DiMMI

Dictionary for Multidisciplinary Music Integration: Interaction

Trento, November 25-26, 2022

PROCEEDINGS

UNIVERSITÀ DI TRENTO

Silvia Sacchetti, Nicola Conci (eds.)

INTERACTION

Proceedings of the 3rd DiMMI conference, held in Trento, 25-26 November 2022

Università degli Studi di Trento

DiMMI. Dictionary for Multidisciplinary Music Integration

Vol. 1 - Interaction

The Dictionary for Multidisciplinary Music Integration (DiMMI) is a proceedings series about the event organized by the University of Trento and the Conservatory "F. A. Bonporti" of Trento and Riva del Garda, in which musicians and representatives of the academic world are called to reflect together on a word of common interest, each from the perspective of their own discipline.

Published by
Università degli Studi di Trento
via Calepina, 14 - I-38122 Trento
casaeditrice@unitn.it
www.unitn.it

Dictionary for Multidisciplinary Music Integration (DiMMI)
Vol. 1 - Interaction

Cover Design: Grafica > UniTrento | IMG > stock.adobe.com

ISBN 978-88-5541-044-1
DOI 10.15168/11572_398471

Table of Content

Dictionary for Multidisciplinary Music Integration (DiMMI) 2022: Interaction

Silvia Sacchetti[1] **and Nicola Conci**[2]

[1] *Dept. of Sociology and Social Research, University of Trento, Italy*
[2] *Dept. of Information Engineering and Computer Science, University of Trento, Italy*

Editorial

The Integrated Multidisciplinary Music Dictionary (DiMMI) is an annual conference in which scholars, researchers, musicians and practitioners are asked to interpret a single keyword within the perspective of their own discipline, seeking to promote the exchange of ideas at the crossroad with other research communities. After two editions, we have decided to launch a series of conference proceedings that collect the extended abstracts presented at the conference. Like the conference, this manuscript collection is multidisciplinary by its very nature.

This volume comprises the papers of all authors who contributed to the 2022 edition of the event, jointly organized by the University of Trento and the Conservatorio di Musica F.A. Bonporti of Trento.

After "dissonance" in 2020 and "rules" in 2021, the keyword chosen for the 2022 edition was "interaction". The term interaction takes on different connotations depending on the context and the discipline in which it is adopted, in relation to the processes, languages, and technologies used. Interaction implies coordination, action, reaction, reciprocal modifications between the entities involved to guarantee communication and exchange. Interaction is present in all processes in which codes, devices, organizations, and people influence each other.

A reflection on exchanges in symbolic, physical and socioeconomic systems, which is even more relevant following the the rise of novel ways of organizing and computing paradigms based on machine learning and artificial intelligence, can provide significant perspectives for interpreting the interactions that characterize our experience and our perception of music, making music and the world of sounds from unusual points of view.

Overview of the Volume

The conference has seen the participation of researchers and practitioners from different countries and academic contexts, and it has been a great opportunity for fruitful discussions that span across a variety of topics. The volume will be introduced by the contribution of three keynote speakers, Ellen Dissanayake (Univ. of Washington), Nick Crossley (Univ. of Manchester), and Almo Farina (Univ. degli Studi di Urbino). These three manuscripts set the stage for locating the conference contributions in context, ranging from the biological origin to musicality (Dissannayake), to social networks in music (Crossley), to ecology (Farina).

In particular, Dissannayake addresses how the first interaction between mothers and infants can build the foundation of a universal interactive behavior between them, commonly known as "baby talk," which, according to the author, can be seen as a forerunner of what eventually becomes music or musicality. Crossley's contribution addresses the issue of social interaction in music and discusses how music can be conceived as a social interaction based on rules and resources to be coordinated. He maintains that networks play a crucial role in combining resources and especially those required for making and enjoying music.

Farina's contribution is instead focused on the analysis of the interactions across natural and human domains, investigating the central role of sound in landscape ecology. The author explores the properties of sound and the possibility of using it as a tool to activate interactions between the natural and the hybrid (created by humanity) worlds.

The second part of the volume follows the conference program, consisting of presentation sessions, poster sessions and demonstration sessions. Presentation sessions include the discussion of theoretical and practical frameworks to promote the idea of interaction in different domains, bringing to the attention of the audience an all round perspective on what "interaction" means, spanning across social sciences, music, philosophy, and technology. The contributions include algorithms, teaching frameworks, multimodal and multidisciplinary interaction paradigms. The poster and demonstration sessions are instead aimed at presenting practical implementations and in-progress works, such as software libraries, user studies, complementing the technical program with tangible experiences that stimulated a fruitful discussion among all the participants.

We hope the reader will enjoy the contributions included in this volume, which, in the perspective of DiMMI, aim at promoting cross-disciplinary debate, and a space where researchers, scholars, and musicians can benefit from the mutual exchange of knowledge.

The First Interaction Foretells Musical Behavior in Our Species

Ellen Dissanayake

School of Music, University of Washington, Seattle, WA, USA

Abstract

Could human musicality be a species-specific behavior like language, laughter, hand-use, cooperation, ethnocentrism, and so forth—something that evolved because it contributed to our ancestors' survival and reproductive success? I propose such a hypothesis based on a universal interactive behavior between human mothers and infants, commonly called "baby talk," which I propose is a forerunner of what eventually became music or musicality.

