World Academy of Science, Engineering and Technology
International Journal of Humanities and Social Sciences
Vol:12, No:9, 2018

The Creative Unfolding of "Reduced Descriptive Structures" in Musical Cognition: Technical and Theoretical Insights Based on the OpenMusic and PWGL Long-Term Feedback

Jacopo Baboni Schilingi

***Abstract*—**We here describe the theoretical and philosophical understanding of a long term use and development of algorithmic computer-based tools applied to music composition. The findings of our research lead us to interrogate some specific processes and systems of communication engaged in the discovery of specific cultural artworks: artistic creation in the sono-musical domain. Our hypothesis is that the patterns of auditory learning cannot be only understood in terms of social transmission but would gain to be questioned in the way they rely on various ranges of acoustic stimuli modes of consciousness and how the different types of memories engaged in the percept-action expressive systems of our cultural communities also relies on these shadowy conscious entities we named "Reduced Descriptive Structures".

***Keywords*—**Algorithmic sonic computation, corrected and self-correcting learning patterns in acoustic perception, morphological derivations in sensorial patterns, social unconscious modes of communication.

I. Introduction

THE neurophysiological study of living beings highlights the incredible complexity of our cognitive perception-reaction systems and the structural refinement of the organs of animal physical activity. Man is one of the living creatures that has developed the most elaborated interpretative and symbolic skills of the animal kingdom. The construction of its social habitat and of its cultural environment is the result of a dynamic and interwoven activity between a techno-cognitive and a symbolic development. It is the ability of developing symbolic languages that distinguished our species from others in its modes of adaptation to its natural habitat. Language is a field of individual and collective experience that opens the possibility of a de-pragmatized curse of existence, partly detached from a simple instinctive reactive behavior to exterior stimulation. This symbolic apparatus is inherited from a genetic and natural basis that, firstly has to be understood as a socioculturally determined environment and, secondly, if we needed to give out some broad figures that have now been well studied relays on a specific cognitive system: a neurophysiological system of more than 80 billion interconnected neurons [1], with a network of several hundred billions synaptic connections, all irrigated with information in the order of almost 50 meter per second. Human life is peculiar, one already notices, from the point of view of its modes of learning defining us as an original species –that sociocultural dimension– and which deeply determine the quality of our living logics in our material environments and through the collective way the transformations of conditions for the elaboration of life in common are thus lead.

The uncountable number of semio-cognitive implications of Man within its life environments encounters in the field of listening technics and auditory phenomena some operational logics particularly developed and interesting to analysis: the intricacy of the relations between the subconscious, the unconscious and the conscious dimensions of sensorial "mind processing". Exogenous and endogenous information in relation to their situation of corpo-phenomenological emergence unfolds theirselves according to an extremely fine cultural and natural creativity.

In this analysis perspective, we speak about recent research in neuroscience as well as in hermeneutics of listening converge on certain traits. The phenomena of formal and substantial categorization of acoustic stimuli respond to at least three types of very specific learning processes involving different memory games and a specific range of interpretative reactions, or so to say, of the cognitive processing of "environmental" information. The first effect of acoustic exposure reveals the establishment of an essentially subconscious treatment of this environmental sound material. Its efficiency threshold is deployed in a time-lapse below the threshold of conscious perception. The auditory subconscious is a necessary area of sensory filtering and information retention allowing, before exceeding a half-second according to the starting period of the stimulation own timescale, and operates a first selective identification within the dynamic information medium housing our existence. This first a-phenomenological layer of consciousness does not intervene in the historical writing of our common futures; this reflex zone constitutes in a tacit inheritance, "muted", and in a taxic dimension in the evolution of our human species. The perceptual logics attached to it do not participate, beyond a general principle of historical adaptation of the living, in the elements of our cultural and conscious *mise-en-scène* of the listening capabilities and, to finally seize the subject which concerns us fundamentally, not either in the "musical" understanding of situations and experiences that universally

