ICMPC 2000 Proceedings paper

From my hand to your ear: The faces of meter in performance and perception.

Caroline Palmer and Peter Q. Pfordresher, Ohio State University

Background.

Although much theoretical and experimental study in music cognition has examined the role of meter in perception, less work has examined the role of meter in music performance. On the one hand, many studies have documented that perception of meter may arise from a variety of acoustic cues, and metrically regular patterns allow more accurate perception of and memory for music (eg. Jones & Pfordresher, 1997; Povel, 1981). Such evidence has inspired theories to posit that the perception of meter arises from attention to surface-level periodicities in a sequence that generate expectancies by driving internal rhythmic oscillations (Large & Jones, 1999). Other work has suggested that meter may serve as a well-learned abstract schema that guides listeners' interpretation of strong and weak beats even in the absence of surface cues (Palmer & Krumhansl, 1990); although meter stems from the musical surface, it is not entirely dependent on surface structure.

In contrast to the perceiver's task, performers do not have to derive the meter; they know it beforehand. Furthermore, they often choose not to emphasize the meter in terms of the acoustic cues found useful for listeners. This may be because expressive nuances in performance are for the most part subtle, and metrical accents interact with many other accents in terms of performers' expressive nuances (Drake & Palmer, 1993). A performance that emphasized meter might even be considered exaggerated and unmusical. Yet evidence from many performance situations suggests that meter is far from irrelevant in performance. Pitch errors in experienced pianists' performances of well-learned music reflect the tactus or metrical level considered most important (Meyer & Palmer, submitted). The precision of performance timing, as measured by deviations in interonset intervals, suggest that some metrical levels are more directly timed than others (eg. Shaffer, 1981). Production of event sequences that match a metrical framework is often more accurate than production of sequences that do not (Povel, 1981). Finally, performance of the complex meters present in polyrhythms demonstrates that metrical complexity is an important dimension of performance (Handel, 1989).

Music-theoretic approaches to meter suggest at least two alternatives for the role of meter in musical structure: as a time-based metric, in which metrical beats are separated by equal time-units, or as an accent-based metric, in which metrical beats are distinguished by accents. Both approaches point to some regularity in the pattern of events. That regularity can be defined in terms of accent strength; meter can be described as an alternation of strong and weak accents, usually in binary or ternary alternation between strong beats (eg. Cooper & Meyer, 1960). The regularity can also be defined in terms of time spans; a beat is defined in terms of a point in time and the time elapsed between one beat and the next offers a source of regularity (eg. Lerdahl & Jackendoff, 1983). The accent-based approach assumes only an ordinal scale, which means that the downbeat is stronger than the second beat and so on. Ordinal scales for meter make the assumption that strong beats are separated by weak beats but do not rely on assumptions about the ratios of timespans between such beats. The timespan approach defines meter as a periodic alternation of strong and weak beats and incorporates an assumption of a ratio scale of events, which means that one timespan has twice the duration of another, and so on. The assumption of an ordinal versus ratio scale for meter has important implications: conclusions such as relational invariance of timing across tempo in performance follow from the ratio-scale metric, but not from the ordinal-scale metric.

Theories of perception and performance tend to differ in the scale assumptions they make. Both music-theoretic (Hasty, 1997) and psychological approaches (Jones, 1976; Large & Jones, 1999) propose that interactions between rhythms of the preceding event structure and rhythmic predispositions of the listener generate expectancies such as meter. These theories assume that the perception of rhythm incorporates an underlying ratio scale and can explain findings such as listeners' detection of timing deviations and categorization of ratio-based time intervals (Jones & Yee, 1997; Large & Jones, 1999).

However, there is less evidence that performance can be explained by invoking similar ratio-based models. For instance, the ratio-scale assumptions conflict with evidence that performers do not use ratio-based time units - timing in performance is always fluctuating. Also, some work suggests that listeners can use temporal cues other than simple ratio intervals to perceive meter (Large & Palmer, in preparation). In addition, ratio-based theories do not explain perceived similarity among the musical events that make up a performance; therefore, it is difficult for such an approach to explain memory confusions that arise in performance errors, such as the common error of substituting the correct event with one intended for a nearby location in the same musical sequence: a serial ordering error. Because of these problems, and the simplicity of the ordinal time scale inherent in the accent-unit approaches to meter, we rely in this paper on an accent-based (ordinal) approach to meter in performance.

