

# Mikrofonie I (1964) by Karlheinz Stockhausen - Between Idea and Auditory Image

Justyna Humięcka-Jakubowska

## I. INTRODUCTION

**Abstract**—Background in music analysis: Traditionally, when we think about a composer's sketches, the chances are that we are thinking in terms of the working out of detail, rather than the evolution of an overall concept. Since music is a "time art," it follows that questions of a form cannot be entirely detached from considerations of time. One could say that composers tend to regard time either as a place gradually and partially intuitively filled, or they can look for a specific strategy to occupy it. It seems that the one thing that sheds light on Stockhausen's compositional thinking is his frequent use of "form schemas," that is often a single-page representation of the entire structure of a piece.

Background in music technology: Sonic Visualiser is a program used to study a musical recording. It is an open source application for viewing, analyzing, and annotating music audio files. It contains a number of visualisation tools, which are designed with useful default parameters for musical analysis. Additionally, the Vamp plugin format of SV supports to provide analysis such as for example structural segmentation.

Aims: The aim of paper is to show how SV may be used to obtain a better understanding of the specific musical work, and how the compositional strategy does impact on musical structures and musical surfaces. It is known that "traditional" music analytic methods don't allow indicating interrelationships between musical surface (which is perceived) and underlying musical/acoustical structure.

Main Contribution: Stockhausen had dealt with the most diverse musical problems by the most varied methods. A characteristic which he had never ceased to be placed at the center of his thought and works, it was the quest for a new balance founded upon an acute connection between speculation and intuition. In the case with *Mikrofonie I* (1964) for tam-tam and 6 players Stockhausen makes a distinction between the "connection scheme," which indicates the ground rules underlying all versions, and the form scheme, which is associated with a particular version. The preface to the published score includes both the connection scheme, and a single instance of a "form scheme," which is what one can hear on the CD recording. In the current study, the insight into the compositional strategy chosen by Stockhausen was been compared with auditory image, that is, with the perceived musical surface. Stockhausen's musical work is analyzed both in terms of melodic/voice and timbre evolution.

Implications: The current study shows how musical structures have determined of musical surface. The general assumption is this, that while listening to music we can extract basic kinds of musical information from musical surfaces. It is shown that interactive strategies of musical structure analysis can offer a very fruitful way of looking directly into certain structural features of music.

**Keywords**—Automated analysis, composer's strategy, *Mikrofonie I*, musical surface, Stockhausen.

This work has been funded by the National Science Centre in Poland appropriations for science in the years 2012–2015 as research project Grant number 2011/03/B/HS2/03935, Program Opus 2.

Justyna Humięcka-Jakubowska is with the Department of Musicology, Adam Mickiewicz University, Poznań, Poland (e-mail: jhj@amu.edu.pl).

THE inspiration for the scientific reflection on Stockhausen's *Mikrofonie I* from the point of view of the interrelationship between the creative idea and the auditory image is the fact of the gradual abandonment of highly parametric approach to electronic music in favor of *composer's instincts* [1] as the arts gained control. According to Zampronha "non-motivated" parametric procedures were replaced, over time, by approaches which acknowledged the "complexity of listening and sound references" [2]. In turn, this change contributed to more perceptual approaches towards musical work itself.

The essence of electroacoustic music at all, and *Mikrofonie I* in particular, is the use of the purely sonoric properties of the sound material for artistic purposes. In the case of electronic music, in a broader sense, any sound within the limits of human perception becomes a material for composition. However, also during live performances conventional sounds can be modified in real time using electronics, e.g. microphones (electroacoustic music). Hence the communicational competence between a composer and a listener often boils down to the ability to recognize communicational situations and the composer's selection of appropriate means for the creation of an auditory perceptible sound shape, the musical surface. This auditory perceptible musical surface results from the composer's auditory representations, which—in turn—reflect his ideas and influence the shape of the perceptual representations, arising in the mind of the listener while consciously listening to music.

In Maconie opinion, regarding to *Mikrofonie I*, "the music of *Mikrofonie I* is derived entirely out of a single complex vibrating body. Not only is the piece integrated in theory, it is also acoustically integrated in fact. For the first time; a perceptual equivalent to totally-organized structure has been found, and it is particularly significant that this has been achieved by very simple means" [3].

To understand Stockhausen's creative ideas of this musical work, one must remember that the composer has written a detailed introduction to the work, reproduced in the CD booklet, and moreover, the preface to the published score includes the "connection scheme." It seems important also that, the "connection schema" isn't identical with the "form schema," which is associated with a particular version, actually realized by musicians during the performance of the musical work. The "connection scheme" contains substantial interchangeable elements that call for the preparation of the specific performance, and also the basic clues underlying all

possible performances [4].

To examine the essential assumptions of Stockhausen, one needs to compare advisable creative ideas with analysis results. It seems useful the use of automated analysis. In this case, Sonic Visualiser program, taking advantage of computer-based techniques of analysis, is designed to offer different ways of visualizing the music as you hear it.

## II. MIKROPHONIE I IN THE CONTEXT OF STOCKHAUSEN'S SPECIFIC CREATIVE IDEAS

### A. Sound Material

The only instrument in traditional sense is a large tam-tam, on which two musicians play, standing on each side of it. But the entire range of sound materials is produced coming from various manner of generation of these sounds, and what is more important, all the sounds are controlled, by means of scan tam-tam with hand-held microphones by two operators standing alongside musicians. The listeners perceive, more or less surprising and complex sounds. This is the effect of transferring of the scanned sounds by two variable-bandwidth filters, and they then pass, filtered and amplified electronically, to separate speakers placed at either side of the stage. Ultimately, one can distinguish some general types of sound qualities. There are the sounds similar to the voice, the sounds coming from room (resonance), the sounds similar to the percussion and texture effects, and also the sounds sometimes perceived as pitch changes in the result of filters modifications of sound.

Before composing *Mikrophonie I* Stockhausen along with his assistant Jaap Spek, who was working in the studio for electronic music at Cologne Radio, has done the experiment. Stockhausen has mentioned, that in the effect of this experiment "We recorded for about twenty minutes and then I walked in and said, let's hear it. And I must say that what we both heard was so astonishing that we started embracing each other and saying, this is unbelievable, a great discovery. We heard all sorts of animals that I had never heard before, and at the same time many sounds of a kind I couldn't have possibly imagined or discovered, not in the twelve years I had worked in the electronic music studio up to the time of that experiment" [5].

