RHYTHMIC STABILITY AS EXPLANATION OF CATEGORY SIZE*

Peter Desain[1] and Henkjan Honing[1,2]

MMM Group, NICI / University of Nijmegen (1)

Music Department, University of Amsterdam (2)

{desain, honing}@nici.kun.nl, http://www.nici.kun.nl/mmm

ABSTRACT

A measure is proposed for stability and the structural integration of a percept for a perceived temporal sequence. This measure is an elaboration of a theory of rhythm perception, named DECO, based on time intervals between any pair of onsets in a temporal pattern. More specifically, it specifies bounds between successive intervals whose durations relate as simple integer ratios. The definition of the internal stability of a mental representation of rhythm depends on the contributions of these bounds. This theory is in accordance with the responses when subjects were asked to identify rhythmic categories from performed temporal patterns.

1. INTRODUCTION

Time perception has been extensively studied in psychophysics. This research is primarily concerned with understanding the perception of isolated time intervals or intervals in isochronous sequences. Intriguing subjective judgements of duration were revealed, such as systematic over- or underestimation (e.g., Nakajima, 1987). However, these phenomena disappear when the stimulus is contained in a larger number of intervals as occurs when rhythms are performed in music. In this paper we will therefore consider a large systematic set of temporal patterns, and introduce a model that predicts the ease of identification of these patterns based on a theory of rhythmic stability.

As an example, consider the sequence of three time intervals [5 2 1], with one unit representing a duration of, e.g., 125 msec. This one second rhythm may have been notated in a musical score as shown in Figure 1a (i.e., a half note tied to an eighth note, two tied eighth notes and an eighth note) and is quite hard to perceive and produce accurately. However, a simple permutation of this sequence, e.g., [2 5 1], though not a simple rhythm, is much easier to recognize and reproduce (see Fig. 1b). Thus, accurate cognitive theories about the perception of time intervals are obliged to take order into account, to be able to deal with the full complexity of rhythmical patterns.

The relative complexity of these example rhythms is to some extend reflected in the musical notation (but note that there are many alternative notations possible). One could argue that the second rhythm (Fig. 1b) is less complex than the first (Fig. 1a) because given a certain metrical interpretation, i.e. a regular recursive subdivision scheme of time durations into two or three parts, it can be notated with less syncopation, i.e. the absence of notes on important metrical boundaries. However, this is no evidence for a distinct complexity of the underlying cognitive representation that arises when such rhythms are listened to, nor of a distinct complexity of the process of listening itself.

Another interesting example is the pattern [1 3 4] and its time reversal [4 3 1]. The first (see Fig. 1c) is normally perceived as a syncopated pattern, with a 'loud rest' (Cooper & Meyer, 1960) occurring after six time units (half way the second and third note) while the second, its reversal, is unsyncopated (see Fig. 1d), and it can be argued that therefore the first is perceived as being more complex (unstable) compared to the latter.

While there are theories formalizing the notion of perceived syncopation (e.g., Longuet-Higgins & Lee, 1984) given a meter, they leave open how one arrives at that specific metric interpretation. There is, however, a variety of computational models proposed that attempt to explain how a meter is induced by a temporal pattern (cf. Desain & Honing, 1999). With these models in mind, it could be that the second rhythm (Fig. 1b) more easily induces a beat, a regular pulse that functions as a mental framework onto which the rhythm percept can be grafted (Povel & Essens, 1985). However, most beat induction theories postulate a process which needs quite some material to allow for a stable organization behavior to emerge. This is especially true for recent beat induction models that arose from within the paradigm of complex dynamics (Large & Jones, 1999). These models are based on coupled oscillators and need several cycles (i.e. a number of beats or bars) before synchronization can occur.

In music theory the importance of grouping of individual notes into larger structures has always been recognized. Most music theoretic analyses (e.g., Cooper & Meyer, 1960) are based on assigning salience or stress to individual notes and thus forming alternating patterns of strong and weak beats. Here the onset of each note, its point of attack, attains the most relevance. Also in the search for psychological processes that govern the formation of higher level (metrical) percepts, the basic events are often the onsets of notes and their relative importance (e.g., Palmer & Krumhansl, 1990). The grouping of strong and weak beats

---

* Both authors contributed equally to this work.

into hierarchies (bottom-up process) is a central concept in these theories. Other models, for example the musical parser described by Longuet-Higgins (1976) take the reverse, top-down, approach and proceed by subdivision of larger metrical units, into two or three units at each lower level. In this paper, we will take a position which is not based on onsets only, and neither on the importance of hierarchies and recursive subdivision of time intervals, but investigate a radically different approach to defining the rhythm percept: a distributed model based on intervals and their emerging rhythmical structure that predicts the stability (ease of encoding, memorizing and recall) of a mental representation of a pattern of onsets.

