Similarity, Invariance, and Musical Variation

STEPHEN McADAMS AND DANIEL MATZKIN

IRCAM-CNRS, F-75004 Paris, France

ABSTRACT: Perceptual similarity underlies a number of important psychological properties of musical materials, including perceptual invariance under transformation, categorization, recognition, and the sense of familiarity. Mental processes involved in the perception of musical similarity may be an integral part of the functional logic of music composition and thus underly important aspects of musical experience. How much and in what ways can musical materials be varied and still be considered as perceptually related or as belonging to the same category? The notions of musical material, musical variation, perceptual similarity and invariance, and form-bearing dimensions are considered in this light. Recent work on similarity perception has demonstrated that the transformation space for a given musical material is limited by several factors ranging from degree of match of the values of auditory attributes of the events composing the sequences to their relations of various levels of abstraction and to the degree that the transformation respects the grammar of the musical system within which the material was composed. These notions and results are considered in the light of future directions of research, particularly concerning the role of similarity and invariance in the understanding of musical form during listening.

KEYWORDS: Similarity in music; Invariance in music; Musical variation; Perception of music

INTRODUCTION

Perceptual similarity underlies a number of important psychological properties of musical materials, including perceptual invariance under transformation, categorization, recognition, and the sense of familiarity. Mental processes involved in the perception of musical similarity may be an integral part of the functional logic of music composition and may thus underly important aspects of musical experience. We are interested in the perception of musical materials and of their musical transformations—that is, "themes" and "variations," but in a much larger sense than is usually attached to this specific form in Western tonal music.

A longer-term goal of this approach is to understand the role of the perception of similarity and invariance, as well as that of the change produced by variations of the original materials in the experience of musical form. We hesitate to say "perception" of musical form since most musical forms are too long to be "perceived" as such. By understanding the mental processing underlying similarity perception, it may be possible to explain in part how listeners form hierarchical and associative relations

Address for correspondence: Dr. Stephen McAdams, IRCAM-CNRS, 1 place Igor Stravinsky, F-75004 Paris, France. Voice: +331-4478-4838; fax: +331-4478-1540.
mail: smc@ircam.fr

among features of a musical piece belonging to a given musical system and thereby explain in some musical situations how they manage or fail to establish these relations. How much and in what ways can musical materials be varied and still be considered as perceptually related or as belonging to the same category? Can there be a repertoire of possible transformational rules that the cognitive system can decipher, allowing listeners to perceive a relation between a variation and a theme and perhaps even the nature of the transformation process itself?

We will first of all consider in a general way the notions of musical material, musical variation, similarity, and invariance, and then what these all suggest for the development of the notion of form-bearing dimensions in music as originally proposed by McAdams.[1] Subsequently, preliminary experiments on musical similarity that we have conducted[2,3] will be summarized. In this work, pitch and duration variations are studied for both tonal/metric and nontonal/nonmetric musical systems. Finally, a number of theoretical issues will be developed in the hopes of stimulating and orienting further work in this area, notably concerning (1) the nature of the representation of musical materials and transformation processes, (2) modularity and dimensional interactions within the realm of music cognition, (3) parallelism and associative structures in music, and (4) the notion of musical development and musical process.

MUSICAL MATERIALS

We take the notion of musical material in a very broad way. Materials may be more conventionally considered as simple figures or themes such as the opening five-note figure of Anton Webern's *Sechs Stücke für großes Orchester*, op. 6 (1909) (only slightly more complex than Beethoven's famous ta-ta-ta-tum, although not as explicitly developed) or slightly more lengthy themes such as a Bach fugue subject. They may also be more fully developed musical ideas such as the thematic materials used by Roger Reynolds in pieces such as *Archipelago* for chamber orchestra and computer-generated sound (1982–1983) or *The Behavior of Mirrors* for guitar (1986). He calls these materials *core elements*.[4] They may vary in duration anywhere from several seconds to several tens of seconds, and any given piece will have several of them, each designed to have distinctive characteristics of pitch materials, durational proportions, textures, and gestural movements. Musical materials can also be complex textures such as the dense blocks of sound used emblematically by Krzysztof Penderecki in *Threnody for the Victims of Hiroshima* for 52 strings (1960), the architectonic gestures often used by Iannis Xenakis in his orchestral works such as *Pithoprakta* (1956), and the rich micromelodic tapestries in early orchestral and choral works by György Ligeti such as *Atmosphères* (1961) and *Lux æterna* (1966). Finally, we should not leave out the realm of electroacoustic music in which spectral and temporal sound structures can be imagined from scratch and become the basis for compositional development or in which recorded sounds can become musical materials in their own right as the composer delves into the inner structure of the sound itself. Notable examples of the latter class are the deep church bell tone and the boy soprano's voice that are the two concrete sound sources from which the computer-generated piece *Mortuos Plango, Vivos Voco* (1980) was composed by Jonathan Harvey or the interpolations between unlikely partners such as a horse's laugh and a baby's cry in *Vox 5* (1986) by Trevor Wishart.

