H., (1989). Categorical perception of rhythmic patterns. *Psychological Research*, 51, 10-15.

Todd, N. P., Lee, C.S., & O'Boyle, D.J. (1999). A sensory-motor theory of rhythm perception. *Journal of New Music Research*, *28*, 1-24.

Tuller, B., Case, P., Ding, M., & Kelso, J. A. S. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 3-16.

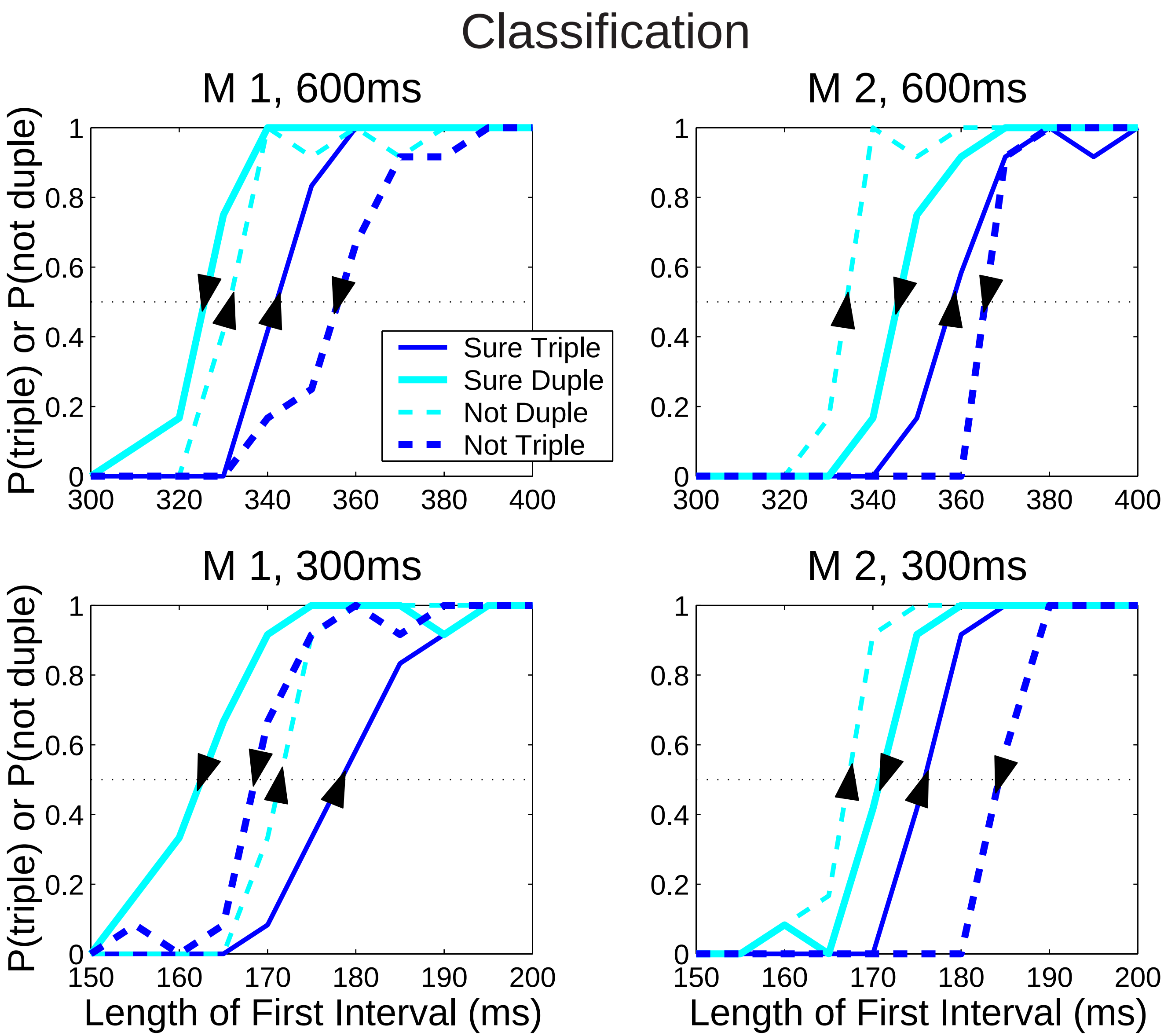
Classification
M 1, 600ms
M 2, 600ms
M 1, 300ms
M 2, 300ms
P(triple) or P(not duple)
Length of First Interval (ms)
Sure Triple
Sure Duple
Not Duple
Not Triple
0
0.2
0.4
0.6
0.8
1
300
320
340
360
380
400
150
160
170
180
190
200

Figure 1