UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Structural Representations of Music Performance

Permalink
https://escholarship.org/uc/item/2dc0124m

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 11(0)

Author
Palmer, Caroline

Publication Date
1989

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Structural Representations of Music Performance

CAROLINE PALMER

PSYCHOLOGY DEPARTMENT, OHIO STATE UNIVERSITY

ABSTRACT

A primary goal of music cognition is to understand mental representations for musical knowledge that allow communication of thoughts and emotions. Theories of musical competence generally model mental representations in terms of structure given in the musical text, and do not model performers' preferential choices of structural content for emphasis. Such choices are an important component of musical interpretation. Two sources of converging evidence are described that support the role of phrases as structures in mental representations for music performance: evidence from expressive timing in skilled performance and from performance breakdowns (errors). The location and amount of expressive timing, and the likelihoods of different error types coincided with musicians' notated interpretations. Evidence from both ideal and non-ideal musical behavior implicate the same structures in representation of musical knowledge, and suggest that individual preferences can explain much variation in music performance.

Introduction

Many theories of music cognition attempt to specify mental representations for musical knowledge; that is, the knowledge that performer, listener, and composer have, that allows communication of thoughts and emotions. The topic of this paper, music performance, provides a rich area in which to study the interaction between representations of musical structure and skilled performance in a naturalistic setting. Mental representations underlying music performance should affect our perception and comprehension of music as well.

Several theories tackle the problem of modelling our musical competence (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Meyer, 1973). One of the most explicit attempts to formalize musical competence is a recent theory that models experienced listeners' knowledge of Western tonal music (Lerdahl & Jackendoff, 1983). The theory predicts the relative importance of musical events, based on a combination of well-formedness rules that specify the set of possible structural descriptions, and preference rules, that resolve among the competing alternative structures. This theory represents a common approach in music cognition: to model structural aspects in terms of information in the musical text, usually deriving one set of structures per musical excerpt.

This ambitious undertaking has met with some success in predicting effects in music perception (Deliege, 1987; Palmer & Krumhansl, 1987). There are, however, several problems in its application to music performance (shared by all current theories of musical competence): the model is not algorithmic (the particular rules and order in which they are applied are not specified), and it is not clear how preferences resolve among alternative interpretations in many ambiguous musical passages. Of most concern for this paper, the theoretical predictions do not take into account the different expressive variations introduced in different performances of the same musical text. These problems illustrate the tasks faced by any theory of musical competence - how to choose among possible structures composed of different constituents, and how to apply procedures to combine or contrast them in their relative importance.

This paper addresses how the musical structures and processes in a mental representation influence music performance. The specific questions include: how does a musical performance communicate some intended meaning or structural content? How do abstract intentions differentiate one performance from another? What kinds of structures and operations comprise the mapping of abstract musical intentions to sounded performance? I will refer to musical intentions as interpretations. Interpretations are the musicians' modelling of a piece according to their own choices of appropriate musical structure for emphasis, such as melody, phrasing, and dynamics (Apel, 1972). A skilled performer plays not only what is specified in the musical text, but adds much intentional expressiveness that does not appear in the text. Small variations in frequency, timing, intensity, and timbre are thought to govern the assignment of musical meaning to expressive performance (Nakamura, 1987; Palmer, 1989; Shaffer, Clarke, & Todd, 1985; Todd, 1985).

Two sources of evidence are presented that support musical phrases as structures common to mental representations underlying music performance. Phrases are a good candidate for musical structure; they are typically described in music as a unit of meaning, often defined by elements at their boundaries (Cogan & Escot, 1976). Pausing and increasing durations at phrase boundaries in both music (Todd, 1985) and speech (Cooper & Paccia-Cooper, 1980) are assumed to be determined by syntactic and prosodic structure. Algorithms have been developed that predict pausing in speech and music by an amount proportional to the hierarchical level or depth of phrase embedding (Grosjean, Grosjean & Lane, 1979; Todd, 1985). Note that these algorithms predict tempo changes from the information in the text, not by the individual performer's intentions, and thus result in one prediction per musical piece.