MUSICAL VARIATION

Given the variety of types of musical materials, there is obviously an infinity of ways that even simple ones can be varied in order to generate new material to fill the needs of musical discourse, all the while maintaining (most often) some kind of perceptual link to the original material so that the listener not only recognizes the relation (implicitly or explicitly), but also senses the kind of variation that has been applied. This trajectory of change is one of the perceptual components of musical development, and our intuition is that it contributes to the sense of larger-scale movement and cohesion in a work. In the simplest (operational) terms, musical variation or transformation may be taken to be simply the change along one or more musical dimensions of a given musical material.

In most musics, the principal dimensions of materials targeted for transformation are pitch and duration, although space and timbre have also become the focus of musical structuring in the 20th century. A complete catalogue of transformation devices is, of course, impossible, so only a few clearly definable types will be mentioned below. The simplest kind of (non)variation is the repetition of the same material at a different point in time. In this case there is absolute identity in a physical sense, and the perception of similarity and the recognition of the return of the material depend only on the listener and the intervening material. Another version would keep the pitch and rhythmic values constant and change the timbre through orchestration. A blatantly clear example of this approach is the piece *Boléro* by Maurice Ravel (1928), in which two themes are alternated AABB four times and then AB at the end. However, each time, the A and B melodies and their accompaniment are orchestrated differently. The similarity and sense of repetition are strong; and it is the pattern of change in orchestration, dynamics, and articulation that creates the expressive trajectory of the work. A more complex approach is to change the timbre of each note, a true Klangfarbenmelodie, as proposed by Schoenberg,[5] and then have different timbral compositions for each new return of the theme. The danger in this case, as will be discussed below, is that the melody may become fragmented perceptually if the timbre change is too drastic, bringing about the assignment of different subsets of notes to different auditory streams and thus compromising the psychological coherence of the melody as a whole.[6]

Transposition of pitch and change of tempo are the simplest kinds of transformations and testify to the fact that the representation of relations between events (exact intervals or number of scale steps related to a given scale) are prominent in musical memory. The perception of similarity of transposed melodies is indicative of the influence of internalized musical systems on recognition and similarity perception. An exactly transposed melody (one that maintains the same pitch intervals between successive notes) will in many cases be perceived as a change in key. A melody that is transposed, keeping the same number of scale steps between successive notes but with pitches that remain within the current key, is often perceived as more similar, as suggested by recognition studies, although this ability depends on musical training, musical context, and task instructions.[7] At any rate, what seems to be the important aspect of the perceptual representation underlying melodic and rhythmic similarity here is the relations among notes—that is, the relative distance in intervals or scale steps. Slight interval changes generally preserve the contour (pattern of ups and downs), a component of melody representation that has been shown to be particularly cogent for the similarity of unfamiliar melodies as measured by false recognitions.[8]

Some studies have shown that the perception of timbral intervals is also possible.[9,10] This work was based on the notion of timbre space in which the perceptual relations among sounds equated for pitch, duration, and loudness are represented mathematically as a multidimensional space. Timbral intervals are defined as oriented vectors in the space, and timbral contours could be defined by the relative patterns of up and down along the various dimensions. A transformation of a timbre melody such as transposition would simply be the translation of the timbre vectors in the space, keeping their lengths and orientations constant. Although the precise reproduction of timbral intervals is problematic with instrumental music, such musical structures are quite possible with synthesized sounds.

A variant on the simple translation operation is the inversion operation, in which the sign of the intervals and contour directions are changed. This technique is used in mirror canon writing and is also quite prominent in the 20th century in serial music.

Another class of transformations involves temporal rearrangement of the musical material. For example, playing a melody backward is called retrogradation and is often found in music of the common practice, romantic, and contemporary periods of Western music. Mirror and cancrizans canons used this technique, and it also became a prominent operation in serial music. A particularly intriguing compositional problem-solving task involves creating a melody that can be played forwards and backwards (the cancrizans or "crab" canon). Often the material that is retrograded may not reproduce the exact interval structure for reasons related to the musical syntax. In such cases, certain prominent intervals or the inverted melodic contour may preserve the sense of similiarity. Take, for example, the opening five-note figure of Anton Webern's *Sechs Stücke*, op. 6 mentioned above, in which the flute plays C♯4-E4-F4-G4 followed by D5 on the muted trumpet. Subsequently the sustained trumpet tone becomes the first note of the retrograded pattern D5 (trumpet), B♭4-E♭4-D♭4 (flute), B3 (French horn). Although the intervals are slightly different, the pitch contour (small skip, two steps, large skip) is inverted, as is the timbral contour, moving from the low-register flute to a bright, muted trumpet in one direction and from the flute to a lower, darker French horn in the other. The mirror image is quite convincing perceptually. Various other techniques of recombination have been used as well, often involving the choice of particularly salient figures from the theme and their subsequent development—a practice quite prominent in Beethoven piano sonatas, for example.