During the lecture MIKROPHONIE I, filmed by Allied Artists in London in 1971, the composer has described the sounds, which the listeners hear in the work, by the words "within each sound channel, every sound event that comes through the speakers is the outcome of a multiple interference of the actions of the three musicians in each group (...) Nobody can say what the sounds really are, other than tam-tam sounds: we cannot describe them purely on the basis of what we hear coming from the loudspeakers. (There something very characteristic of our conventions of analysis—not only for new music, but in other areas too—that the further we go into the microcosm, the more we have to describe what we are observing in terms of the tools we are using.) So while we are getting more or less what our instruments can provide, there remains a lot of mystery about the sounds produced

directly in the tam-tam itself" [5].

### B. Stockhausen's Aims: The Conception of the Work

Stockhausen's aim, which is included in the conception of the work, has been stressed in the title of the composition. On the one hand, he wanted to emphasize, that the microphone is played as a musical instrument. On the other hand, by means of suffix "-phony," Stockhausen would suggest collectivity of the performance, as in symphony.

Moreover, Stockhausen wanted to give the audience the creative idea of such music, in which—regarding to its narrative flow—listeners can observe liaison and progression between two groups of three performers (tam-tam—microphone—filter) almost always playing alternately (*tutti* is planned only three times).

The creative idea contains itself also in the ways in which the musical structures were composed. It is worth noting at this point that Stockhausen has applied regarding to the musical structure of *Mikrophonie I* of the alternative name, that is mean, "moment." But it seems important here, that the "moment" in the process of musical narration of the work tend to be self-contained, because it does not gradually develop out of specific features from the previous musical structure. According to Stockhausen, the musical structures are "independent" here.

The differences in the range of their contents and characteristics determine three relationships for every connection between two successive structures (*entsprechend* that is to say similar; *verschieden* that is to say different; and *gegensätzlich* or opposite). In the consequence of these possible types of successive connections, a musical narration of *Mikrophonie I* may be—respectively—supporting (*unterstützend*), neutral (neutral), or destroying (*zerstörend*) to the mutual relations of every connection between the two structures. In turn, a perception of connections, which are created in such a way, may lead to feelings increasing (*zunehmend*), constant (*konstant*), or decreasing (*abnehmend*) flow of musical narration [6].

### C. Connection Scheme and Score

Because the performers have to choose the order of the structures, although according to composer's general idea of creation of variable musical narration, the "form scheme" of *Mikrophonie I* is different as to particular performance. Nevertheless Stockhausen has made the "connection scheme," which indicates the ground rules underlying all versions of performances.

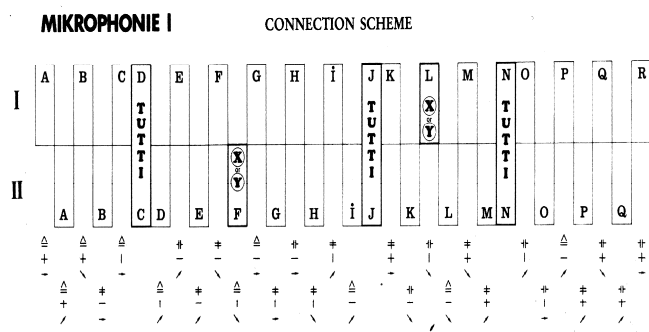


Fig. 1 *Mikrophonie I*—Connection Scheme [4]

On the basis of the “connection scheme” one can note, that Stockhausen predicted three *tutti* moments in the range thirty-three moments in all, but he didn't determine which of them occurs at which point—this is left free for performers—but they have to occur at set points of musical narration. Moreover, there are also two solo moments (X and Y), whose position is also fixed in the musical work, though again their ordering is interchangeable: either X or Y can come first.

The graphic signs below the “connection schema” have to do with mentioned above three relationships for every connection between two successive structures (top line) and three ways of impact the new structure in relation to the previous one (middle line), and finally, three types of feeling of the musical narration flow (bottom line).

The *Mikrophonie I* is a music notated graphically and more or less conventionally for live performers in relation to timing and pitch. As Stockhausen has emphasized “The actions of each group are indicated in a score. The score is divided into three layers. The upper layer is for the exciter who plays on the surface of the tam-tam, the middle layer is for the microphonist, and the bottom layer is for the player of the filter and potentiometer. The tam-tam player's part is subdivided into three registers: high, middle and low, indicating relative pitch. The microphonist's part also subdivides into three, but indicating relative distance, moving across the tam-tam, from the point of excitation on the surface. The for the filter player a similar subdivision of part into three indicates high, medium and low registers of filtering (actually the filter that we use has nine more or less equidistant frequency divisions between 30 and 12,000 cycles per second, so the player can work on a basis of three levels per register). All the time measurements are drawn to scale, with more precise indications for tempo and rhythm. The thickness of the notated lines for the tam-tam player indicates relative intensity: the thicker the line the more intensely you play. For the microphonist the thickest line means closest to the surface of the tam-tam, the thinnest line the furthest away, moving perpendicular to the plane of the tam-tam” [5].

During the creative process of *Mikrophonie I*, after a phase of experimentation and structuring, it soon turned out that regarding to *Klangfarben* effects of liaison and continuity—difficult to achieve electronically—a system of notation based on a description of how to play was inadequate. Two people

rubbing a piece of rubber on a tam-tam could give two diametrically different results. Instead, Stockhausen used a scale of descriptive terms, words that describe something about the way the sound resonates. The words are the names of the all thirty-three moments of *Mikrophonie I*, but these are rather vague terms. The names, as closely as possible describing the sounds of the moments (often in an onomatopoeic fashion) and other names in relation to the moments, are listed in the booklet's English translation of the German. The one word in German sometimes requires two words in English translation, and even sometimes this is only approximation the core meaning of the German word.