Another possible metrical distinction to consider between perception and performance of music is the role of memory. Both perception and performance require integration of musical events over time in memory. Related work in psychology of memory suggests that behaviors as diverse as speech, categorization, and decision-making reflect temporal constraints on short-term memory. Developmental work suggests that older children show increased temporal persistence of auditory sensory memory relative to younger children, as well as increased storage of phonological (verbal) information (Gathercole, 1999). A related finding suggests that children make relatively more serial order errors than adults (Brown et al, in press). One explanation offered for these findings is that younger children's slower mental rehearsal leads to faster decay of information over time. If memory demands in general play a larger role in performance than in perception, then temporal constraints on memory may be more apparent in music performance, especially at faster tempi.

Two problems arise in the performer's memory for sequences of events: knowing what to do next (the serial order problem) and knowing when to do it (the relative timing problem). Early work in memory for sequences of words, tones, and other lists showed that when we are required to remember a sequence of items, we often remember the items but not the order in which they occurred (Gathercole, 1999). This result suggested that there is an important difference between remembering the items in a sequence, and remembering the order in which those items occurred. However, these two dimensions may not be separate in memory for hierarchically organized sequences such as music. For example, both speech errors and music performance errors tend to reflect sequence events intended for elsewhere in the sequence that arose from the same phrase rather than from a different phrase, suggesting that mistakes in serial order are not random but instead reflect hierarchical constraints on memory for the sequence (Garcia-Albea et al, 1989; Palmer & van de Sande, 1995). The question we address here is: is this scope constraint on how much of a musical sequence is accessible in memory based on ordinal or ratio-scale properties? That is, are elements within a sequence related in memory in terms of their ordinal properties (such as same or different phrase), their ratio properties (such as twice the duration or half the duration), or both?

Modeling Meter in the Serial Ordering of Performance

We describe a two-part model of how performers retrieve musical events from memory and organize them in music performance. In this model, meter serves as an accent-based ordinal grid that provides performers with a schema or general, abstract framework which, combined with temporal constraints of short-term memory, predicts the likelihoods of the correct musical events being retrieved from memory and performed. The theoretical assumption that events are related metrically on an ordinal scale is based on definitions of metrical accent strength derived from metrical grids, as used in Western tonal music theory (Lerdahl & Jackendoff, 1983) and English metrical phonology (Liberman & Prince, 1977).

Metrical grids define a sequence of musical events in terms of their accent strengths on an ordinal scale; take for instance a metrical grid for 4/4 meter, which defines 4 hierarchical levels of accent, with binary alternation of strong and weak beats at each level. The highest hierarchical level in this grid corresponds to the downbeat of each measure, which is aligned with 4 accents: one for the sixteenth-note level, one for the eighth-note level, one for the quarter-note level, and one for the half-note level. Such a grid would apply to all pieces whose time signature is 4/4, or 4 beats per measure (metrical unit); similar grids have been proposed to represent ternary meters, with two weak accents between strong accents at one level of the grid. Perceptual evidence supports the notion of metrical grids; listeners tend to infer patterns of strong and weak accents in the absence of sufficient information (Palmer & Krumhansl, 1990) and musicians tend to have more well-defined metrical hierarchies (Jones & Yee, 1997; Palmer & Krumhansl; 1990). Frames or schemas are instantiated in musical pieces by the distribution of events that fall at each accent location (Palmer & Krumhansl, 1990; Palmer, 1996). Thus, sufficient cues across many musical pieces instantiate for the musician the metrical frame through exposure to the musical style.

The advantage of applying metrical grids to performance is that they offer a bidirectional (past and future) framework that can guide the planning of sequence events. Consider as evidence the serial order errors that arise in music performance. Although timing has been the focus of most performance studies, performance errors (breakdowns that result in unintended events) provide more information on memory and planning constraints. Figure 1 shows a typical distribution of pitch errors from different performances of a Bach prelude in D-Major, in 4/4 meter; the histogram of error frequencies is shown in terms of the distance of the error (unintended pitch) from where it was intended to have been performed (Meyer & Palmer, submitted). Two patterns are evidenced here (see also Palmer & van de Sande, 1995, for similar error distributions with different musical pieces); the first is that errors are more likely to reflect events from nearby distances than from faraway distances; the distribution is symmetrically distributed around the present event (distance 0), and approaches a normal distribution. The second pattern is that the histogram shows periodic peaks at binary distances from the present event. These peaks suggest a periodic rise in the accessibility of that information to the performers. These events at binary distances (multiples of 2) can be more similar to each other in metrical accent for this piece in 4/4 meter than events at other distances.