All of the preceding examples are based on rearrangements of symbolically notated events. However, in electroacoustic music, where the material is a continuous sound waveform, the operations need not be limited to predefined events. Granular synthesis, for example, takes bits of recorded sound from different parts of the original sample, applies a smoothing envelope to them and then recombines them in ingenious ways to create rich sound textures.[11] The larger the grains, the more the original sound quality "comes through" in the recombination. Roger Reynolds has used a technique involving segmentation and rearrangement in several pieces, applied either to sound samples (as in *Archipelago*, 1982–1983) or to note lists representing symbolic events (as in *Variation* for piano, 1988).[4] He has developed two classes of algorithms (SPLITZ and SPIRLZ) that transform the original samples within a space of variation constrained by the variables of the algorithm. The first one segments the sound sequence into equal-sized units and then plays the odd ones in one order and the even ones in reverse order. The spacing and timing of the segmented unfolding can be controlled. This gives a kind of alternation between the front and back ends of the sound sequence that move toward the center and then back to the extremities. The technique is obvious when applied to musical gestures with a clear trajectory, but can give stunning results in other cases even though the listener may not be able to follow explicitly the algorithm's logic. As for granular synthesis, the transformed sequences may have a more or less strong similarity to the original material depending on the size and ordering of the segments. When applied within a piece to several different themes, the transfomation device itself may in some cases take on an identity across materials that is stronger than the relations between the original and transformed materials. The perception of these kinds of materials is currently under study in our laboratory in collaboration with the composer.

Staying within the realm of electroacoustic music, there are a plethora of digital techniques now that allow an extremely fine-grained analysis of a sound's spectral and temporal structure, and then either a direct modification of certain of these parameters or even a modeling of their behavior and a subsequent resynthesis with changes in the model parameters. This kind of sonic elaboration is a class of transformation that remains quite close to the timbre and sound source identity of the materials to be transformed. Compelling examples of this approach can be found in Trevor Wishart's *Vox 5* (1986), Jean-Claude Risset's *Sud* (1985), and Jonathan Harvey's *Ritual Melodies* (1990). In particular, in the Harvey piece models of various kinds of non-Western instruments and vocal styles are used to create interpolations between the sound universes of each, including not only the sound qualities themselves, but also the styles of playing (e.g., forms of vibrato, tremolo, and portamento).

By far the most widely used type of variation involves change in the surface level structure while some invariant is retained at a higher hierarchical level. This kind of technique is quite common in theme-and-variations forms in music of the common practice period. The notions of elaboration and reduction have been formalized to a large extent by Lerdahl and Jackendoff,[12] although their applicability to contemporary music as proposed by Lerdahl[13] has been questioned.[14] Indeed, Dibben suggests that since structurally important events that "stand for" sets of subordinated events in tonal music do not have the same kind of representation in atonal music, it may be necessary to consider more associational structures to explain formal coherence in this latter kind of music, as will be discussed at the end of this paper. Evidence for similarity among musical materials related by a common reduced structure in tonal/metric music is clear (see, e.g., Refs. 15–17) and will not be belabored here. However, notions of elaboration also exist in contemporary music, although it is not clear to what extent commonalities of mental processing exist between the different musical systems.

Some of the psychological questions that arise from this brief and extremely schematic consideration of the possibilities of musical variation may include the following: How far can one transform original musical material before it becomes completely new, and perceptually unrelated—that is, what are the bounds of perceptual similarity? What does the answer to this question tell us about the representations of musical materials and the nature of the psychologically realistic processes of transformation that can operate on them?

SIMILARITY AND INVARIANCE

At the crux of the issue of musical variation are the notions of perceptual similarity and invariance, *if* the varied material appears to have at least some relation to the original material. What we actually mean by "similarity" is not easy to pin down operationally. Let us just assume that it is some degree of match between the properties and features of two materials. The experimental question is thus to understand what properties contribute to perceived similarity and how the degree of match between these (to-be-determined) properties determines its strength. We will speculate on the nature of various kinds of properties and levels of abstraction that might contribute to musical similarity as judged directly by listeners.

Properties that contribute to similarity may be of various levels of abstraction, moving from the most concrete (specific values of attributes of individual events in the sequence) to the most abstract (higher-level reductions of the hierarchical event structure, if such exists for the material being heard, or harmonic contours that provide a kind of skelelton for a tonal work). The values and relations thus analyzed from the musical sequence may also be explicitly tied to the temporal order of the musical sequence or experienced under transformations of temporal order or even as distributions with reference to specific order.