Jacopo Baboni Schilingi is with the Conservatoire de Musique du Pays de Montbéliard, France (e-mail: jbs@baboni-schilingi.com).

inhabit the human worlds. Nevertheless, the subconscious is not entirely absent from the logic of cultural and musical interpretation of the hearing data process: it even supports its possibility because of the cognitive economy of the information processing that takes place there and because of the overture that it authorizes in terms of semiotic or semantic analysis of these sound stimuli. The subconscious consists in a neurocognitive infrastructure that is itself dependent of the anthropological trajectories of the material design and acoustic exposure of different environments. But by the fact of its non-integration to the domain of symbolic codification of its processes of information treatment, the subconscious cannot be considered other than as a zone of cognitive development related rather to the evolution of the human species in global terms and not so much as integrated in the logics of historical-cultural formation.

The field of historico-cultural formation of the hearing, being an issue as much individual as a collective one, finds its relevant scene in terms of perception in the lapse of time lasting the time of 40 milliseconds in the cognitive experience of exposure to sound stimuli. Without reducing its extreme complexity, we can nevertheless identify the stakes, differentiated but having related parts, of the two polarities that are those of the memorial and interpretative dimensions, which are respectively the domain of the psychic unconscious and that of the mediated consciousness. What is first called the unconscious is an area of symbolic and physiological memorization of the relationship to our living environments. This memory, however, is extremely expensive to revive directly through a voluntary intellectual effort, but not inaccessible so far, as evidenced in particular by the practices of free association and hypnosis promulgated by psychoanalysis. These technics are nowadays considered for their subjective effectiveness as well as for the objective anchoring of the events they participate to reveal in this memorial and interpretative stratum of the psychic unconscious. This cognitive zone of the unconscious is a very powerful analogic and metaphorical processing area. The information which penetrates there preserves the temporal imprint of the moment which saw them appearing, but their significant forms, their sense (functionally and symbolically speaking), undergoes a deep reconfiguration. These mnemonic traces can “come back to life” under formal or substantial association of extremely diverse ideas.

Finally, we must mention the different cognitive and interpretative processes that are part of what is called "consciousness", that form of thought that emerges from or apart from the two previous domains of cognition that we have just described. Consciousness is the area of reflexive and procedural processing of information that comes to us from our natural and social environments. Significant shaping of stimuli develops mainly in a medium-term direct memory processing and from a relatively short-term working memory process. It is the area of cognitive activity, the most directly determined by the historical and cultural dimensions of an individual's existence.

The complete and profound apprehension of musical listening cannot be undertaken without this analysis of the cognitive economy that governs its conscious, unconscious and subconscious experience. The music, its forms, its repertoires, the whole spectrum of the auditory and transcendental experience that can be made out of it, involves an extremely complex branch of connections and distribution of information in terms of cerebral cognition, involving the three temporal "floors" and/or types of memorial and interpretive infrastructures that we have just described. Music, a complex multi-sensory and expressive phenomenon, relies, in its modes of cognitive integration and perceptual learning, on a particular exchange of cognitive information between these different domains of thought, a process that doubly involves the body, its attitudes, its sensory arrangements and instruments as much as the spirit and its operations of symbolization, subjectivation and objectification of concrete and material sensitive situations. It is precisely in this context that we conduct our theoretical work on the Reduced Descriptive Structures of the perception and the development, in computer-based music generation, throughout the software suites of OpenMusic and PWGL intervenes. We achieved this issue by proposing a programming model that reproduces two of the principal operations in the development of auditory percept-action structures interfacing between the different levels of memory, awareness, and significance of the acoustic events presented above: corrected and self-correcting classification or learning processes producing classes of eco-cultural aspects of perception and so, in terms of sound perception and musical recognition.