Metrical grids can be used to predict similarity among sequence events that have similar accent strengths (ie, are aligned with the same metrical level in the grid). Within each level in a metrical grid, events that are nonsequential have greatest similarity. As a result, events from farther away are more similar to each other metrically than events close by. This prediction is attractive because similarity-based interference, a common cause of serial ordering errors in performance, often arises among nonadjacent events (Palmer & van de Sande, 1995). Figure 2 shows an example of an abstract representation of metrical accent strength in 4/4 meter, in which repeats throughout the music in a cyclic fashion; 0 degrees defines the onset of each metrical bar (the downbeat), aligned with the greatest number of accent levels in the grid. Thus, the number of total event locations in one cycle of the grid, *n*, is determined here by the number of divisions from the lowest level to the highest level. The metrical accent strength of each event, *m*, is represented by the length of each vector, which is equal to the number of metrical levels in the grid with which that event coincides. Note that the accents are not temporally defined; the grid can stretch or shrink to fit the tempo of the sequence. Time is defined in terms of a serial (proximal) component of the model, described later.

The first component of the model, metrical similarity ($\mathbf{M_x}$), defines the similarity in metrical accent strength between sequence events. The absolute difference in metrical strength between an event at position *i* and another event at distance *x* (position *i+x)* is computed and divided by the sum of the metrical accents for the two events. That difference is subtracted from 1 to form a similarity metric, as follows:

Equation 1:

$$\mathrm{Sim}\,(m_{(i)}, m_{(i+x)}) = 1 - \frac{|m_{(i)} - m_{(i+x)}|}{m_{(i)} + m_{(i+x)}} = 1 - \frac{\Delta m}{2\bar{\bar{m}}}$$

The righthand side of Equation 1 reflects the fact that this function for metrical similarity is a form of Weber's law. This is psychological appealing because it captures the perceptual analogue that listeners are more sensitive to a given accent difference among a pair of events when presented in a context of low-intensity accents, than in a context of high-intensity accents.