Concerning surface values, the surface content of a material and its transformation (e.g., specific pitches, durations, loudnesses, and timbres) may have more or fewer events in common along the various auditory dimensions. Also, the temporal order of these events may be the same or may be transformed to a greater or lesser extent. Exact repetition would, of course, have the highest similarity, and completely different values would have no similarity. However, under some situations the exact same number of events with specific attributes in completely different orders (a random rearrangement of the events, for example) may still be experienced as similar due to identical statistical distributions of material properties. As mentioned above, however, musical material can also be represented in terms of relations among attributes in a more or less precise way: exact intervals, relations within a musical system such as a scale, contours, and so on. The perception of similarity would thus be based on a match between relations abstracted from the material attributes rather than directly on the values of the attributes themselves. A melodic contour in Béla Bartók's *Music for Strings, Percussion and Celesta* (1936), for example, might undergo an expansive transformation in which all of the intervals are increased, and yet the whole stays within a certain pitch system, and the pitch contour is preserved. Since the rhythmic material is relatively similar between the versions, listeners may hear a strong link between this variation and the original.

Finally, as also mentioned previously, research[16,17] has shown that materials that are quite different in surface structure (affecting specific surface values and the relations between adjacent events) may still be perceived as similar if they share an underlying structure, such as the reduced structure proposed by Lerdahl and Jackendoff.[12] In these cases, it is necessary to postulate that the representation of the event structure of the material is organized to some extent in a hierarchical fashion and that it is some level of this hierarchy that is compared between versions. The experimental question here concerns the depth of the hierarchical representation that can be used to derive a sense of similarity between materials that differ drastically in surface structure. It would seem that it is this kind of structural abstraction that would be one of the main bases of the musical invariance underlying theme-and-variations forms.

From these considerations, we propose that there are (at least) three types of similarity that should be taken into consideration, all the while understanding that they are overlapping and non-mutually exclusive categories. The first is relatively abstract and concerns similarity of the statistical distribution of surface values and their derivatives as well as of surface relations of first and higher orders, including perhaps even transitional probabilities from one value or relation to another. This kind of similarity would seem to be involved in the perception of similar textures, where it is not the exact values that count but the probability of occurrence and of transition that define the texture. The second kind concerns figural similarity, specific patterns of attributes (perhaps associated intimately across auditory dimensions) that give perceptual landmarks within an extended material and that allow its recognition when a transformed version still contains the same identifying pattern, or something very similar to it. This notion of figural similarity is akin to the notion of imprint proposed by Deliège.[18] The third kind of similarity is structural and relies on the abstraction of structural invariants related to the event hierarchy, often defined by underlying harmonic and metric templates that are maintained constant in cases of elaborative or reductive transformations of thematic materials.

FORM-BEARING DIMENSIONS

A framework has been developed concerning the psychological constraints on form-bearing dimensions in music.[1] Among these psychological constraints, some seem particularly cogent for the present concerns. Form-bearing dimensions are those along which invariant relations can be configured. For them to be useful in musical discourse these relations must be easily discriminable. Further, in evaluating the degree of similarity between thematic materials and their transformations, time has necessarily passed between the two, often several tens of seconds in real music. The configurations must thus be memorizable to some degree of precision either in their absolute form or in some abstracted relative form.

In real music, several dimensions may be varied for a given material. For example, transposing a melody in pitch that is played by the same instrument maintains the same rhythmic structure, and while globally the timbre may change slightly, it remains roughly constant. One psychological constraint on form-bearing dimensions that merits extensive consideration is the resistance of perceived invariance on one dimension to independent variation on other dimensions. In other words, to what extent are the dimensions processed independently or do they interact? This raises the modularity question that has been addressed in both theoretical and experimental terms.[19–21] Both the psychological and the musical issues are extremely complex here, as it becomes clear that one may have varying degrees of identity and difference on different dimensions, and the question would be to know how strong the sense of similarity was under certain conditions, or even whether similarity could be maintained. The perception of such transformations would depend on several factors. At the level of the stimulus, it would depend on the way the composer used variation across dimensions and on the types of interactions between dimensions, which may indeed by asymmetric. For example, in Olivier Messiaen's *Mode de Valeurs et d'Intensités* for piano (1950), many dimensions are used: pitch, dynamics, articulation, duration. However, there is a coupling between the values of one and those of the others. Krumhansl has shown[22] that listeners pick up on these interdimensional correlations and can judge that new material fits or does not fit with rules as developed by the composer. In integral serialism, on the other hand, the composer may have series of pitches, dynamics, articulations, and durations, each series having a different number of elements, and they thus become uncorrelated with one another, as in Pierre Boulez' *Structures* for two pianos (1952). It would be interesting to repeat Krumhansl's experiment on such material. However, she and her colleagues have shown that listeners do hear different transformations of row forms (retrograde, inversion, and retrograde inversion) as having a similarity relation to the original series in twelve-tone rows of Schoenberg.[23]