What we call "learning" is divided into two types of differentiated behaviors, depending on the fact that it involves or not, a socio-symbolic mediation in the process of acquiring knowledge. The second case mentioned, which we focused our attention, is the one called self-correct learning and is characterized by the autonomy of the system-actor involved in perceptual learning. This type of learning is complementary to the first type, the corrected learning, necessarily involving a socio-symbolic[1] or socio-technical[2] mediation. While the social and interactive structure of corrected learning is already the subject of a very rich and comprehensive analysis by sociologists, psychologists and anthropologists, self-correcting learning has for the last several decades gathered around itself and in an unresolved debate (still widely open in terms of interdisciplinary scientific analysis), both researchers in linguistics, systems studies (including cybernetic sciences), mathematicians and, with a growing actuality, computer scientists are developing artificial intelligence systems capable to simulate self-correcting learning abilities. Self-corrected learning thus consists, in synthesis, in an ability to discriminate formal structures through a situation of complex sensory exposure from the only cognitive resources of an

[1] Here we are obviously thinking of the role of language in the socialization of individuals to the imaginary worlds that give meaning to human cultures.

[2] Here, for example, we think of the principle of attentional guidance as a key-process in mediation of eco-perceptual knowledge in the global developmental process of young children.

individual considered in ecosystemic isolation. We can imagine the complementarity of these two postures of the learner as a personal journey rallying different postures of knowledge facing a series of empirical situations. Think about a dotted line, where the dots are the situation of corrected learning moments giving to the learner's parkour some kind of directional vector and to the lines between these dots as the individual, "isolated", path undertaken to link the next destination. This path is characterized by learning times where the signifying arrangement of a situation, and therefore of the pragmatic experience that is made of it, is sometimes strongly determined by knowledge and by knowledge mediated by institutional and culturally determined symbolic forms[3] rather than in the situation where this (time-mediated situation would be mostly) determined by the simulation and the individual mental projection of the possibilities of realization of a concrete action. This first learning marker, directly and decisively involving social and symbolic resources, is considered a one of a corrected learning situation. The second marker indicates another polarity regarding the exploitation of cognitive resources necessary for any development of knowledge, and thus for the establishment of a learning pattern. This second development demonstrates the idiosyncratic capacities for the elaboration of knowledge and a second pole for the rooting of learning processes, that of mental simulation, that of self-correction and the equally epigenetic dimension of the knowledge that can be made of the world. This world that reaches us, if we are so engaged physically and socially, but also that makes sense of the fact that we inject, by means of cognitive simulation, heuristic analysis resources that are specific (specifying) and characteristic (individualizing) for each individual. In other words, this particular double-flipped point-of-view

> "emphasize the growing conviction that cognition is not the representation of a pre-given world by a pre-given mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs" [2].

This second type of learning is today the object of an efficient and operational delegation to computer systems based on extremely powerful algorithmic calculation technologies and recursive computation structures based on the arrangement of virtual neural layers. Its implementation by machines rests, schematically, on the capacity of a closed system to infer meaning propositions, to meta formalize a message or information on the basis of different processes of heuristic, combinatorial and projective analysis. It is precisely, to stick to an epistemological and theoretical language, the nature of the functions configured at the heart of these algorithmic machines and devices of reproduction of the sense of musical listening mentioned previously and developed since the 1980s, within and from IRCAM and then taken over by the *Prisma* group [3]. These same technologies, applied to the analysis and the musical composition, allow us today to explore and to better understand the springs of these mechanisms of perceptive learning, in particular, those based on a principle of self-correction [4]. It is through the sequencing of complex sono-acoustic data and the generation of Reduced Descriptive Structures that these principles are concretely apprehended and implemented [5]. The Reduced Descriptive Structures correspond to data complexes allowing, by a phenomenon of radiation, to reach a given goal (an informational correspondence) by saving on the process of recomposition of all the possible chains of meaning that derive from the pragmatic sensory morphology [6] of a sound statement. The reduced descriptive structures are thus based on principles of essentialization of the acoustic morphology of these sound utterances.