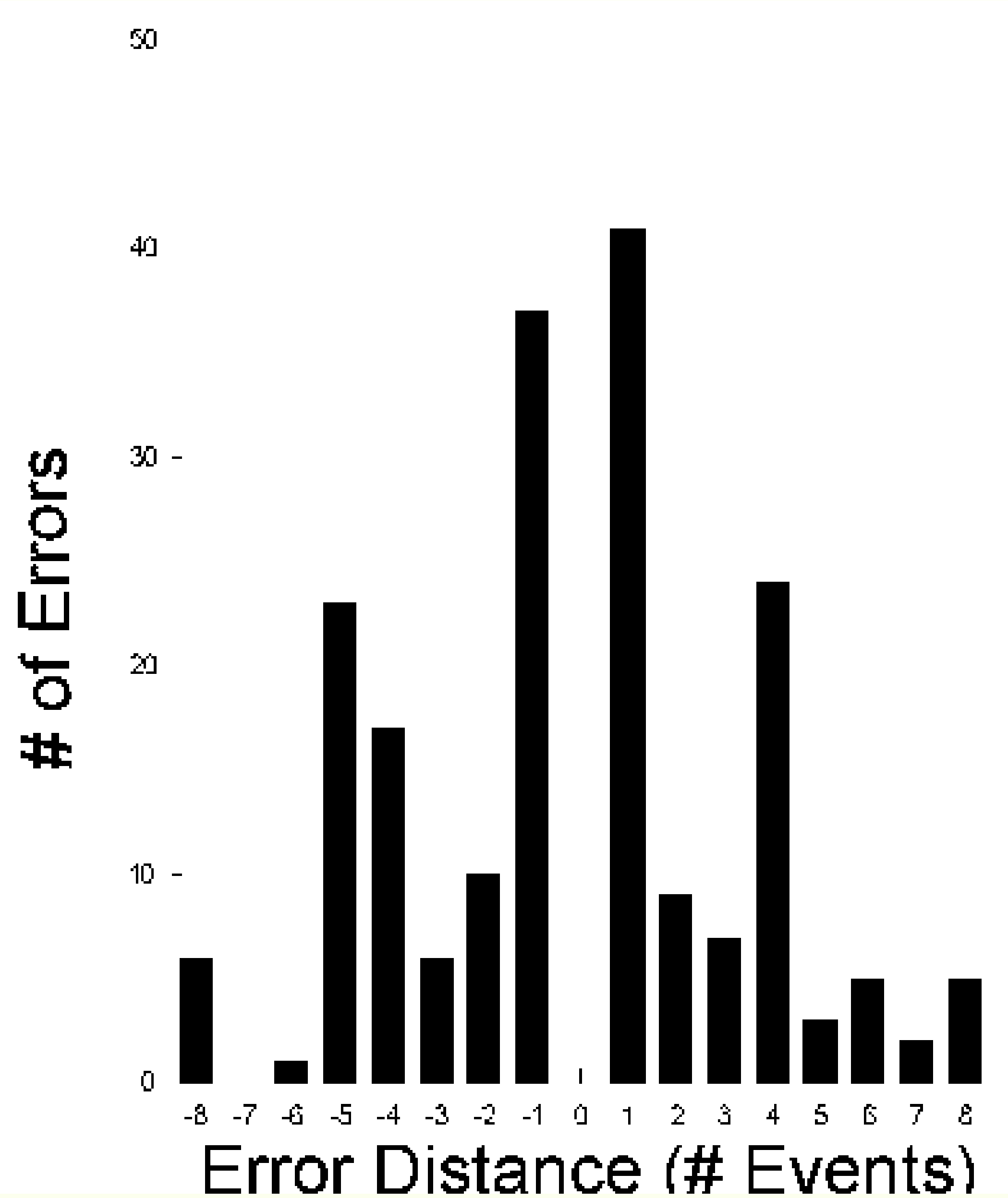

Figure 1: Frequency histogram of the number of serial order (pitch) errors by event distance, in performances of Bach Prelude in D-Major (from Meyer & Palmer, in preparation).