I will describe two effects of individual performers' interpretations on mental representations for music performance. The two sources of evidence are: from timing variations in skilled performance (ideal performance) and from performance breakdowns (errors). A set of piano performances were collected, along with the phrasing interpretations of the pianists. All the performances described here were recorded on a computer-monitored Bosendorfer concert grand piano, containing optical sensors and solenoids. This instrument allows precise measurement, recording, and playback of timing and velocity (loudness) parameters of a performance, without affecting the touch or sound of the acoustic instrument.

Expressive Timing in Music Performance

I will first describe a study that addresses the effect of phrasing interpretation on musical performance by comparing each performer's intended structural content with expressive timing variations in the sounded performance. Six professional pianists were asked to perform the same musical piece, a Piano Prelude by Chopin (the musical text was held constant). Afterwards, pianists notated their phrasing interpretations on an unedited musical score. Figure 1a contains a 5-measure excerpt from that piece, with one of the pianist's notated phrasings above the musical score.

A common method for emphasizing phrasing in music performance is to alter the relative timing of events by using rubato (changes in tempo). The experimental goal is to find correspondences between the intended phrasing and the use of rubato. The timing variations in each recorded performance were therefore analyzed relative to each pianists' phrasing interpretation. Shown in Figure 1b are the rubato patterns from one performance, expressed in percent deviation from a mechanically regular performance. The deviations were calculated relative to each pianist's mean tempo or rate; thus, a large positive value indicates slowing down, and a negative value indicates speeding up. The correspondence between amount of rubato and intended phrasing is most evident at phrase boundaries (indicated by gaps in Figure 1b); phrase endings are accompanied by large positive values (indicating slowing down), relative to beginnings of phrases.

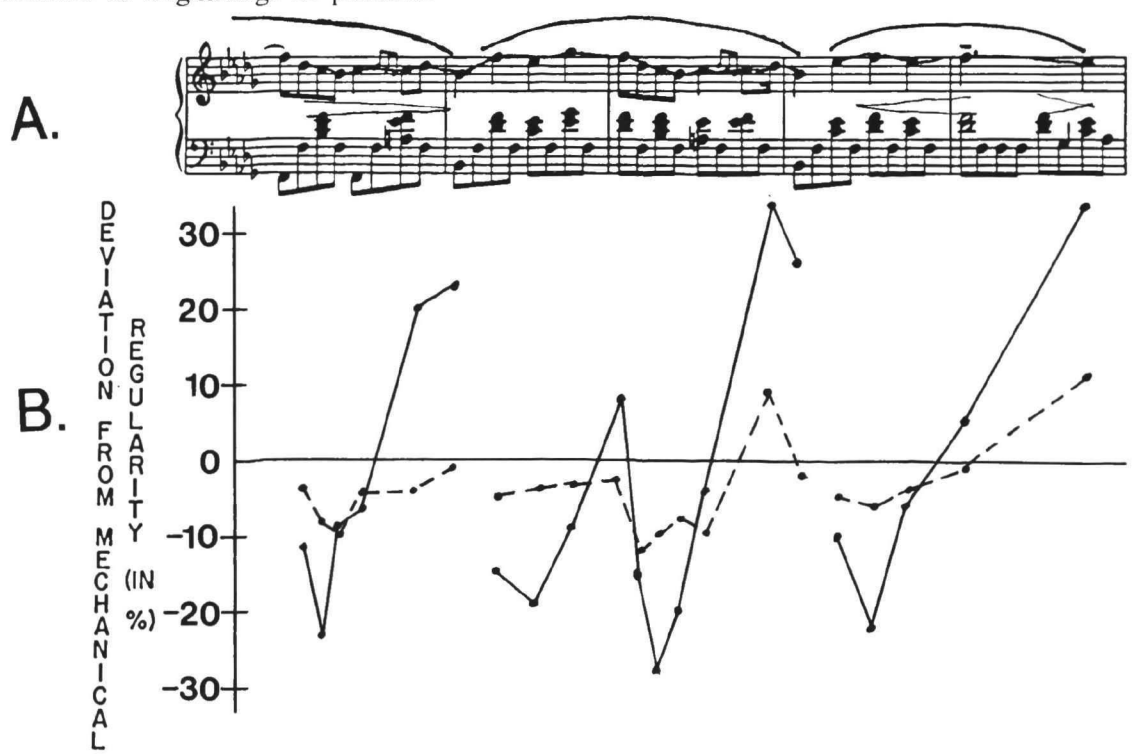

Figure 1. A pianist's performance of the Prelude in D-flat Major by Chopin. a) Notated phrasing of performer. b) Temporal deviations from mechanical regularity of the score.