In practice, the algorithmic sequencing of sound utterances allows the updating of the key principles of repetition and redundancy that structures a complex sound set. This analysis is not reduced to human speech, but integrates all different forms and modes of sound communication, whether artistic or not. The first experimental results of this automated procedure, which was carried out using the *fv-morphology* plugin in the PWGL environment, made possible to obtain concluding results from audio-recorded bird samples within the IRCAM laboratories-Paris.

The procedure consists of a discrete reduction of the whole extract of a sound utterance (bird song, musical melody, spoken speech) into a maximum of coherent and autonomous units. This procedure is necessarily dependent on the acoustic resolution of the sample as well as that, correlative at least to the part, of that of the digital encoding of the sample [7, p.4]. The rendering is presented as an ad hoc alphabet composed of all these singular sound forms composing the initial sample. This repertoire of isolated units is completed by the set of possible concatenations of "strings of sound characters" [8]. The second degree of algorithmic processing imposed on this alphabet of sound singularities and to the repertoire of their effective associative forms is to quantify what is called the editing distance, also called "distance of Levenshtein" separating these sequential statements [9]. This mathematical measurement accounts for the degree of similarity between two strings. A more intuitive understanding of the results of this second analysis of editing and overall distance at the scale of an entire sample is given by the graphical representation of a "tree of recovery". It illustrates the extent and compositional diversity of a sample with respect to the measurement of this editing distance.

[3] This is the case with knowledge acquired, educated or imposed by the force of authority, power or mimetic group reactions as in the case of crowd dynamics or even of certain political groupings. We thus think of the disciplinary instruction of knowledge in certain environments such as those of the "classical" school environment, prevalent in institutions such as the army, or responding to cultural institutions with strong asymmetries of the symbolic relations of power as in the industrial work environment, of male political domination in the public space or in the public policy space to cite only examples with the most stereotyped traits. These stereotypes should rather be used to illustrate an idea that they cannot accurately reflect concrete situations.

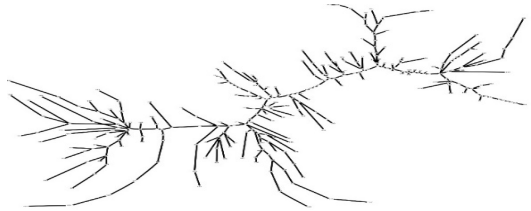

Fig. 1 Example of a "Tree of recovery" generated from a whistling bird's audio sample

Fig. 1 [10] shows us the network of branches and nodes that structures the coherent "internal" organization of a sound sample. How does this endogenous organization of content affect our perception and how does it relate to our different memory systems? If it is not conceivable, neither wanted, to close this question today, we would like to undertake and open an original line of reflection around the context of artistic creation. Considering that there is no creativity ex nihilo, and as far the debates on the "ontology" of works of art never escape the rule of their ideologization, it seems possible to open a third way for understanding the specific communication effects linking, on the one hand an artist-producer, and on the other a spectator-receiver, around a work of art in a effectively significant and operating way by considering the intermediation of Reduced Descriptive Structures in the individual cognition process and by according its role as a necessarily-rooted function in cultural reference classes generation.

The symbolic charge affecting any attention turned towards a work of art would effectively rest, in its pre-semantic and even pre-ontological dimensions, on the activation of this network of neuropsychological computation which characterizes much of the mental processes of consciousness and the sensory perception that determines any expressive possibility of a cultural object and therefore an artistic work. The communication system that is thus engaged then rests on a "natural" understanding –conscious, unconscious and perhaps partly related to our subconsciousness– of the internal organization of a work and, following this directly, to the generation of perceptual structures updated on the basis of analogical and experimental mathematical and computational principles [11]. These same principles illustrate, in a synthetic and controlled form, the algorithmic games that we have described throughout our discussion. Contrary to the thoughts that would see in this vision of communication's effects around cultural creation, an apology for the "repetition of the same", the hypothesis raised here is that what one calls creativity corresponds to the operating effects from the phenomenon of re-elaboration of the most distant parts of the original. That is, to stimulate the farthest "branching areas" from the nodes as represented by the "Trees of recovery".