A psychological factor that needs to be considered is the resistance of perceived invariance/similarity to perturbation by intervening material. We often presume as score gazers that recognition of the return of a theme is unaffected by new intervening material. However, work in auditory short-term memory has shown that relations of attribute similarity between the to-be-remembered information and the intervening material can degrade recognition and discrimination performance. This kind of auditory memory does seem to be somewhat modular, however, in that intervening material that is different on the criterial dimension but similar on other dimensions does not necessarily degrade performance.[24–28] At a more complex level, it is unclear at present what kinds of contextual effects may exist that would render some intervening material more or less perturbing of the memory trace of the original material and thus hinder its recognition. If such were the case, then the "success" of a given theme in contributing to a sense of coherence in a musical form would certainly be compromised.

Given that materials can vary on different dimensions and that there have been strong tendencies in many cultures to focus primarily on pitch and duration, another psychological question that comes to the fore is the extent to which listeners can learn to focus on changes along a dimension not usually used in one's past experience as a structuring dimension in musical discourse, such as timbre or space, for example. That listeners do not have experience in listening for such structures would probably show up in psychological testing, but we could not conclude from such results that the use of these dimensions was impossible or doomed to failure from the outset. So the question of perceptual learning will need to be addressed within the framework of contemporary musical experience.

Finally, one of the crucial problems of the perception of musical transformations in relation to their original material concerns perceptual coherence. Variations that affect stream formation or change the availability of properties of the material in working memory may, indeed, render what is perceived unsimilar to the original material, as relations among perceptual attributes would appear to be computed on events belonging to the same auditory stream.[6,29]

EXPERIMENTS ON MUSICAL SIMILARITY

Having drawn with broad strokes a rough and largely intuitive account of what musical similarity might be and what its role in music might be, let us examine the few studies that have specifically addressed this question, focusing in particular on some of our own recent work that takes a few small steps toward the answer. Several experiments have studied the implication of musical training, musical memory, pitch hierarchy, and musical context in the perception of tonal and nontonal transformations and their similarity ratings.[8,23,30–33] For example, similarity perception is not a symmetrical phenomenon: the similarity perceived between a melody A and a melody B does not automatically predict the similarity perceived between B and A.[30] Finally, although participants show an ability to recognize the type of transformations applied to a nontonal material presented isochronously, this ability tends to decline when transformations are multidimensional—that is, when both pitch and rhythm dimensions vary simultaneously.[8,23,31] All these facts, however, do not explore or explain the computational basis of melodic similarity perception. Is it plausible to imagine that melodic similarity perception could rely on general mental processes that are common to all listeners, independently from the melody's morphology (that is, for example, number of pitches, rhythmic figures, and metronomic speed) and the listener's musical aptitude?

In recent work, we asked the following questions: What is the nature of musical transformation space, and how is it constrained by the type of musical material and musical training? To what extent do pitch and duration components of musical materials contribute independently to global similarity judgments and discrimination focused on each? The first question was addressed in Experiment 1 and the latter in Experiment 2. The full details of the experiments are found in References 2 and 3.

EXPERIMENT 1

Three tonal melodies were used as reference material. Following Serafine *et al.*[16] a lure melody was constructed for each one such that the surface characteristics were very similar, while the underlying reduced structure was different. For each melody, 16 transformations were composed that differed from the original in pitch content, rhythmic content, or both. The transformations could be either elaborations or reductions of the original material, reduction being taken in the sense of a simplified abstract representation of a musical structure that keeps structural information such as important notes in the tonal and metric hierarchies.[12] Further, the pitch and rhythmic transformations either respected tonal/metric syntax (reasonable variations within the scale and meter) or departed from it (unlikely chromatic and unmetrical variations). Corresponding structural transformations were applied to the lure melodies as well. The melodies were verified with respect to the desired properties by a practising composer (Joshua Fineberg) working at IRCAM.

Musician and nonmusician listeners heard one of the three reference melodies followed by a test melody. The test melodies could be the reference or its corresponding lure or any of their transformations. The listeners made a similarity rating on a scale of 1 to 9, with the high end corresponding to maximum similarity.

The experiment was designed to test several hypotheses:

(1) Musicians should be able to hear similarity to greater degrees of transformation if the transformations respect the syntactical rules.
(2) If listeners are sensitive to commonalities at certain levels of hierarchical reduction, transformations that respect the reduction should be more similar than those that violated it (including the lure transformations).
(3) Drastic transformations in both pitch and rhythm should render the variation as dissimilar to the reference as unrelated material, represented by the lure transformations.
(4) The same types of transformation should have similar effects on the three reference melodies, although local morphological differences may prevail due to figural properties of the melodies.