There are the German names original in capital letters: ÄCHZEND: groaning, creaking – BELLEND: baying, barking – BERSTEND: bursting – BRÜLLEND: bellowing, bawling – BRUMMEND: growling (low buzzing) – DONNERND: thundering – FAUCHEND: hissing, spitting – FLÖTEND: fluting – GACKERND: cackling – GELLEND: yelling – GERÄUSCH: noise – GRUNZEND: grunting – HAUCHEND: exhaling (like a breeze) – HEULEND: howling – JAULEND: wailing – KLÄNGE: pitched sounds – KLAPPERND: clacking – KLATSCHEND: clapping – KLIRREND: clinking, jingling – KNACKEND: cracking – KNALLEND: banging, clanging – KNARREND: grating – KNATTERND: chattering, flapping – KNIRSCHEND: crunching, gnashing – KNISTERND: crisping, crinkling – KNURREND: grumbling, snarling – KRACHEND: crashing – KRÄCHZEND: cawing – KRATZEND: scratching – KREISCHEND: shrieking, screeching – LÄUTEND: pealing, tolling – MURMELND: murmuring – PFEIFEND: piping, whistling – PIEPSEND: cheeping – POSAUNEND: tromboning – PRASSELND: spattering, jangling – PRELLEND: slapping, rebounding – QUAKEND: croaking, quacking – QUIETSCHEND: squeaking, squealing – RASCHELND: crackling – RASSELND: clashing, clanking – RATTELND: rattling – RATTERND: clattering – RAUSCHEND: rushing, rustling – REIBEND: rubbing – RÖCHELND: choking (rattling in the throat) – ROLLEND: rolling – RUMPELND: rumbling, thumping – SÄGEND: sawing – SCHARREND: scraping – SCHLÜRFEND: shuffling, slurping – SCHNARCHEND: snorting, snoring – SCHNARREND: twanging, rasping – SCHWIRREND: whizzing, whirring – SINGEND: singing (whining) – TÖNEND: ringing, resounding – TOSEND: roaring – TRILLERND: trilling, tinkling – TROMMELND: drumming – TROMPETEND: trumpeting – TUTEND: hooting – UNKEND: keening (or mourning with “u”-timbre) – WINSELND: whimpering – WIRBELND: whirling – WISCHEND: wiping, swishing – WISPERND: whispering – ZIRPEND: chirping – ZUPFEND: plucking. These words still gave the opportunity to notate abstract sound in a way that actually makes it possible to play the same rubbing sound twice. Thus, the possibility to structure the composition was present.

This unique way of describing applied sounds was the result of such composer's thoughts: “I first tried to describe the actions; in fact I initially worked out a score which gave

instructions such as 'Take a plastic box of a particular size (I would give the exact size), and hold it against the surface of the tam-tam at a specific angle (and I tried to describe the angle), and then scratch it with a quick movement.' etc. This score became incredibly complicated in its description of the playing materials, and I felt it was absurd to have to write a score in 1964 saying this sort of thing. So I abandoned completely the description of the process in terms of measures and tools. I didn't even want to prescribe the materials any more: whether rubber, or wood, or metal or glass. Then I said, what shall I do? How can I recover this sound world, even approximately, in any future? And how can this piece become a model for similar processes? What I did was the following: I made a scale. This scale went right back to the dawn of music, so to speak, to the kind of language the technician and I used when we were talking about the kinds of sounds we produced (...) "[5].

### III. AUDITORY IMAGES

The nature of electronic music yields a distinct form of communication between the composer and the listener. The listener, as well as the composer, experience electronic music within a broader domain of cognitive associations to sounds. With the introduction of sounds that are not distinguished by the listener as being musical, the cognitive response to genre involves a distinct process of meaning attribution. Increasingly, it is emphasized that the ways in which the listener experiences meaning of electronic music are similar to the ways in which the mind constructs meaning in everyday life. The perception of environmental sounds cannot be described by the means of the traditional musical language. Similarly, the features of sonic events that are used in electronic music fall outside the vocabulary of this language.

There are some important observations on the perception of environmental sounds. It is believed that our categorization of auditory stimuli is based on semantic features rather than perceptual ones. It seems that the emergence of meaning in the everyday mind binds to the cognitive processes, which integrate experience from our multiple perceptual domains and forms a centralized understanding of the world around us. Listening, also listening to electronic music, gives rise to meaning as a direct manifestation of the everyday cognitive process of finding meaning in the world around us. An experience in one domain almost always invokes connections with another. In turn, electronic music certainly uses knowledge from other domains as part of its meaning.

It must also be noted that meaning varies within the context. And therefore there may be kinds of meanings that can be explained in common language and shared or they may seem more idiosyncratic and personal. Gary S. Kendall argues that "Sounds may take on particular meaning depending on their context. They can take on the role of opening, closing, marking boundaries, and so on. (...) another source of artistic meaning is the highlighting of novel sounds in the context of the art work. For example, electroacoustic art can strive to bring the acoustic background into the foreground of the listener's attention. These sounds could be unnoticed or

simply inaudible." [7]. One must also remember that ubiquitous contextual variation in meaning "is an inherent component in the meaning construction process" [8].

On the other hand, the essence of electronic music is often the use of the purely sonic properties of the sound material for artistic purposes. Listeners receive a certain message issuing from the composer effectively when they respect the rules for the interpretation of the semantic units contained in that message. Hence the communicational competence between a composer and a listener often boils down to the ability to recognize communicational situations and the composer's selection of appropriate means for the creation of an audibly perceptible sound shape. This audibly perceptible sound shape results from the composer's auditory representations, that is to say, from his auditory ideas (arising in the composer's mind). In turn, the configuration of these ideas is the basis for the mental representation of the musical work in the mind of the listener while consciously listening to music.

Electronic music is a difficult object of perception, because while listening to it the perceiver experiences the auditory images of highly complex musical structures. These produce highly complex and variable impressions, which for the listener may cause difficulties with perception. During the auditory reception of music, the listener makes a kind of selection of impressions, and extremely important here is the process of learning. The listener learns to differentiate the distinctive features of impressions. When a given piece of music is "heard through" many times, the ability to differentiate these features is enhanced, thanks to the process of learning.

Auditory perception is based on the identification and differentiation—through observation—of the sound images of heard music. The task of identification essentially involves the listener seeking the mental representation of a given sound image. Differentiation, meanwhile, concerns the assessment as to whether two sound images derive from the same configuration of musical structures or from different configurations. Research into perceptual processes in respect of the identification and differentiation of sound images indicates that these are processes for which different areas of the cortex are responsible. We conclude from this that among potential listeners to a given piece of music there may exist different levels of ability to identify and differentiate its sound images. This is conditioned by the cortical properties of their brains, since there may arise certain anatomical differences between the two hemispheres.