After performing the musical excerpt, each pianist was asked to play it again, except this time, to play it mechanically, not to add any expressive interpretation to the musical text (the line around zero in Figure 1b reflects a strictly regular performance). The dashed line in Figure 1 reflects the pianist's "mechanical" performance; the relationship between rubato and intended phrasing disappears in the mechanical performance, suggesting that the use of rubato is both voluntary and part of the performer's intention to emphasize phrasing.

I will turn to a study of phrasing in a more ambiguous musical example. Pianists consistently notated the same phrasing for the Chopin excerpt described above. In contrast, they differed often in choice of phrasing for a Brahms Piano Intermezzo, shown in Figure 2 with two phrasing interpretations (notated by different pianists). To study the uniqueness of the mapping from phrasing interpretation to timing variations, one pianist was asked to play the same music twice, with the two different phrasing interpretations. Each performance was analyzed in terms of the different interpretations.

The graphical notation shown in Figure 3 demonstrates the expressive timing variations in the two performances. This nontraditional notation offers the advantage of increased resolution of temporal and intensity (loudness) information. Pitch height is represented on the abscissa, time on the ordinate axis; black denotes loud musical events, white denotes quiet. The two boxes display the first six notes in the melody (in this excerpt, the highest pitches) for each performance. The intended phrasing is marked with lines. The gaps coincident with intended phrase boundaries indicate an increase in expressive timing (pauses occur at phrase boundaries). As shown here, the mapping of interpretation to performance is unique. When a performance is examined in terms of an interpretation other than that intended by the performer, the timing variations no longer coincide; only the performer's intended phrasing characterizes the timing accurately.

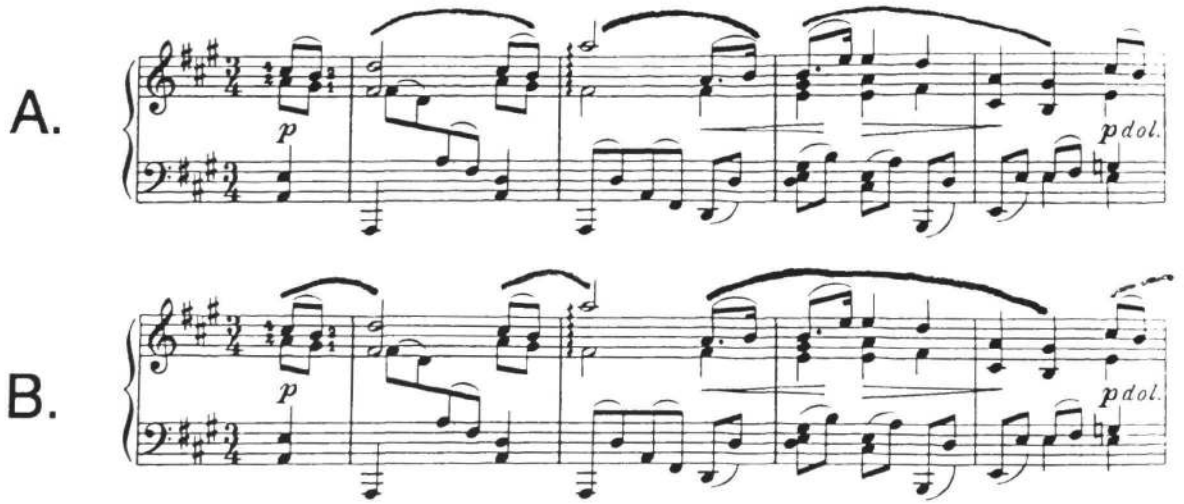

Figure 2. Two phrasing interpretations (A and B)
of opening section of Piano Intermezzo, Opus 118 No. 2, by Brahms.