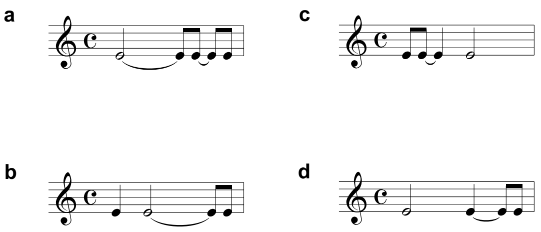

Figure 1. The time interval patterns **a** [5 2 1] and **b** [2 5 1] as notated in common music notation, with **b** being easier to recognize and reproduce. And the time interval patterns **c** [1 3 4] and **d** [4 3 1], with **c** considered syncopated and **d** unsyncopated *(see accompanying sound examples).*

2. DECO, A DECOMPOSABLE THEORY OF RHYTHM PERCEPTION

In Desain (1992) a theory is presented to explain various phenomena in the perception of temporal patterns. The mental representation of rhythm proposed there is based on the following hypotheses:

- Each time interval between any pair of onsets in a temporal pattern forms a basic component of the rhythm percept.
- Only neighboring time intervals, sharing an onset, may form a bound.
- A stable bounding between two neighboring intervals is only formed if their durations relate as a simple integer ratio (e.g., 1/2, 1/1 or 2/1).
- The strength of a binding may depend on the ratio (e.g., 1/2 may be stronger than 1/3; Fraisse, 1982).
- The strength of a binding may depend on the duration (e.g., intervals may bind stronger around a preferred range of 600 ms [Fraisse, 1982]).
- For near-integer ratios the interaction strength is characterized by a Gaussian distribution around the exact ratio.
- The bound of two neighboring intervals anchors their shared onset: each onset may thus receive stability from any interval pair it divides. The binding strengths combine in an independent, additive way.

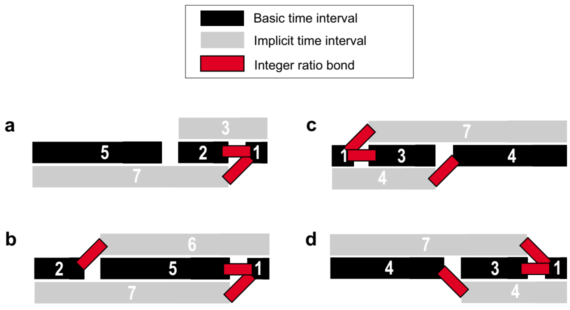

Figure 2. The time interval patterns **a** [5 2 1], **b** [2 5 1], **c** [1 3 4] and **d** [4 3 1] (black), and all intervals implied by these rhythms (light gray), plus the simple integer ratio bonds between neighboring intervals (red/dark gray). Pattern **b** is more stable than a because it has more bounds, and pattern **d** is more stable than **c** because the ratio 3/1 is stronger than 1/3. See Figure 1 for music notation.

These hypotheses are extended here with the additional rule for determining stability of the whole percept, reflecting that the instability of one onset may not be compensated for by additional stability in the other onsets:

- The stability of a pattern is determined as the geometric mean of the stability of the onsets it comprises.

(Note: this is still a simplified definition of the overall stability of a complex temporal pattern. We are currently investigating alternative definitions that take advantage of the intrinsic rhythmic structure.)

In Figure 2a and 2b the interval components of the sequences [5 2 1] and [2 5 1] are shown, together with the bonds between them. As one can see, in the [5 2 1] pattern, one constituting time interval is indeed not bound to the others, predicting a rhythm that is more difficult to perceive, memorize and produce.

For [4 3 1] and [1 3 4] the difference in stability is not reflected structurally in DECO (see Fig. 2c and 2d) but in the strength of the various integer bonds that are present in the pattern, suggesting that the binding strength, in this case, should be stronger for [3 1] than for [1 3].

Thus, not unlike how chemical bounds between atoms give stability to a molecule, or how agreement between constituents keeps the grammatical structure of a sentence in natural language together, the rhythm percept achieves stability from integer ratio binding between neighboring time intervals.

In Desain & Honing (1989) this representation was used to construct a relaxation network that functions as a quantizer: it transforms performed musical durations on a continuous time scale to their integer-valued counterparts in the score. In Desain (1992) the consequence of such an approach for rhythmic expectancy was formalized. In Desain & Honing (1994), by interpreting expectancy as dynamic attending (Large & Jones, 1999), a beat induction model was constructed. But though tested on individual examples, and achieving promising results, the theory has not yet been validated systematically. In this article we will present evidence for the assumptions of DECO, summarized in the hypotheses mentioned above. We obtain this evidence directly from a quantization task: the identification of rhythmic patterns in a categorical rhythm perception experiment.