The retention of simple information obtained as a result of sensory registration takes place in the store of sensory information, which enables this information to be kept in the cognitive system until the phase of emotional assessment and then the phase of the recognition of the content of the stimulus (perceptual categorization). In perceptual categorization, the listener seeks in his mind the category to which the incoming information is best suited. The experienced listener also effects a schematization. Categorization and schematization are not mutually exclusive processes; indeed, they are

mutually complementary, since the construction of the standard that serves as a model in memory recall is based on the memorization of an ever greater number of the distinctive features of the heard music, which are employed in their integration into a schema. Cognitive psychology links categorization with schematization in the sense that it treats the schema as a part of the perceptual process, a part which—as Neisser states—“is internal to the perceiver [listener], modifiable by experience and somehow specific to what is being perceived. The schema accepts information, which are received by sensory surfaces, and is changed by the information” [9]. This reasoning contains the suggestion that there exists a certain set of innate perceptual elements, within which, besides the senses, there also exist schemata that control them. The schema, meanwhile, acts as a prototype for the classification of patterns into well-defined categories.

The everyday mind characterizes and organizes much of the activity in the world around us by utilizing the term *event*. This concept has quickly become a metaphorical way of organizing and understanding the complexity of our auditory experience and participates in the construction of meaning in electronic music—not only in relation to the listener’s understanding of single sounds, but also the understanding of groups of sounds and compositions as wholes. Universal properties of events can be captured in a cognitive schema.

The listener, in the phase of perceptual categorization, makes use of a certain store of ready-made innate schemata and of schemata of “generalization” employing prototypes which have the character of arithmetic means or moral values (the most frequently occurring images). In each case, the schemata, through their relative constancy and their connections with other parts of the auditory experience, enable the listener to recognize sound images which on the surface resemble nothing that the listener could have previously perceived.

In everyday life, the experience of events is interwoven with the flow of felt experience in its “distinctive qualities”: tension, linearity, amplitude and projection. The everyday mind relates the dynamics of energy flow to felt experience, an innate part of every listener’s feeling capacity and an intrinsic component of how the everyday mind experiences meaning. The embodied basis of such energy flow explains how listeners so easily relate sonic events to physical gestures. The felt qualities of experience are clearly recognized by listeners. According to Kendall, what is important here, is that “After the conclusion of an *event*, access to that *event*’s history is limited to what was bound to the *event*. (...) Often what is retained is a general characterization of the flow dynamics, a sense of the texture of the energy flow that could be captured with words such as *rough*, *bumpy*, *grainy*, *smooth* or *flowing*. When recalling the event, the listener may more easily retrieve an overall sense of the *event*’s energy than the pattern of its flow.” [7].

In the case of the perceptual categorization of the sound images of electronic music (although not only), the role of invariants shaping a prototype may be played by the repetitiveness of a particular configuration of sensory features,

which determines the qualitative relationship among sound images (linked to such parameters of the musical work as performance means, articulation, the pitch range of the musical material employed, dynamics and agogics).

The organization of memory and the limits to our capacities for remembering have a strong influence on the way we perceive, and consequently on the shape of sound images and their temporal boundaries. Memory affects the listener’s decisions as to when given sound images or their configurations end and others begin, and also as to the way in which sound images are connected with each other in the mental representation of heard music. Elements of musical structures occurring 50 milliseconds apart (which corresponds to a frequency of 20 sound events per second) accumulate, creating a level of blending of the sound events (often termed their “fusion”). In the auditory experience, this level is linked to the forming of the sensory features of musical structures, and the boundaries of this level result from limits as to the speed at which the neurons can process incoming information [10].

The musical structures responsible for sound events occurring at intervals greater than 63 milliseconds participate in the second level of auditory experience. They are separately distinguishable, but are not so far apart from one another on the timescale as to cross the time limit of short-term memory (ave. 3–5 seconds per event). In the case of electronic music, these levels can be defined as levels of the forming of timbre and time patterns. Within the range of this level, there ultimately occurs the successive, simultaneous or successive-simultaneous grouping of elementary events endowed with particular sensory features. Proximity of pitch-range, of time, a similarity among the dominant sensory features, a common type of motion and continuity to the flow of the music determine the integration of these events, which are subject to the above-mentioned criteria for grouping. The main difference between the level of grouping and the level of blending of events is that on the latter level the listener registers the boundaries between single formed sound objects with specific sensory features, whereas on the level of grouping he registers temporally expanded patterns comprising numerous events. The difference concerns the timescale. The boundaries between events on the level of blending have an immediate character, not exceeding the length of ultrashort (sensory) memory, but at the level of grouping, the events are sufficiently extended in time that short-term memory is required for their perception. However, these temporal differences are not absolute. Moreover, the duration of the recollections from sensory and short-term memory may coincide [10]. This is also the level of the forming of the listener’s sound images and the composer’s auditory ideas.

The musical structures responsible for sound events occurring at intervals greater than 63 milliseconds participate in the second level of auditory experience. They are separately distinguishable, but are not so far apart from one another on the timescale as to cross the time limit of short-term memory (ave. 3–5 seconds per event). In the case of electronic music,

these levels can be defined as levels of the forming of timbre and time patterns. Within the range of this level, there ultimately occurs the successive, simultaneous or successive-simultaneous grouping of elementary events endowed with particular sensory features. Proximity of pitch-range, of time, a similarity among the dominant sensory features, a common type of motion and continuity to the flow of the music determine the integration of these events, which are subject to the above-mentioned criteria for grouping. The main difference between the level of grouping and the level of blending of events is that on the latter level the listener registers the boundaries between single formed sound objects with specific sensory features, whereas on the level of grouping he registers temporally expanded patterns comprising numerous events. The difference concerns the timescale. The boundaries between events on the level of blending have an immediate character, not exceeding the length of ultrashort (sensory) memory, but at the level of grouping, the events are sufficiently extended in time that short-term memory is required for their perception. However, these temporal differences are not absolute. Moreover, the duration of the recollections from sensory and short-term memory may coincide. This is also the level of the forming of the listener's sound images and the composer's auditory ideas.

Configurations of the succession of sound images extended for a time exceeding the limits of short-term memory create mental representations of music that refer to the formal level of its experience. This level of experience is most often described by the listener metaphorically, the metaphors concerning movement within a physical space. Movement in large physical spaces requires the use of long-term (permanent) memory. The listener says of groupings of sound events formed in this way that they are "earlier" or "later", and describing the auditory experience he indicates that he becomes "lost" in the music, "finds himself in a particular place." The formal level and its articulation are linked to the structure and limitations of permanent memory. In contrast to patterns existing on the level of grouping, in which short-term memory is engaged, segments or sections on the formal level exist within a timescale that is too great for memory to be able to embrace them all "in the present."