The data revealed no systematic differences between the two populations of listeners. This last finding is concordant with previous studies that have specifically examined effects of musical training on musically realistic tasks.[34] Transformations that respected the reduced structure were judged significantly more similar to the reference than those that did not. This effect was particularly strong for chromatic transformations and for combined chromatic and nonmetric transformations, many of which were judged as dissimilar as the lure melody and its transformations. Finally, the three reference melodies were different with regard to the perceived effects of the transformation types, demonstrating the importance of local surface features in the perception of melodic similarity and the fact that a transformational algorithm must necessarily interact complexly with the to-be-transformed material. However, rank correlations between reference melodies were very high, indicating similar qualitative effects of the set of transformations applied to all three melodies.

EXPERIMENT 2

The answer to one question that cannot be easily gleaned from the results of the first study is the degree to which pitch and rhythmic materials interact in the perception of similarity. The second experiment was designed to address this problem more explictly. We, therefore, introduced a simple computational model based on a modular hypothesis that could represent the basis of the mental processing underlying melodic similarity perception. If the computational model turns out to be reliable for tonal melodies, can it also predict similarity perception of nontonal melodies and thus indicate general processing mechanisms independently of the musical system? One may wonder whether or not a modular approach is relevant to the study of music perception in general and of similarity perception in particular. Examples taken from vision, language, audition, and neuropsychology lead us to consider the modular hypothesis in order to clarify the specificities of similarity perception. Interestingly, aside from modular considerations in music theory,[19] the only consequential empirical research done within an explicitly modular perspective comes from the neuropsychological work of Peretz and collaborators.[20,35,36]

Based on previous findings, which argue in favor of independent pitch and duration processing structures,[33,37] the main hypothesis posited here is that a similarity judgment made between a reference melody and a transformation of that melody could be understood to a certain extent as an additive integration of two similarity coefficients calculated by the two hypothetical pitch and duration modules. While such an additive approach is certainly a simplistic way to consider the relations between the two modules, such a general principle should be applicable to any musical system without regard to the nature of the underlying musical grammar, the tonal and twelve-tone systems in our case.

To test the independence, a paradigm similar to that of Garner and Morton[38] was used in which discriminations of changes on one dimension were made in the presence or absence of changes along another dimension and vice versa by different groups of listeners. A third group was required to make both discriminations. We coupled this discrimination task with a similarity judgment task. This paradigm was used in order to investigate whether the listeners could recognize a pitch or duration pattern independently from the values taken by the other dimension. Note that since the judgment is focused on a single dimension, at times in the face of variation on the other dimension, perfect discrimination performance across all conditions indicates an ability to ignore the other dimension, implying perceptual independence.

Four reference melodies were composed: one major, one minor, and two twelve-tone melodies. In addition to the original pitch or rhythm patterns, three variations of each dimension were created. These variations were designed to be progressively less similar to the reference pattern. The pitch and rhythm patterns were crossed, giving 15 transformations of each reference. The melodies were all composed to achieve the desired properties by a practising composer (Jacopo Baboni-Schilingi) working at IRCAM. The transformations could thus differ only in pitch, only in rhythm, or in both. Listeners were divided into three groups: pitch discrimination, rhythm discrimination, or both. They first made discrimination judgments on the pitch and/or temporal dimension and then judged globally on a 10-point scale the strength of similarity for each type of transformation in relation to the reference melody.

The discrimination results globally corroborate the modularity hypothesis, although complete perceptual independence was not observed. In particular, the pitch and duration modules do not exhibit the same processing properties, as duration processing seems to be less permeable to pitch information than pitch is to duration (at least for tonal melodies). This is reflected when comparing participants' discrimination performance on the two musical dimensions for the two musical systems. For all melodies, duration configurations were better discriminated than pitch configurations; and, in particular, identical pitch patterns were judged as being progressively different when the rhythm pattern was, indeed, more and more removed from the reference, even though the pitch pattern was identical. By deduction this finding is coherent with existing empirical data in melodic similarity perception, where it has been shown that the duration dimension was the major one used by participants for establishing similarity judgments.[39] The pitch permeability to duration configuration was limited to tonal melodies, suggesting a functional interaction between the two in that musical system that is decoupled in the twelve-tone system.

Variations among listeners' similarity judgments showed that tonal and nontonal transformations were globally perceived in different ways. While the tonal transformations were hierarchically distinguished, this was not the case for the perception of the nontonal transformations, where the pitch and duration modalities were not hierarchically differentiated. Further, whereas the perception of tonal transformations appeared to be perceived in a systematic way over repetitions and for the two reference melodies, this was not the case for the nontonal transformations, where the range of variation of listeners' judgments was found to be much greater. Consequently, it seems reasonable to infer that similarity perception of tonal and nontonal melodies depended at least partly on different mental processing mechanisms.