3. CATEGORIZATION OF MUSICAL TIME: EXPERIMENT

Listeners to music do not perceive rhythm on a continuous scale. Instead, rhythmic categories are recognized which function as a reference relative to which the deviations in timing (so-called expressive timing) can be appreciated (Clarke, 1999). It is amazing that listeners can quite easily quantize (i.e. extract discrete rhythmic categories from) a musical performance and, while memorizing and reproducing the discrete rhythmic structure, may not even be aware that there were such large deviations from mechanical timing. This seems to indicate that an autonomous process of categorical rhythm perception takes place, which maps time intervals on a continuous time scale to a finite set of duration categories (Clarke, 1999) and thus, for example, extracts a symbol sequence like [2 1 5] from a sequence of time durations (in sec) like [0.192 0.105 0.567].

We will use the results from a categorization experiment that explored systematically a space of all possible temporal patterns (including musical and unmusical ones) of four onsets and a fixed duration (Desain & Honing, in press). These were sampled on a fine temporal grid using an identification task with semi-open responses. We aim at predicting the amount of occurrence of a rhythm in the results of this identification task (i.e. how often is a certain rhythm identified), independent of the issue whether true categorical perception exists or where the categories are perceived in the stimulus space. The amount of occurrence of a rhythm is used as an indicator for stability of the rhythm percept, reasoning that a stable rhythm will occur more often because a stable rhythm percept functions as a wide "basin of attraction" (cf. Desain & Honing, 1989) and may "catch" more responses. Thus we restrict ourselves here to predicting the size of a rhythmic category in the space of responses.

This experiment revealed that some rhythmic patterns were identified considerably more often than others. As an illustration: the pattern [5 2 1] (see Fig. 1a) occurred twice as response, while the pattern [2 5 1] (see Fig. 1b) was chosen 37 times out of the 210 chosen response patterns. DECO indeed claims the latter to be more stable.

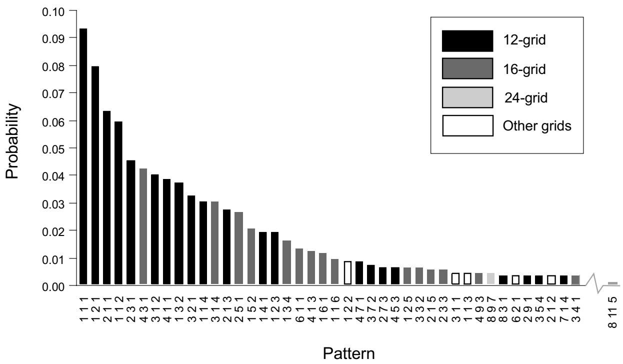

Figure 3. The distribution of the responses in the categorization experiment, expressed as the proportion of all response patterns chosen. Gray tones indicate the grid-size needed to notate the response pattern.

For all 210 different response patterns that were given in the categorization experiment, 172 could be expressed as a sequence of integers summing to a power of two and three only (see Fig. 3). However, this set attracted 97% of the response choices (all possible patterns). This means that the stimulus patterns presented were recognized to a large extent on a discrete, metrical scale: a grid that can be thought of as resulting from dividing the total duration of the stimulus recursively in two or three parts. For example, 144 responses can be represented using 96 as the smallest subdivision. These rhythmic patterns account for 95% of all response choices.

The large effect of order of a temporal pattern can be appreciated when we compare data across different permutations. The correlations of the response proportions with various permutations (excluding patterns invariant under the permutation) are on average 0.83. The largest correlation is between the patterns and their reversal (0.91), which are structurally similar, in de DECO sense.

4. RESULTS

Because a finite domain is needed to test the theory we restricted our explanations to rhythms on a 96 grid. Furthermore, we ignored rhythms which contained very small durations, because the experimental data did not allow the accurate testing of model predictions for these patterns. For each of these 2080 rhythms the stability was calculated. Correlating these predictions to the response proportions yielded a high agreement between the DECO model and the data (correlation is .73). In this test only five parameters were used: the ratio's were restricted to 1/3, 1/2, 1/1, 2/1 and

3/1, and the duration scaling was disabled. The optimal values ratio parameters reflect common intuition about their relative importance. The simplest 1/1 ratio is most salient, followed by symmetric duple ratios. The triple ratios are more complex than the duple ones, furthermore they exhibit an asymmetry: the multiplication 1 followed by 3 is more stable than the subdivision of 3 followed by 1. A more elaborate version of the model (7 parameters) was able to explain even more of the variance (correlation is .87). Figure 4 presents observed versus predicted stability for the example patterns to illustrate the prediction of the relative stability of their permutations.