According to Kendall "Whatever the context, the demands of the moment often surpass the listener's mental capacity to fully assimilate what is heard. But listeners are able to grasp the gist of what they hear, even when the details are too complex to follow. *Gist* is a working sketch for meaning, a snapshot of essential relationships. In one sense *gist* enables the listener to keep up with the essentials of ongoing experience and in another sense it is a product of the listener's understanding of what is essential in the current context. *Gist* separates foreground from background, the prominent from the inconspicuous. What *gist* does not generally include is a sense of outcomes or consequences" [7]. Additionally, broad sequences of events on this level do not automatically conserve their order in time. This order must be reconstructed—it is not a given characteristic. Thus for the links between the configurations of sound images on the

formal level to be discovered, they must find their way, at least in part, into consciousness (by being summoned or recalled) from the permanent memory [10]. Rapid auditory events in electronic music—similarly as sound events in everyday life—occur within a short time frame. In this case it seems that the *gist* guides the listener to the level of organization that is the primary carrier of significance and meaning—not the level of the individual auditory events, but the composite of the whole that is heard as a single event.

At this point, it is very important to realize that all the listener's experiences from the above-described levels are temporally interconnected as they occur. In actual fact, the different levels of experience are no more than differences between individual ways of processing information in the memory. Therefore, as Kendall emphasizes, *gist* and events are part of the listener's meaningful understanding.

Also important in music cognition are memory processes which enter the realm of non-declarative memory (implicit knowledge, which occurs automatically and exclusively in such a context in which it was assimilated). These are priming, sensitization and habituation.

With the effect of priming, each stimulus presented sufficiently early contains a trace that modifies the correctness and facility of the recognition of stimuli appearing at a later time. Priming acts bi-directionally, and so the identification of further stimuli can be made easier or more difficult [11]. Sensitization increases the attention devoted to stimuli which differ—be it only in some minor detail—from stimuli previously memorized. The memory has the chance to alter the information it contains. Habituation, meanwhile, is the opposite of sensitization, as it involves the reduction of attention devoted to the analysis of stimuli that are already familiar [10]. Although it does occur that one of these memory processes dominates, barring exceptional situations these processes mostly act together, responsible for the evolution of our memory.

Finally, it is worth mentioning that, in the creation of meaning in the cognitive paradigm, of importance are such contextual clues as the composer of the work, knowledge of previous works that have been heard, attitudes toward sonic material, and so forth, and also, the confrontation of the meaning in relation to the spatial idioms in audio reproduction, which are linked to the technology of electronic music, creating some unique possibilities for the artistic meaning.

#### IV. AUTOMATED VISUALISATIONS

While research on sound perception up to the last decades of 20th-century has focused rather on the assessment of physical attributes of sounds for sorting tasks, recent experiments based on event perception underline the fact that source identification subdue psychoacoustic features of sounds in categorizing complex stimuli. The "cognitive reconnaissance" introduced in this section of the paper aims at exploring some of the possibilities that open up when musical research has been extended to include the techniques of automated analysis of music which can be applied together

with the MIR (Music Information Retrieval) techniques to analyze the musical material discussed in paper. The method of analysis by the means of an automated visualization, that is used here, generates in objective manner the visualizations, thus giving us an insight into the auditive shape of the musical work. Dynamic visualization refers, for instance, to the shaping of the pitch, rhythm and timbre of sound with the purpose of producing an analogue of the process of musical perception. Time-frequency (spectral) visualization makes it possible to get insight into the complex features of musical structure. This analysis should verify whether the forms of mental representations of the musical surface of the analyzed work correspond to the form as it was imagined by the composer.

Sonic Visualiser is an application for viewing and analyzing the contents of music audio files. The work SV does is intrinsically processor-hungry and (often) memory-hungry, but the aim is to allow working with long audio files.

*Mikrophonie I* had its premiere on the occasion of the music festival *Reconnaissance des musiques modernes* in Brussels on 9th December 1964. In this research was used the audio file recorded on the CD from Stockhausen-Verlag. This is the *Brussels version*, i.e. the version of the work that sounded in the world premier by Aloys Kontarsky, Alfred Alings (tam-tam); Johannes G. Fritsch, Bernhard Kontarsky (microphones); Karlheinz Stockhausen (filter operator and potentiometer controller 1); Jaap Spek (filter operator 2); Hugh Davies (potentiometer controller 2, aiding Spek).

The sounds we hear are mostly noises, i.e. sound events of undefined pitch, according to the tam-tam's sound character, where it occurs, is mainly generated by the processing of sound through microphone/filter as the composer describes. Alternatively, pitch is defined by the implements used and their interaction with the surface of the tam-tam.

From the point of view of the perception of the entire work one can distinguish two general features. Firstly, *Mikrophonie I* is characterized by extremely stretched musical space. This is the effect of moving between two extremes. On the one hand there is the strong inward concentration in many of the "moments" as a result of the close-held microphones, which allow the conversion of inaudible vibrations to sound. On the other hand there are the passages that powerfully explode through loud sounds of the tam-tam.

Secondly, there is a kind of the tension game. This has emphasized by David Huron's theory of expectations—Imagination-Tension-Prediction-Response-Appraisal = ITPRA [12]. Tension prepares listener for the imminent event through arousal (motor preparation) and the focusing of his attention (perceptual preparation). The tension is palpable in *Mikrophonie I* in a strong degree. One can observe the specific action-reaction game, in which one musical "moment" responds to another, and which forms a continuous moments chain through the music. The tension that causes the expectation of the response of the given musical "moment" to the sound and specific motion of the preceding one is due by the succession of events, by the trajectory of energy change, the decay of sound and also by thought-out selection of pauses

and their duration. Through tension a continuum between sounds is created.

The visualization of acoustic signals can take several forms, but not all of them are equally useful for the musicological interpretation of auditive experiences. For example, the so-called temporal representation of the signal, treated as a mathematical function of time, reveals many details concerning the soundwave, but their visualization is not fully compatible with the auditory images of that signal. For this reason, the visualization of an audio representation of analyzed music also takes the form of spectrogram (sonogram), that is, a temporally variable spectral representation that shows in the form of an image how of information about the analyzed musical work, that one can obtain from the visualization of its acoustic representation, depends on the software we use. We currently have the possibility of retrieving from spectrograms such information about the analyzed music as a profile of the shape of the sound wave and data about phrasing, segmentation and pitch components.