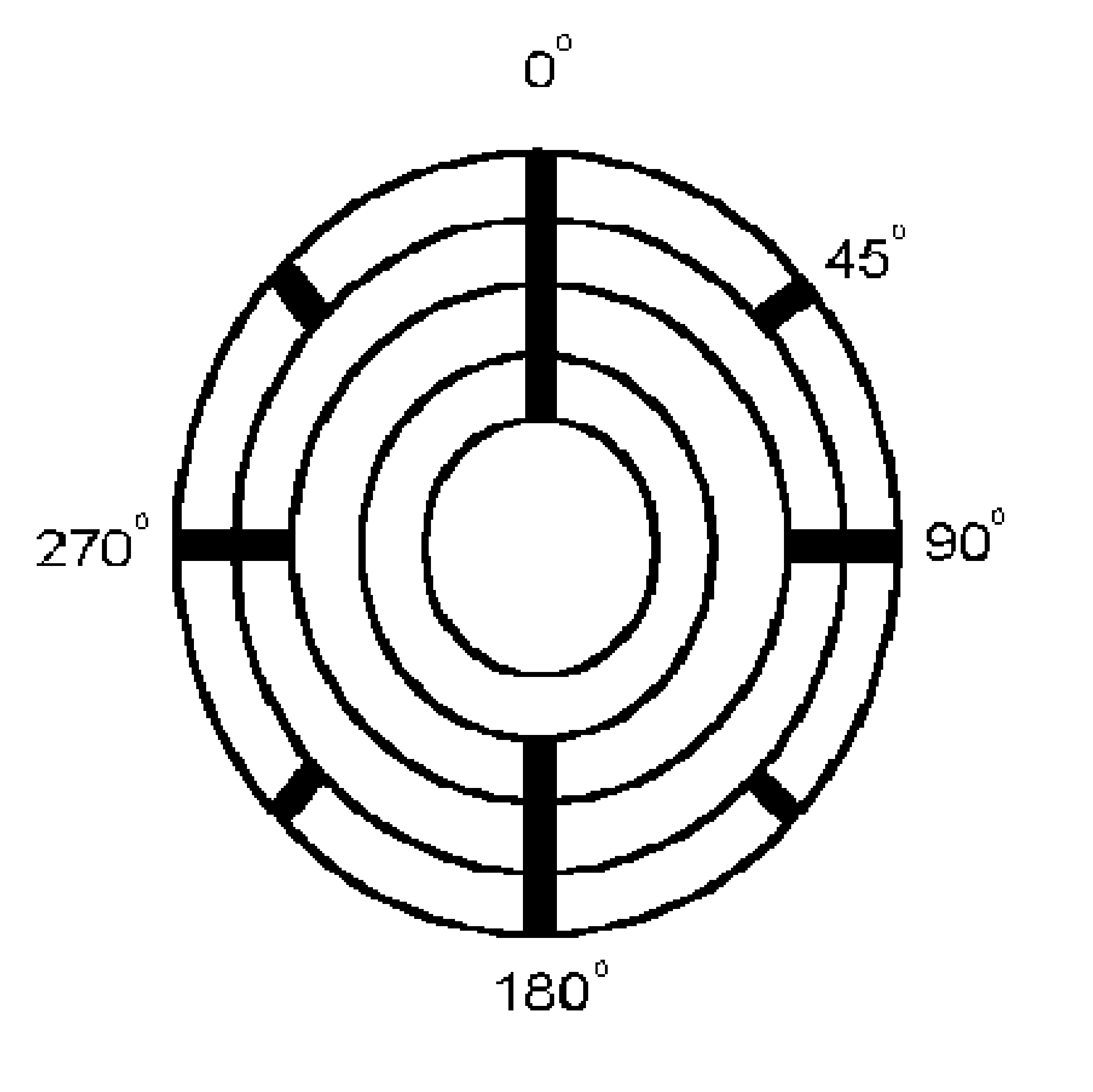

Figure 2: Circular representation of metrical accent strength for 4-tier metrical grid. Metrical strength of each event is represented by the length of each vector; concentric circles represent each level in the grid.

Similar contrast functions have been used in vision to model the detection of luminance differences (Michaelson, 1906). This metrical similarity measure is summed across all positions in a sequence and divided by the total number of positions *n*, to generate the vector of metrical similarity values across distances from the current event, $\mathbf{M}_{\mathbf{x}}$, as follows:

Equation 2:

$$M_x = 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{|m(i) - m(i+x)|}{m(i) + m(i+x)}$$

The second component of the model, serial proximity ($\mathbf{S}_{\mathbf{x}}$), captures the fact that memory for sequence events is less accessible the farther away they are from the performer's present position in the sequence. Event strength (**S_x**) is assumed to be maximal at the current position and equal to 1; event strength for other sequence events decreases both with absolute event distance (x) from the current event and as event duration (*t*, defined here as seconds per event) increases in the following nonlinear relationship:

Equation 3:

$$S_x = a^{(x/t)}$$

This function takes an initial activation *a*, a value between 0 (no activation) and 1 (total activation) that represents temporal constraints on short-term memory. The exponent (x / t) refers to number of events per unit time (similar to beats per minute in musical terms) and leads to two predictions. First, the larger *x* is (the farther away an event from the current event), the weaker the event strength. Second, the smaller *t* is (as tempo gets faster), the weaker the event strength. Thus, the serial component represents a proximity-based combination of decay (over elapsed time) and interference (over intervening events). Sequence events from the future and the past will decay faster as the number of intervening events increases and as the rate increases.

The model makes a basic assumption common to many formal models of memory, that sequence elements can be represented as vectors of relations among elements. The metrical similarity and event strength of each sequence element at each distance from the current event are represented in **M** and **S** vectors, respectively. Position within the vector represents comparisons among sequence events at different positions and distances; the vector size is equivalent to one metrical cycle.

Finally, the two components of the model are combined in a multiplicative fashion to predict relative event strength or activation for any event x at time t as the product of metrical similarity and serial activation (**$S_x \cdot M_x$**) function. The relative activations of sequence elements at each distance x from the current event are then normalized to determine relative error probabilities for each sequence event. The error probabilities for each event distance from the present event reflect the fact that sequence events from greater distances have greater event strength in some cases than sequence events from smaller distances. This is psychologically appealing because it reflects Garrett's (1980) caveat that although speech errors often reflect access to sequence events from some distance from the error location, it does not follow that a speaker has access to all intervening events. This model is the first to make specific predictions for which elements are more or less accessible from various sequence distances.