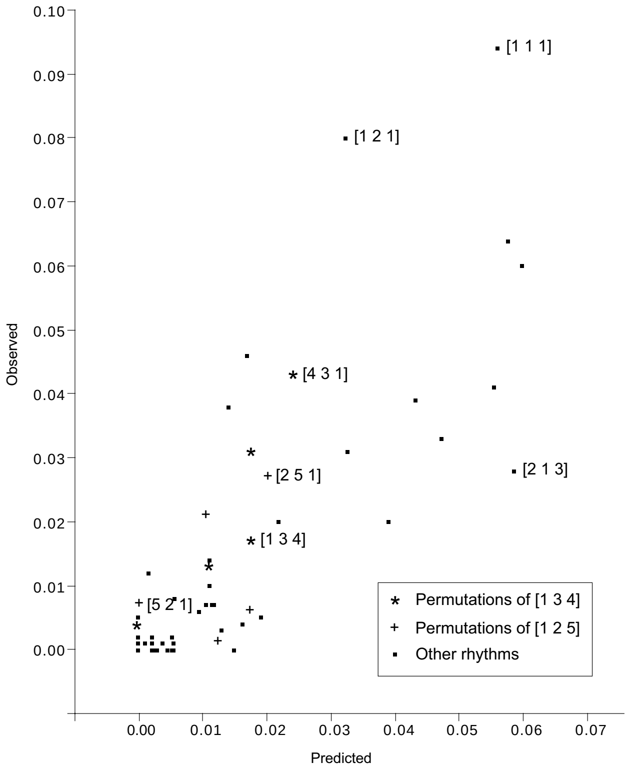

Figure 4. Observed versus predicted stability for all responses given the elaborate model.

Interestingly, for the latter model there are two clear outliers. The rhythm [1 2 1] is chosen more often than the stability measure predicts, the rhythm [2 1 3] is chosen less often than predicted. We plan to study families of patterns (e.g., those that are made up of the intervals 1, 2 and 5) to get an better understanding of how their internal structure is related to the notion of stability,

5. CONCLUSION

DECO proved successful in predicting the empirical results of a categorization experiment – predicting 75% of the variance in 2080 data points with only a few parameters. While the results reported here are still preliminary, they are a promising first step in providing evidence for the hypothesis that an essential component of a rhythm percept are the implied time intervals, and that its stability depends on integer ratio bonds between neighboring intervals. Furthermore, only very simple ratios seem to be needed for accurate predictions. Still, the main challenge is to show that this measure of stability also holds for other sets of empirical data. Furthermore, we need to improve the definition of rhythmic stability as connectivity of the time intervals making up a temporal pattern.

6. REFERENCES

1. Clarke, E.F. Rhythm and Timing in Music. In *Psychology of Music, 2nd Edition* (ed. Deutsch, D.) 473-500 (Academic Press, New York, 1999).
2. Desain, P. A (de)composable theory of rhythm perception. *Music Perception*, 9(4), 439-454, 1992.
3. Desain, P. and Honing, H. Quantization of musical time: a connectionist approach. *Computer Music Journal*, 13(3), 56-66, 1989.
4. Desain, P. and Honing, H. Advanced issues in beat induction modeling: syncopation, tempo and timing. In *Proceedings of the 1994 International Computer Music Conference*. 92-94. San Francisco: International Computer Music Association, 1994.
5. Desain, P. and Honing, H. Computational Models of Beat Induction: The Rule-Based Approach. *Journal of New Music Research*, 28(1), 29-42, 1999.
6. Desain, P. and Honing, H. Modeling the Effect of Meter in Rhythmic Categorization: Preliminary Results. *Journal of Music Perception and Cognition*. Sapporo: Japan. (In Press). [see http://www.nici.kun.nl/mmm/time.html]
7. Large, E. W. and Jones, M.R. The dynamics of attending: How people track time-varying events. *Psychological Review*, 10(1), 119-159, 1999.
8. Large, E. W. and Kolen, J. F. Resonance and the perception of musical meter. *Connection Science*. 6(2,3), 1994.
9. Longuet-Higgins, H. C. and Lee, C.S. The rhythmic interpretation of monophonic music. *Music Perception*, 1(4), 424-441 (1984).
10. Longuet-Higgins, H. C. The perception of melodies. *Nature*, 263, 646-653 (1976).
11. Nakajima, Y. A model of empty duration perception. *Perception*, 16, 485-520 (1987).
12. Palmer, C. and Krumhansl, C. L. Mental representations of musical meter. *Journal of Experimental Psychology: Human Perception and performance* 16(4), 728-741 (1990).