The indicator of the specification of the timbre of *Mikrophonie I* is representation, which—as the result of temporal parameterization—shows the amplitude of the analyzed signal as a function of time and define the dynamic timbral profile of isolated sound events (the moments). The audio waveform of the musical work shows the overall shape of the signal. This is particularly useful for showing the evolution of the signal over a longer time and it is used specifically for detecting sound events. The shape of the musical surface in this case—*Mikrophonie I*—is compatible with continuously varying processes and accumulated power.

The work starts itself from three structures (number 1, 2 and 3) which are similar in respect of their duration and—what is more important—of their sound content. Stockhausen described these "moments" as, respectively, Quakend (for the first group of performers), Knistern-Hacker (for the second group of performers) and Winselnd-Jaulend (for the first group of performers, again). Composer wanted to create the impression of continuous succession of short events, which on the visualization (Fig. 2) are shown by similar shapes of the curves energy with short duration of attack and slightly longer duration of the decay of sound. The relationships between all three structures are supportive each other, although the relationship between structure 1 and 2 is constant, and between structure 2 and 3 is increasing, because the latter has got more complex content. Sounds there are more drawn-out and sometimes overlap each other, what is linked with the names of "moments"—Knistern-Hacker vs. Winselnd-Jaulend. This fragment of *Mikrophonie I* exemplifies the kind of noise with very dry and direct sounding. The dryness of sound seems to be the outcome of high amplification of very low-level signals. In such situation the articulation of the structures doesn't excite the plate's resonance, whereas the generated sounds are picked up by a close-held microphone. After structure 3, the pause appears (the line on the visualization), and its role is to create dramatic effect, before the advent of the structure 4.

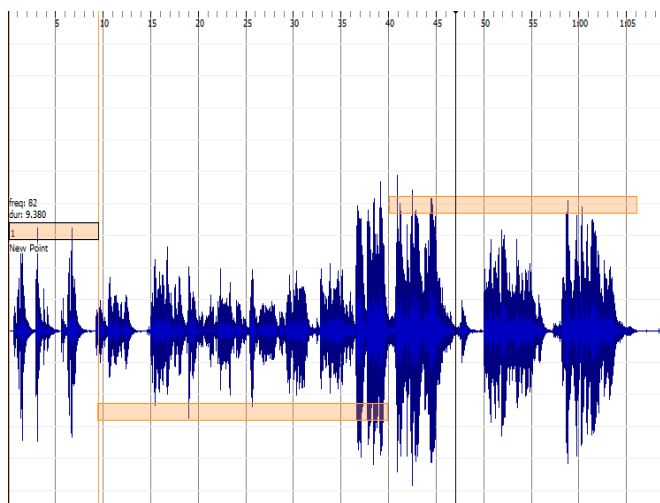


Fig. 2 *Mikrophonie I*—the audio waveform (structure 1-3 and pause)

The audio waveform of *Mikrophonie I* confirms its polyphonic texture with changing density. For example, from the structure 4, which is opposite to previous one, and at the same time it is destroying in relation to previous relationship, inducing the impression of increasing tension, *Mikrophonie I* becomes more polyphonic. Different kinds of sound are played simultaneously. On the visualization of structure 4 (Wirbelnd-Trommelnd-Knarrend, for the second group of performers, Fig. 3) and 5 (Rasselnd-Ächzend-Donnernd, for the first group of performers, Fig. 4) one can note continuous chain of dense and more volume of audio waveforms with sharp attacks. On the top layer of the visualization there is the form of spectrogram, in which the all bins have shapes of wavelike oscillations caused of swelling and deflating of deep sounding plate resonance. These sound motions are generated by repeated cycles of moving the microphone towards, and again away from, the plate of the tam-tam.

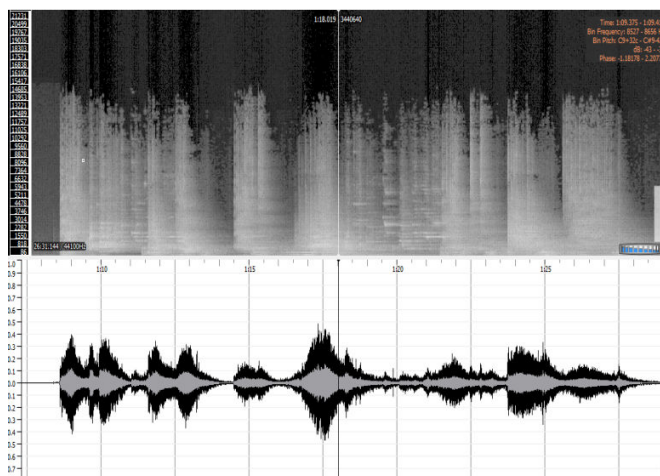


Fig. 3 *Mikrophonie I*—the waveform and spectrogram of structure 4

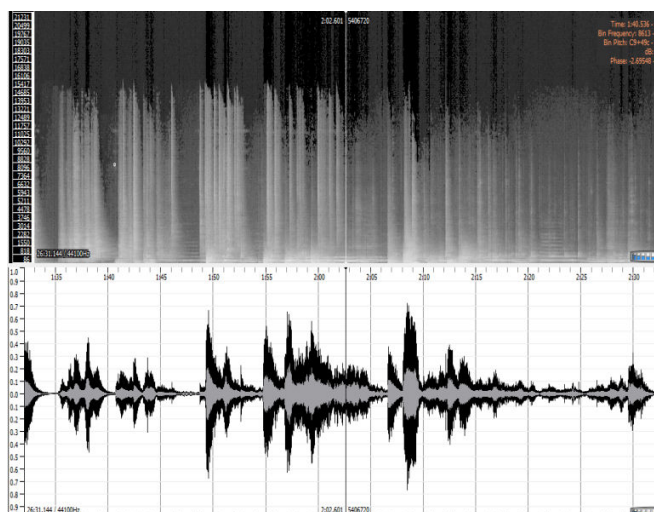


Fig. 4 *Mikrophonie I*—the waveform and spectrogram of structure 5

Sometimes, when the “moment” of *tutti* (structure 19, “Tutti 157”) plays a special role in musical narration, then the polyphonic sounds are highlighted and enhanced as the result of their rendering in stereo panorama. The idea of Stockhausen was to *Mikrophonie I* tends, as it unfolds, to move from more similar “moments” to contrasts. However, only in this “moments”—structure 19—both groups of performers come together in a musical structure, which stands out from others, because all kinds of “moments” are played and heard at once and in succession. One can see how complex and complicated the graphics in the score are for this structure (Fig. 5).