Under the perceptual independence hypothesis, we claimed that similarity perception could be approximated by an elementary additive model. The model decomposes the listeners' similarity ratings into two numerical values, basic mean pitch and duration values for each level of transformation of that dimension without change in the other dimension. The additive combination of the mean values for pitch and duration gives the model prediction. For three of the four reference melodies (including both tonal melodies and the first nontonal melody), listeners' mean similarity judgments for the nine two-dimensional transformations were highly correlated with model predictions. So although some differences between the tonal and nontonal melodies exist, these melodies also seem to have certain commonalities with regard to processing in rhythmic and pitch similarity. There are, nonetheless, deviations of the data from model predictions for each transformation, and these will need to be examined in detail to determine their cause, particularly for the second nontonal melody. There are perhaps salient surface or structural features not taken into account by this simplistic model that are playing a role in the perceived similarity.

CONCLUSIONS/QUESTIONS

From the data, it is clear that the space of possible variation of thematic material that still maintains a link of perceptual similarity to the original is limited. Although conditions of direct comparison are different from those in real music, where intervening material or competing material would often be present, a number of factors can be targeted provisionally for more contextually realistic work in the future. A transformation that maintains a majority of exact values or relations among values in the correct temporal order will be more strongly similar than one that changes all values and all relations. Although this has not been explicitly tested, it may be that maintaining the same distribution of values or relations could still elicit a sense of similarity if compared with material that had a dramatically different distribution of attribute values. Further relations of similarity can exist at more abstract levels of a hierarchical reduction, indicating something about the hierarchical nature of representation of the mental representation of musical material in both tonal and nontonal music. From examining the similarity judgments on elaborative and reductive transformations, it would seem that these kinds of transformations retain similarity to the extent that the transformations do not violate the "grammaticality" of the musical system within which the material has been conceived, suggesting a close link between comprehension of the musical system and the categorization processes that one may hypothesize to operate in the recognition of variations of an original theme.

Another issue of major import for the notion of musical similarity and invariance is the notion of perceptual independence between patterns on different musical dimensions. While there seems to be some degree of independence between pitch and duration structures at a global level, there are also asymmetrical interactions in which duration changes more strongly affect the perceived similarity of identical pitch patterns than is the case for the reverse. More refined experimentation will be necessary to probe the nature of the interaction between the dimensions that gives rise to such asymmetries. Also, future research should clearly study more thoroughly different musical systems, including those of other cultures, in an attempt to distinguish the more general cognitive processes from those that are specific to a particular musical grammar.

Finally, future research should examine the role that the perception of musical similarity plays in the apprehension or experience of musical form. The process of similarity perception necessarily involves comparison across time of different manifestations of musical material and thus involves memory processes. One might imagine that the recognition of similarity could thus establish some kind of associative link between different points in the piece and thus contribute in some way to a sense of formal cohesion over longer time spans. Further, one not only recognizes the link between two manifestations, but can also appreciate the difference between the two, a perception that perhaps contributes to a sense of the kind of change or the trajectory of change that has taken place. This trajectory may contribute in itself to a sense of musical development over longer time spans and also reinforce the integration of materials at different points in time into a greater whole. There are numerous theoretical and methodological obstacles to be overcome in the design of appropriate experiments to test these contributions in ongoing music listening.

ACKNOWLEDGMENTS

This work has benefited from fruitful discussions with several composers (Jacopo Baboni-Schilingi, Benjamin Carson, Joshua Fineberg, and Roger Reynolds) and a fellow psychologist (Emmanuel Bigand), to whom we express our gratitude.

REFERENCES

1. McAdams, S. 1989. Psychological constraints on form-bearing dimensions in music. Contemp. Music Rev. **4**(1): 181–198.
2. Matzkin, D. 1998. La perception de similarité musicale peut-elle être étudiée dans le cadre de la théorie de la modularité et de l'indépendance perceptive? Mémoire de DEA, EHESS. Paris.
3. Matzkin, D. & S. McAdams. Does similarity perception in music involve systematic mental processing regardless of the nature of the musical system? Submitted for publication.
4. Reynolds, R. 2001. Form and Method: Composing Music. Contemporary Music Studies. Harwood Academic Press. London.
5. Schoenberg, A. 1978. Theory of Harmony. (Original German publication 1911). University of California Press. Berkeley, CA.
6. McAdams, S. & A.S. Bregman. 1979. Hearing musical streams. Comput. Mus. J. **3**(4): 26–43.

