A model of beat induction accounting for perceptual ambiguity by continuously variable parameters

Richard Parnscott
Faculty of Music, McGill University
555 Sherbrooke West, Montreal, Quebec, Canada H3A 1E3
parscott@music.mcgill.ca

A quantitative model of rhythm perception is outlined in which all perceptual variables are real numbers rather than integers. Each variable is interpreted as an estimate of the probability that an "average" listener will notice or become aware of a given rhythmic percept or structure. In cases of multiple concurrent percepts, continuously variable parameters may reflect subtle variations in perceptual importance or salience. Continuous variables in the model include the phenomenological accent of individual sound events, the likelihood of specific foot-tapping responses, the relative strength of metrical accents, and the likelihood of the perception of specific meters (3/4, 4/4, etc.). Foot-tapping responses may be determined either by pulse or by meter saliences. The relative importance of pulse and meter for foot-tapping may depend both on musical style and on the musical sophistication of the foot-tapper.

A number of influential models of musical rhythm perception have been based on discrete variables and yes-no decisions. The models of Longuet-Higgins and Lee (1982) and Lee (1991) process a sequence note by note, continually revising hypotheses about the positions of downbeats in the light of new evidence. The model of Povel and Essens (1985) assigns phenomenal accents to the notes of the rhythm, but is restricted to integer values (1 for weak accents, 2 for strong). The music-theoretic approach of Lerdahl and Jackendoff (1983) similarly implies discrete values of accentuation, and yes-no decisions for pulse and meter.

In the present paper, a quantitative model of rhythm perception is outlined in which all perceptual variables are assumed a priori to be continuous. This feature helps the model to account in a realistic way for observed ambiguities and multiplicities in rhythmic responses. Each continuous variable in the model is initially understood as a probability – an estimate of the likelihood that an "average" listener will notice or become aware of a given rhythmic percept or structure. For example, a simple rhythm in which successive durations are unequal (e.g., 3:2:1:3:2:1, etc.) can evoke several different equally-spaced pulse trains (otherwise referred to in the literature as clocks, rhythmic levels, strata, or pulse sensations), whose perceptual saliences may be estimated by asking experimental subjects to tap out the underlying beat of the rhythm, and counting the number of responses with a given period and phase (Parnscott, 1994).

In cases of multiple concurrent percepts, continuous variables may be used to estimate the relative saliences of the percepts. For example, meter is perceived when two or more consoant pulse sensations or clocks are perceived at the same time. These individual pulse sensations are more or less salient on a continuous scale. The salience of the corresponding meter sensations also vary continuously.

ICMC Proceedings 1994 83 Foot-tapping
The phenomenal accents of the individual events of a sequence may also be regarded as continuous variables. In general, greater phenomenal accents are more likely to contribute significantly to pulse and meter sensations than are lesser phenomenal accents. In the case of syncopations, phenomenal accentuation may determine the probability that a given event or group of events will cause a listener to switch mentally from one pulse or meter to another.

Pulse saliences are most salient in the vicinity of a moderate tempo of approximately 100 events (or foot-taps) per minute, corresponding to a period of 600 ms (Fraisse, 1982). Departures from moderate tempo cause pulse salience to decrease in a gradual manner; again, a continuous variable is the most appropriate way to model this effect. In Parnscott (1994), the effect of tempo on pulse salience is accounted for by a broad, bell-shaped function that is symmetrical with respect to the logarithm of pulse period. Most of the area under the curve is contained between the values 400 and 900 ms, or 50 to 200 events per minute.

In the model, pulse responses and their estimated saliences are derived from the temporal pattern of phenomenal accents by a pattern-matching procedure. First, a template of equally-spaced events is hypothesized. Then matches are sought between elements in the template and real events. For example:

```
events  • • • • • • • • • • •
template  • • • • • • • • • • •
matches  | | | | | | |
```

The model looks at each successive pair of matches. For each pair, it calculates a contribution to pulse salience (a real number) according to a prescribed formula (Parnscott, 1994, equation 4). Finally, it adds contributions to pulse salience (above, there are only two).

Phenomenal accents are enhanced if they occur either near the start of the rhythm (primacy effect) or just before the time of observation (recency effect). Primacy and recency may best be modeled by continuously variable functions, e.g., exponentials whose decay times correspond to the duration of the psychological present – a procedure that is potentially more sensitive to subtle variations in timing and accentuation than procedures based on discrete variables (e.g., Longuet-Higgins & Lee, 1982; Rosenthal, 1992).

The relative strength of metrical accents, and the relative salience of different metrical interpretations, may be derived directly from pulse saliences by adding up the saliences of pulses that contribute to specific metrical accents or meters (Parnscott, 1987) – again, using continuously variable parameters. This procedure is superior to the method proposed by Brown (1993) in that it can determine not only the meter (whether 3/4 or 4/4, etc.) but also position of the downbeat; however, only Brown’s model accounts for thematic repetition.

Foot-tapping responses may correspond either to pulses or to meters as predicted by the model. The difference may depend on musical experience. Musicians may be more skilled at keeping track of several pulse sensations at once than are nonmusicians; if so, the responses of musicians may be determined predominantly by meter, and of nonmusicians predominately by pulse. For example, some musicians may tend to tap out the slowest pulse in a given meter (corresponding, perhaps, to a typical notated measure), making their foot-tapping responses slower than those of nonmusicians. The relative importance of pulse and meter for foot-tapping may also depend on the musical style to which the tapper is accustomed. These may be interesting topics for future research.