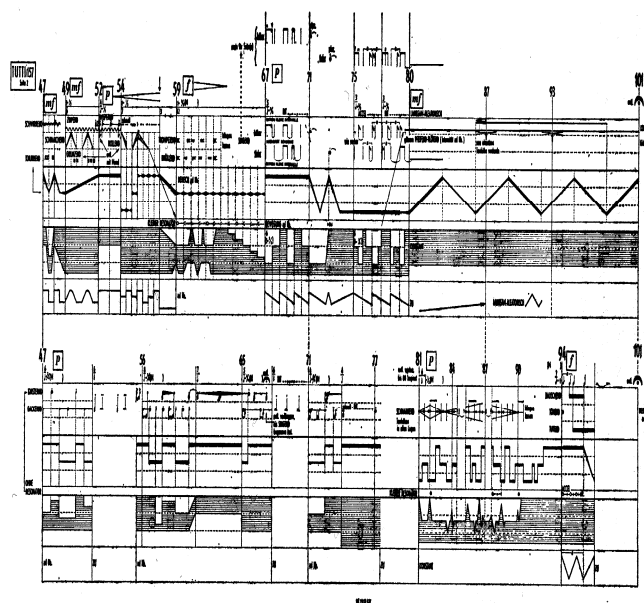


Fig. 5 *Mikrophonie I*—the fragment of score (structure 19) [3]

On the visualizations (Fig. 6) there are: the audio waveform (top layer), spectrogram (middle layer) and melodic range spectrogram (bottom layer). All layers show very complex



temporal patterns. In the range of structure 19, quite often there exist a flow of sounds in the sense of traditional fluidity. These sounds are responsible for directional shape of this fluidity and for shape of surface of melodic range spectrogram. On spectrogram in turn one can distinguish different, regarding to density, places, with more linear passages. All is correspond to the presence of all kinds of “moments” in structure 19 and all kinds of excitation of tam-tam.

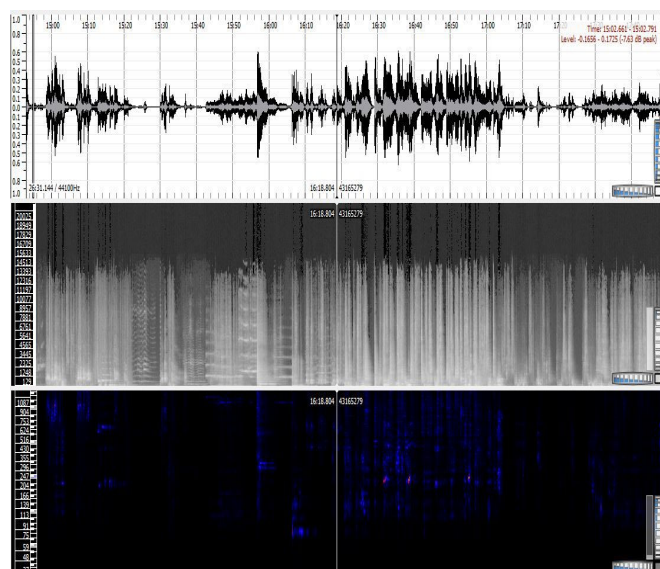


Fig. 6 *Mikrophonie I*—The visualization of structure 19

But—as a contrast—the concluding passage with fade-out (structure 33, Trompetend-Brüllend for the first group of performers with continuation of structure 32 for the second group of performers—Pfeifend-Flötend—) includes “moment” with excitation of tam-tam through rubbing of the edge of its metal plate by means of bow. Additionally, the surface of tam-tam are scratched or rubbed by the first group of performers by means of various kinds of objects, and in the same time the sounds are picked up by changing distance of close-held microphones. Together with the start of fade-out, the movements become on the tam-tam more and more slow and gentle.

On the visualization of structure 33 (Fig. 7) can be seen the gentle curves of energy, flatter trajectories of attack with a stretched-out decay of sound. Only twice the curves of energy of waveform are increasing, but its duration is very short in compare to the duration of whole moment. In turn bottom layer indicates of the peak frequencies spectrogram, which is specific because of the very scattered events in pitch space and time. Such content of spectrogram confirm the nature of sound events appearing in the range of this final structure.

But there are also some moments with pauses. Particularly in structure 12 (Fig. 8, Ypsilon, for the second group of performers) there are short events of very gradual, calm and quietly generated, and timbre and motion of sounds create of the passage, which is similar to previous one, although

following structure is very contrast to it. The relationship between structure 12 and neighbors is constant, but the content of it is destroying regarding to structure 11, and is neutral to structure 13.