The model makes a further prediction for the absolute mean distance between any serial order error and its target pitch. The mean range is computed as the weighted sum of the error probabilities at each sequence distance multiplied by each sequence distance. This is shown in Equation 4.

Equation 4:

$$\text{Mean Range} = \sum_{x=1}^{n} norm(S_x \cdot M_x) \cdot x$$

A final prediction of the model is that for any two tempos t1 and t2, such that t1 is less than (faster than) t2, the mean absolute range for t1 will be smaller than the mean absolute range for t2. This follows from the predictions of the product model (combination of metrical and serial components): the serial component decreases activation of sequence elements from farther distances faster for t1 than for t2, in essence damping the effect of metrical similarity for events form larger distances. This fact, combined with the fact that sequence events at closer distances are more accessible at fast than at slow tempi, and sequence events at farther distances are more accessible at slow than at fast tempi, account for the general prediction that events will be accessible from greater sequence distances on average at slower performance tempi than at faster tempi. .

Palmer, Pfordresher and Brink (in preparation) tested the model's predictions in two experiments. Pianists performed simple musical excerpts during which both practice and production rate (tempo) were varied to test the predictions of the model. Increased practice and slower rate both led to fewer errors. Performances at slower tempi generated a larger range of planning, with sequence elements arising in errors from greater distances. Furthermore, more errors reflected nearby elements when the music was performed at a faster tempo, and more errors reflected distant elements when performed at a slower tempo. In a second experiment, novice child pianists performed the same task. The performances showed relatively more serial order errors, consistent with psychological theories of short-term memory processes that predict faster decay of information for children than adults (Brown et al, in press). Because the initial activation parameter of the serial component predicts faster decay of information in memory for children than for adults, the novice performances showed less contribution of metrical frames to range of planning than the expert performances (these findings are described further in Palmer, Pfordresher and Brink, in preparation).

Implications:

We have presented a model of a metrical framework based on accent similarity that guides the retrieval and organization of musical events during performance; the metrical component of this model highlights some features that may be unique to performance. First, the metrical similarity component of the model predicts a symmetrical influence of past and future events relative to the present. This feature may be specific to performance because memory for past and future events may be simultaneously available and the weighting of memory may be symmetrical. Perceptual tasks in contrast may reflect lower memory load but the burden of generating expectations for upcoming events from past events may force listeners to have asymmetrical influences of past and future on processing of current events. Second, the reliance on metrical grids requires only an ordinal-scale assumption about metrical similarity, which is sufficient to generate predictions about memory retrieval in music performance.

The proximity-based decay component of the model specifies how temporal constraints of short-term memory can influence serial order of events in performance. The metrical and serial components interact to moderate the influence of meter on a performer's scope of planning at different tempi. These psychological consequences of time in memory for musical sequences may also extend to other aspects of music performance, such as those related to tempo effects on musical interpretation and relational invariance of motor programs.

Are principles such as metrical similarity and temporal proximity common across music perception and performance? Strangely enough, similarity and proximity principles are more commonly found in perceptual theories but rarer in performance theories. The model of metrical similarity described here is simple because it has relatively few parameters: production rate, metrical grid size, and initial activation strength. The first is established by the experimental conditions; the second is schematic (general and abstract) and acquired through exposure, and the third reflects temporal constraints on short-term memory and is posited to increase with age. In principle, none of these parameters need be specific to music: even the metrical schemas, which might be most specific to musical styles and periods, resemble in fundamental ways those proposed for language (Hayes, 1984), and perception of their component periodicities has been modeled with dynamical systems (Large & Jones, 1999).

The model's simplicity also gives rise to its limitations. Perhaps most important is its adherence to only one dimension of similarity among musical events: that of meter. Research in music performance has documented other musical dimensions that influence similarity judgments or confusion errors, including melodic contour, tonality, harmony, rhythm, and timbre (cf. Palmer & van de Sande, 1993, 1995). Another limitation is the model's inability to explain how the metrical grid is learned. Statistical analyses of frequency distributions of note events across metrical positions document the common compositional technique of establishing a meter by putting more notes in positions of metrical strength (Palmer 1996; Palmer & Krumhansl, 1990), but how these are acquired in memory is unsolved. More recently, dynamical systems models have been posited that track events over time and generate expectancies for when events will occur, based on prior sequence structure (Large & Jones, 1999; Large & Palmer, in preparation). Oscillators with adjustable period and phase components may respond to periodicities represented at each level in a metrical grid, offering an explanation of how metrical frameworks are acquired.