7. DOWLING, W.J. 1986. Context effects on melody recognition: scale-step versus interval representations. Music Percept. **3:** 281–296.
8. DOWLING, W.J. 1972. Recognition of melodic transformations: inversion, retrograde, and retrograde inversion. Percept. Psychophys. **12:** 417–421.
9. EHRESMAN, D. & D.L. WESSEL. 1978. Perception of timbral analogies. IRCAM, Report 13. Paris.
10. MCADAMS, S. & J.C. CUNIBILE. 1992. Perception of timbral analogies. Philos. Trans. R. Soc. Lond. Ser. B. **336:** 383–389.
11. ROADS, C. 1985. Granular synthesis of sound. *In* Foundations of Computer Music. C. Roads & J. Strawn, Eds: 145–159. MIT Press. Cambridge, MA.
12. LERDAHL, F. & R. JACKENDOFF. 1983. The Generative Theory of Tonal Music. MIT Press. Cambridge, MA.
13. LERDAHL, F. 1989. Atonal prolongational structure. Contemp. Music Rev. **4:** 65–87.
14. DIBBEN, N. 1994. The cognitive reality of hierarchical structure in tonal and atonal music. Music Percep. **12:** 1–25.
15. STOFFER, T.S. 1985. Representation of phrase structure in the perception of music. Music Percept. **3:** 191–220.
16. SERAFINE, M.L., N. GLASSMAN & C. OVERBEEK. 1989. The cognitive reality of hierarchic structure in music. Music Percept. **6:** 397–430.
17. BIGAND, E. 1990. Abstraction of two forms of underlying structures in a tonal melody. Psychol. Music **18:** 45–60.
18. DELIÈGE, I. 1989. A perceptual approach to contemporary musical forms. Contemp. Music Rev. **4:** 213–230.
19. TEMPERLEY, D. 1995. Motivic perception and modularity. Music Percept. **13:** 141–170.
20. PERETZ, I. & J. MORAIS. 1989. Music and modularity. Contemp. Music Rev. **2**(1): 279–294.
21. FODOR, J. 1983. The modularity of mind. MIT Press. Cambridge, MA.
22. KRUMHANSL, C.L. 1991. Memory for musical surface. Mem. Cognit. **19:** 401–411.
23. KRUMHANSL, C.L., G.J. SANDELL & D.C. SERGEANT. 1987. The perception of tone hierarchies and mirror forms in twelve tone serial music. Music Percept. **5:** 31–78.
24. STARR, G. & M. PITT. 1997. Interference effects in short-term memory for timbre. J. Acoust. Soc. Am. **102:** 486–494.
25. SEMAL, C. & L. DEMANY. 1991. Dissociation of pitch from timbre in auditory short-term memory. J. Acoust. Soc. Am. **89:** 2404–2410.
26. SEMAL, C. & L. DEMANY. 1993. Further evidence for an autonomous processing of pitch in auditory short-term memory. J. Acoust. Soc. Am. **94:** 1315–1322.
27. DEUTSCH, D. 1970. Tones and numbers: specificity of interference in immediate memory. Science **168:** 1604–1605.
28. DEUTSCH, D. 1972. Mapping of interactions in the pitch memory store. Science **175:** 1020–1022.
29. BREGMAN, A.S. 1990. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press. Cambridge, MA.
30. BARTLETT, J.C. & W.J. DOWLING. 1988. Scale structure and similarity of melodies. Music Percept. **5:** 286–313.
31. FRANCÈS, R. 1972. La perception de la musique. Vrin. Paris.
32. HALPERN, A.R., J.C. BARTLETT & W.J. DOWLING. 1998. Perception of mode, rhythm, and contour in unfamiliar melodies. Effect of age and experience. Music Percept. **15:** 335–355.
33. PITT, M.A. & C.B. MONAHAN. 1987. The perceived similarity of auditory polyrhythms. Percept. Psychophys. **41:** 534–546.
34. BIGAND, E. 1990. Perception et compréhension des phrases musicales. Ph.D. thesis. Université de Nanterre (Paris X). Nanterre, France.
35. PERETZ, I. 1993. Auditory agnosia: a functional analysis. *In* Thinking in Sound: The Cognitive Psychology of Human Audition. S. McAdams & E. Bigand, Eds: 199–230. Oxford University Press. Oxford.
36. PERETZ, I. & R. KOLINSKY. 1993. Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. Q. J. Exp. Psychol. **46A:** 301–325.

37. PALMER, C. & C.L. KRUMHANSL. 1987. Independent temporal and pitch structures in the determination of musical phrases. J. Exp. Psychol. Hum. Percept. Perform. **13:** 116–126.
38. GARNER, W.R. & J. MORTON. 1969. Perceptual independence: definition, models, and experimental paradigms. Psychol. Bull. **72:** 233–259.
39. MONAHAN, C.B. & E.C. CARTERETTE. 1985. Pitch and duration as determinants of musical space. Music Percept. **1:** 1–32.