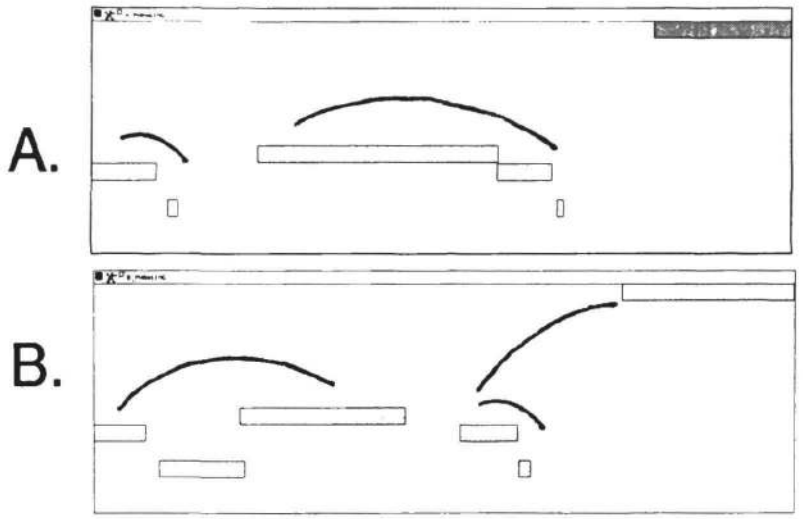

Figure 3. Graphical notation of performances of the two phrasing interpretations (A and B) of opening section from Brahms' Piano Intermezzo.

Thus, the use of rubato to emphasize phrasing accommodates different interpretations of the same music and disappears in the absence of interpretation (as seen in the mechanical performance). These findings support a framework of mental representations for music performance in which expressive timing is directly related to the performers' choice of phrases as constituent structures. However, phrases may serve as constituents at a global level, realized only in performance output; they may be combined from other, more primitive units of representation. A second source of evidence for phrases as basic units of mental representation is described, that argues against this conclusion: evidence from performance breakdowns, or errors.

Performance Errors

Performance breakdowns, or errors resulting in unintended output, are fairly common in many behaviors and often result from conflict among several possible actions, thoughts, or plans. From an analysis of the types and amounts of errors, it is possible to construct a theory that specifies what kinds of structures compete, and what kinds of processes operate on these structures in mental representations for music. If phrases are a basic constituent in mental representations, then performance errors should reflect processes operating on phrase structures. I will describe performance errors and how they coincide with phrasing interpretations. Eight different performances of the same Brahms Piano Intermezzo with ambiguous phrasing structure (shown in Figure 2) are described, including four performances of each phrasing interpretation. The piano performances were recorded on the same computer-monitored musical instrument. Because the computer detects wrong pitches and rhythms, we do not have to rely on a listener's capabilities. This allows us to avoid a problem encountered with speech errors; because errors frequently result from competing items that are similar in class, form, or sound (Dell, 1986; Garrett, 1975), they often go unnoticed by a listener.

Errors in piano performance can involve elements of pitch, duration, or both. I will describe errors restricted to pitch elements, and how they coincide with the constituent structures indicated by each performer's phrasing interpretations. I will concentrate on three error types: deletion errors, perseveration errors, and substitution errors. Deletion errors occur when an intended musical event is dropped or missing in the performance. Perseveration errors occur when an intended musical event is inappropriately repeated at a later time in the performance. Substitution errors occur when an intended musical event is replaced or substituted by an unintended event.

The frequency of occurrence of each error type differed significantly, relative to the performers' intended phrasing. The first error type, deletion errors, was more likely to occur within (71%) than between (29%) phrase boundaries. This finding is similar to phonemic slippage effects in speech, in which a phoneme is deleted more often within a word than at word boundaries. Deletions are found most commonly in various recall tasks for unimportant events, presumably because activation or relative strength of unimportant events is less than that of important events. In this musical context, those events within a phrase appeared to be most susceptible to deletion; this finding coincides with explanations of phrases defined by elements at their boundaries (Cogan & Escot, 1976).

The second error type, perseveration errors, was more likely to occur between (100%), than within (0%), phrase boundaries. Notice that this is in contrast to the frequencies of the deletion errors; both error types are affected by intended phrasing, in alternate ways. Spreading activation models predict that perseveration errors result from the activation or strength of intended elements being lower than that of unintended (perseverating) elements (Dell, 1986). The more strongly activated elements dominate the intended elements. Thus, within phrase boundaries, the intended elements are strong, while between phrases, unintended elements are more likely to intrude.