II. Conclusion

It would be a profound mistake, both philosophical and epistemological, to confound in the analysis the processes of computer simulation, so advanced they have become today, with the principles of formalization of anthropological knowledge and the genesis of natural systems of perception that guide our cultural and day-to-day achievements. Nor is it useful, for our purpose here, to ontologize computer systems by infusing them with human characteristics. Computer science, and in particular here computer music, is gaining an expanding computing power, and its field of application and experimentation are gaining in power as well, as they extend in the scope of possibilities. The work carried out within the *Prisma* group has made it possible to develop experimental and didactic tools to reproduce, in a controlled framework, phenomena of sound cognition and thus to propose analyzes on the complexity of the logic of acoustic perception. We wanted to go beyond the framework of a laboratory proof-reading to open the field of reflection to artistic musical creation.

We argue that creativity finds expressive and perceptive resources as much through controlled learning processes, through different instances and techniques of symbolic mediation, but also that there are self-correcting processes of learning and sensorial classification operating in creative, artistic and musical communication. This second mode of formation is based on the generation of Reduced Descriptive Structures of perception throughout the life of individuals. These structures, represented graphically by the "Trees of recovery" and based on endogenous modes of organization of objective environmental data (acoustic sources), are all vectors of unconscious communication that innerves the lived and future landscape of aesthetic experiences for any individual and potentially connects it to different virtual communities of other people. Among these communities, some people find the cognitive and imaginary resources or/and mobilize specific technics to propose artworks, in which perceptive radiation echoes in others and animate this "latent and silent ear" inscribed in the *mille-feuille* of the conscious and unconscious memories of one-self's body and mind, but also in that, collective, of his sociocultural partners and co-participants.

References

[1] Herculano-Houzel, S. (2011). "Brains matter, bodies maybe not: the case for examining neuron numbers irrespective of body size", in *Annals of the Ney-York Academy of Science,* 1225, pp. 191–199.
[2] Varela, F.J., Thompson, E., Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*, MIT Press. p. 9
[3] Baboni-Schilingi, J., Voisin, F. (1999). *Morphologie: Documentation OpenMusic*, 3ème édition, Ircam, Paris.
[4] Mc. Adams, S., Deliège, I. (1995). *La Musique et les sciences cognitives.* Mardaga
[5] Giordano, B., McDonnell, J. & McAdams, S. (2010). "Hearing living symbols and nonliving icons: Category specificities in the cognitive processing of environmental sounds", *Brain and Cognition*, 73, pp. 7-19.
[6] Giordano, B.L., Guastavino, C., Murphy, E., Ogg, M., Smith, B.K. & McAdams, S. (2011). "Comparison of methods for collecting and modeling dissimilarity data: Applications to complex sound stimuli", *Multivariate Behavioral Research*, 46, pp. 779-811.
[7] Voisin, F. (2011). *Dissemblance et espaces compositionnels*, Conservatoire de Musique du Pays de Montbéliard, p.4.
[8] Deller, J. R., Proakis, J. G. and Hansen, J. H. L. (1999). *Discrete-time processing of speech signals.* Wiley-IEEE Press, New-York, p. 649-675.
[9] Myers, E. W. (1986). "An *O(ND)* Difference Algorithm and Its Variations", *Algorithmica*, 1, Springer-Verlag New York Inc., pp. 251-266.
[10] Voisin, F. (2011). *Dissemblance et espaces compositionnels*,

Conservatoire de Musique du Pays de Montbéliard, p.5.
[11] Chemillier, M. (2007). *Les mathématiques naturelles*. Paris : Odile Jacob.