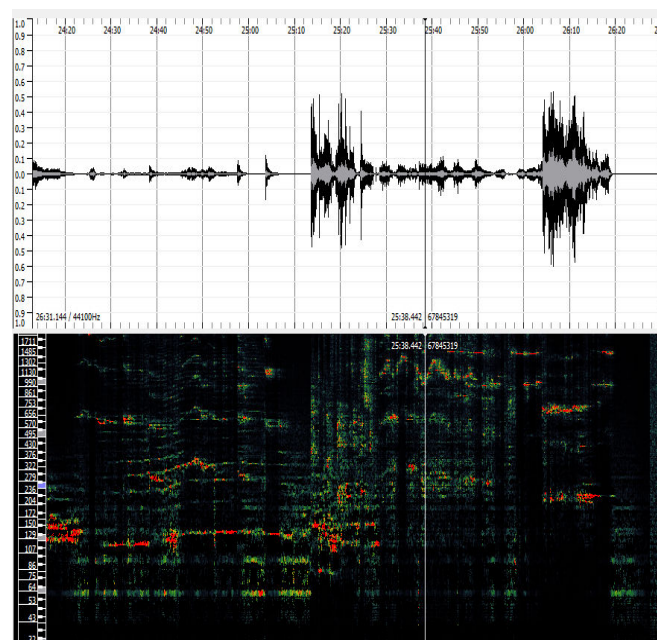


Fig. 7 *Mikrophonie I*—the visualization of structure 33

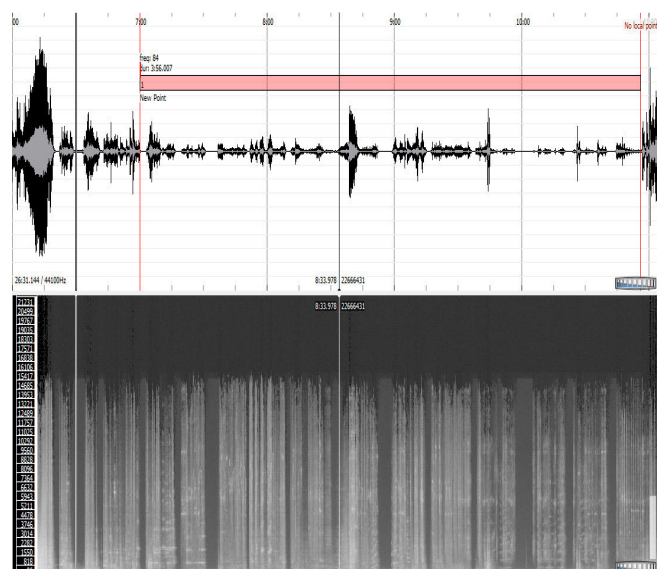


Fig. 8 *Mikrophonie I*—the visualization of structure 12

On the visualization of structure 12 is shown the range of it, and also the end of structure 10 (this is the climax of it), the whole structure 11 and the start of structure 13. Such content of the visualization aims to show the relationships between them. Moreover, spectrogram on the bottom layer confirms the specific nature of structure 12. One can say that there is the non-directional moment with the regular, stretched shape of waveform and stable cyclic content of spectrogram. The

shapes of the curves energy on waveform and all bins on spectrogram confirm the presence of short events. Additionally, vertical bands on spectrogram, with different darker color, correspond with the pauses appearing in the range of this structure.

## V. CONCLUSION

With *Mikrophonie I*, Stockhausen opened the way for a more flexible use of electronics than what was possible with the serial techniques he employed in the 50s, and this piece for real time electronics became the first (in Stockhausen's work list) in a number of such works written during the 60s and 70s.

The density and complexity of musical structures meant that they were not designed to be grasped in such a way as they unfolded in real time during the performance. Stockhausen found a new way of thinking and working on his music in organizing musical continuity and form.

Comparing Stockhausen's statements and notes-the form of auto-reflection in regarding to ideas and creative process-and the automated visualizations of auditory images of *Mikrophonie I*, one can note, that between them, there is an analogy. This analogy confirms not only a huge consciousness of composer regarding to possibilities of creation, but also a huge acoustics and technique knowledge.

Music of this density and complexity demands to be heard repeatedly, or to be studied outside of performance time. The question is, how musical context can affect the listener's way of perceiving the musical work, and, how can musical form, whatever it is, continue to exist during the musical experience. To find the answer and the knowledge, one must know what kind of compositional ideas were born during the creative process.

## REFERENCES

- [1] J. Dack, "Karlheinz Stockhausen's Kontakte and Narrativity," *eContact* 2.2, ix 1999. (Accessed 20 April 2015) <http://cec.sonus.ca/econtact/SAN/Dack.htm>.
- [2] E. Zanpronha, "Gesture in Contemporary Music: On the Edge between Sound Materiality and Signification," *TRANS-Transcultural Music Review* 9 (article 17), 2005. (Accessed 20 April 2015) <http://www.sibetrans.com/trans/article/181/gesture-in-contemporary-music-on-the-edge-between-sound-materiality-and-signification>.
- [3] R. Maconie, *The Works of Karlheinz Stockhausen*. London: Oxford University Press, 1976, pp. 187, 184.
- [4] R. Toop, *Six Lectures from the Stockhausen Courses Kürten 2002*. Kürten: Stockhausen-Verlag, 2005, p.179.
- [5] K. Stockhausen, *Stockhausen on Music. Lectures and Interviews Compiled by Robin Maconie*. London, New York: Marion Boyars, 1989, pp. 78-80, 82-83.
- [6] R. Maconi, *Other Planets. The Music of Karlheinz Stockhausen*. Lanham, Maryland, Toronto, Oxford: The Scarecrow Press, 2005, p. 257.
- [7] G. S. Kendall, "Meaning in Electroacoustic Music and the Everyday Mind," *Organised Sound*, vol. 15, no. 1, pp. 63-74: 65, 66, 68, 2010.
- [8] S. Coulson, *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press, 2001, p. 17.
- [9] U. Neisser, *Cognition and Reality: Principles and Implications of Cognitive Psychology*. San Francisco: W. H. Freeman and Co., 1976, p. 54.
- [10] B. Snyder, *Music and Memory*. Cambridge, Mass.: MIT Press, 2001, pp. 23-25, 31-45, 69-72.

- [11] R. Sternberg, *Psychologia poznawcza*, trans. Ewa Czerniawska and Anna Matczak Warsaw: WSiP, 2001, p. 69 (Eng. orig. *Cognitive Psychology*, 2nd. edn (London, 1999)).
- [12] D. Huron, *Sweet Anticipation. Music and the Psychology of Expectation*. Cambridge, Mass.: MIT Press, 2006.

**Justyna Humięcka-Jakubowska** is an assistant professor at the Department of Musicology of Adam Mickiewicz University in Poznań (Poland). Also graduated from the University of Technology in Poznań (electrical engineering). She received her PhD in musicology from Adam Mickiewicz University, Faculty of History in 2005, and post-PhD degree in musicology from Adam Mickiewicz University, Faculty of History in 2014. She specializes in the areas of analysis, history, theory and aesthetics of 20th-century music and also music perception, cognition, music acoustics and physiology of hearing.

She is the author of two books: *Scena słuchowa muzyki dwudziestowiecznej* [The Auditory Scene of 20th Century Music] (Poznań: Rhythmos, 2006) and the edition of her postdoctoral thesis: *Intuicja czy scjentyzm: Stockhausen – Berio – Nono – Ligeti – Xenakis – Grisey* [Intuition or scientism: Stockhausen – Berio – Nono – Ligeti – Xenakis – Grisey] (Poznań: Publishing PTPN, 2013). She has also written a number of articles, which are published in musicological journals and also conference proceedings.