The third error type, substitution errors, did not coincide with musical phrase locations (they were just as likely to occur within as between phrases). Instead, substitution errors were related to type of unit: chords, or simultaneous note events, were substituted for each other more often (100%) than were elements within chords (0%). This finding suggests that errors are not just a result of attentional biases toward phrase boundaries, because not all error types are affected by positional similarity (that is, similar phrase locations). Instead, substitutions are more likely to involve elements of similar form.

Thus, while the substitution errors reflect processes operating on type of unit or form, the deletion and perseveration errors reflect processes operating on positional information (reflected in frequency of occurrence of errors at similar locations), determined by the performers' phrasing interpretations.

Conclusions

Both the use of expressive timing and the likelihood of various performance errors support the notion of phrases as structures in mental representations for music. The timing

measurements from skilled piano performances demonstrated expressive variations coinciding with intended phrase structures (as determined by each performer). The flexibility of expressive timing was emphasized by the accommodation of different phrasing interpretations in the same piece, and by reduction of expressive timing when pianists attempted to play without phrasing interpretation (mechanically). The analysis of performance errors also implicated phrases as units in music representation. Susceptibility to error was directly related to positional information specified by the performers' intended phrasing. These two sources of evidence converge on the same mental structures determined by performers' choices of musical constituents, not by the musical text.

Ideal and non-ideal (errorful) performances may tap different cognitive resources, such as those hypothesized in a distinction between competence and performance. Thus, the evidence from expressive timing may represent "idealized" competence and the evidence from breakdowns "other performance variables". However, both sources of evidence were derived from highly skilled, well-learned performances under natural conditions. And both cases provide evidence that intentions to emphasize structural content directly affect musical behavior. Both the ideal and non-ideal musical behavior implicate the same structures in mental representations underlying music performance.

Instead, differences between well-formedness processes (constraints on mental structures) and preferences (probabilistic influences that choose among the lawful alternatives) may govern the formation of mental representations for music. Much of expressive performance can be explained by the performer's individual preferences, rather than by a generalized set of rules deriving musical structure from a given text. Thus, preference rules may play a larger role in performance than do well-formedness rules (which are already constrained by the composer). This implies that preferences play a larger role in comprehension of (sounded) performances than they do in comprehension of (notated) musical text: the performer's preferences can determine the likelihood with which the listener assigns a particular structural interpretation to a performance, reinforcing the communication of musical thought.

ACKNOWLEDGEMENTS

I would like to thank Mari Jones and Jordan Pollack for their comments, and the members of the Music Cognition Group at the M.I.T. Media Lab, where some of this research was conducted.

REFERENCES

Apel, W. (1972). Harvard dictionary of music (2nd ed.). Cambridge, MA: Harvard Univ. Press.

Cogan, R.D., & Escot, P. (1976). Sonic design: The nature of sound and music. Englewood Cliffs: Prentice-Hall.

Cooper, W.E., & Paccia-Cooper, J. (1980). Syntax and speech. Cambridge, MA: Harvard University Press.

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

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. Psychological Review, 93, 281-321.

Deliege, R. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. Music Perception, 4, 325-360.

Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed), The psychology of learning and motivation (pp. 133-177). NY: Academic Press.

Grosjean, F., Grosjean, L., & Lane H. (1979). The patterns of silence: Performance structures in sentence production. Cognitive Psychology, 11, 58-81.

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

Meyer, L.B. (1973). Explaining music. Berkeley: University of California Press.

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. Perception & Psychophysics, 41, 525-553.

Palmer, C. & Krumhansl, C.L. (1987). Temporal and pitch structures in determination of musical phrases. Journal of Experimental Psychology: Human Perception & Performance, 13, 116-126.

Palmer, C. (1989). Mapping musical thought to musical performance. Journal of Experimental Psychology: Human Perception & Performance, 15, 301-315.

Shaffer, L.H., Clarke, E.F., & Todd, N.P. (1985). Metre and rhythm in piano playing. Cognition, 20, 612-77.

Todd, N.P. (1985). A model of expressive timing in tonal music. Music Perception, 3, 33-59.