Musical meter provides an important testing ground for comparing the role of temporal sequence structure in perceptual tasks (such as beat tracking) and motor tasks (such as music performance). Comparisons between these domains may lead to the identification of different constraints on attention and memory processes, as well as some similarities. For example, the psychological constraints in the metrical similarity model described here reflect general principles that can be tested in perceptual tasks, a necessary step in bridging the gap from the pianists' hand to the listeners' ear.

Acknowledgements

This research was sponsored in part by NIMH grant R01-45764 to the first author. Reprint requests should be addressed to Caroline Palmer, Psychology Dept., Ohio State University, 1885 Neil Ave., Columbus Ohio 43210, USA, or to palmer.1@osu.edu.

References

Brown, G.D.A., Vousden, J.I., McCormack & Hulme, C. (in press) The development of memory for serial order: a temporal-contextual distinctiveness model. International Journal of Psychology.

Cooper, G., & Meyer, L.B. (1960). The rhythmic structure of music. Chicago: University of Chicago Press.

Drake, C. & Palmer, C. (1993) Accent structures in music performance. Music Perception, 10, 343-378.

Garcia-Albea, J.E., de Visa, S. & Igoa, J.M. (1989) Movement errors and levels of processing in sentence production. Journal of Psycholinguistic Research, 18, 145-161.

Garrett, M.F. (1980) Levels of processing in sentence production. In B. Butterworth (ed,), Language production: Speech and talk (pp. 177-220) London: Academic Press.

Gathercole, S.A. (1999). Cognitive approaches to the development of short-term memory. Trends in Cognitive Science, 3, 410-418.

Handel, S. (1989). Listening: an introduction to the perception of auditory events. Cambridge: MIT Press.

Hasty, C.F. (1997) Meter as rhythm. New York : Oxford University Press.

Hayes, B. (1984). The phonology of rhythm in English. Linguistic Inquiry, 15, 33-74.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. Music Perception, 13, 489-516.

Jones, M.R. & Pfordresher, P.Q. (1997). Tracking musical patterns using joint accent structure. Canadian Journal of Experimental Psychology, 51, 271-290.

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.

Large, E.W., & Jones, M.R. (1999) The dynamics of attending: How people track time-varying events. Psychological Review, 106, 119-159.

Large, E.W,, & Palmer, C. (In preparation). Temporal response to music performance: Perceiving structure in temporal fluctuations.

Lerdahl, F., & Jackendoff, R. (1983) A generative theory of tonal music. Cambridge: MIT Press.

Liberman, M.Y., & Prince, A. (1977) On stress and linguistic rhythm. Linguistic Inquiry, 8, 249-336.

Meyer, R.K. & Palmer, C. (submitted). Temporal control and planning in music performance.

Palmer, C. (1996) Anatomy of a performance: Sources of musical expression. Music Perception, 13, 433-454.

Palmer, C., & Krumhansl, C.L. (1990) Mental representations of musical meter. Journal of Experimental Psychology: Human Perception and Performance, 16, 72-741.

Palmer, C., Pfordresher, P.Q., & Brink, D. (in preparation). Music errors, speech errors, and cognitive constraints on sequence production.

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 457-470.

Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. Journal of Experimental Psychology: Human Perception and Performance, 21, 947-962.

Povel, D.J. (1981) Internal representation of simple temporal patterns. Journal of Experimental Psychology: Human Perception and Performance, 7, 3-18.

Shaffer, L.H. (1981) Performances of Chopin, Bach, and Bartok: Studies in motor programming. Cognitive Psychology, 13, 326-376.

Sloboda, J.A. (1983). The communication of musical metre in piano performance. Quarterly Journal of Experimental Psychology, 35, 377-396.

Thomasson, J.M. (1982) Melodic accent: Experiments and a tentative model. Journal of the Acoustical Society of America, 71, 1596-1605.

Back to index