1 Background

My approach to the origin of musicality is based on information and ideas from many subjects:

- *Human prehistory/archaeology* (the deep human past)
- *Ethology* (the biology of the behavior of animals in their natural environment)
  Human ethology: the biology of human behavior—e.g., musical behavior or "musicality"
- *Anthropology*
  Physical: anatomy and physiology of music-making (singing, instrument playing)
  Cultural: musical behavior in various cultures—similarities and differences
- *Psychology*
  Developmental: how musicality develops in the individual from birth to maturity
  Cognitive: what mental capacities contribute to musicality? How does music express and generate emotions?
- *Neuroscience*
  Cognitive and affective: roughly, making and understanding music and responding to it
- *Aesthetics*
  Philosophy of art; knowledge and experience of individual arts

There are at least five reasons that suggest that musicality is a biological adaptation (or predisposition). Music-making is: (a) found in every known society, past and present; (b) "costly"— even (and especially) in subsistence societies, it takes time and effort that could be devoted to activities more obviously relevant to survival (such as hunting, gathering, or preparing food; making implements, clothing, and other necessities of daily life; courtship, and even resting); (c) like other evolutionarily important things (such as eating, socializing with others, having sex, caring for children, being warm and safe, and playing), music-making is a source of pleasure; (d) without being taught, very young children are predisposed to engage in musical behavior: they move to music, sing, and dance (as well as engage in other arts—playing with words, wearing costumes, making believe); (e) most culturally-important social activities are accompanied by music.

2 The Biological Origin of Musicality

2.1 The first relationship

People all over the world speak and act differently with infants than with older children or adults. Why should this be so? Though physically helpless, newborns are socially precocious. Immediately alert at birth, they respond to human voices and faces more than to any other sound or sight. Soon after birth they look into their mother's eyes in what is called mutual gaze. A mother talks to her baby in a soft, high-pitched, sing-song voice, repeating words and phrases. Although the infant does not understand the words, it pays rapt attention. If its attention strays, the mother will respond by altering her vocalizations in order to regain the baby's regard; if it starts to fuss or cry, the mother again adjusts her sounds to calm and soothe it.

Not only vocalizations are altered (repeated, exaggerated, and elaborated), but there are exaggerated facial expressions (open mouth, wide eyes, wide sustained smiles, raised eyebrows) and head movements (head bob backwards, nodding). Maternal body movements in the interaction include rhythmic pats and strokes, as well as hugs and kisses. These vocal, visual, and gestural signals are exaggerated forms of ordinary friendly sounds, expressions, and movements that adults use with each other when they express familiarity, agreement, agreeableness, or readiness for contact. Fathers and other fond relatives also find themselves behaving like this, which they would never dream of doing with anyone else. It is biological evolution that induces both mother and infant to engage in these unusual behaviors because they are important for the baby's survival and the mother's reproductive success, even if they don't have a clue why they are acting this way. And surprisingly, it is the infant who comes into the world ready to "teach" its caretaker to make special sounds and funny facial expressions by responding to these with smiles, wriggles, and coos. If an adult talks to a baby as it does to another adult, the baby will look away and fuss. Mothers quickly catch on to what a baby wants (and needs), which is certainly not adult conversation.

2.2 Resemblances between baby talk and music

Looking at a transcript of baby talk as one hears and watches a video of mother-infant engagement, one can note musical features (melody, rhythm, variations in volume and speed, and manipulation of expectation) and a similar temporal structure (phrases, repeated motifs or themes, framed episodes with a consistent expressive mood, theme and variations, and evidence of mutual coordination). Using frame-by-frame microanalysis one can see coordination in time: imitation, joint action (chorusing), and turn-taking. In both music and baby talk, movement is essential, and both behaviors may have similar results or effects: pleasure, "wordlessness" or "indescribability," social regulation and enculturation, attunement and entrainment, and emotional bonding ("self-transcendence").

2.3 But (evolutionarily speaking) how did we get from baby talk to music?

2.3.1 Anatomical adaptations from our ancestral past

In the millennia after 1.8 million years ago, in the early evolution of our genus, Homo, two momentous adaptations began to collide. The first was gradual upright walking on two legs (bipedality), which produced over time a deluge of anatomical changes, including reshaping and relocating the opening of the spinal cord, restructuring the rib cage and the bones of the inner ear, altering the anatomy of the hips, lower limbs and the ankles, toes, and soles of the feet, reconfiguring of joints and body musculature. Significantly, among these required anatomical adaptations was a reshaped pelvis and consequently a narrowed birth canal. The second anatomical change was in brain size (encephalization), increasing between the earliest hominid, Australopithecus (508 cc.) and Homo erectus (973cc), between four and three million years ago, and even more dramatically between H. erectus (perhaps H. habilis or H. ergaster, which has a longer vocal tract) and modern humans, who have a brain size of about 1400 cc.

2.3.2 The "obstetric dilemma" and its evolutionary solutions

The alterations to the female pelvis required by bipedality included gradual shortening from fore to rear so that the trunk became centered over the hip joints and thereby reduced fatigue during upright walking. And gradual adult brain enlargement entailed a larger and larger head of the unborn baby. Anatomical and behavioral adaptations further occurred to allow a successful birth and rearing of infants.

A. Anatomical adaptations: (a) the infant skull became compressible at birth: the fontanelle (or "soft spot"); (b) cartilage between the two halves of a mother's pelvis could stretch slightly during childbirth; (c) significant brain growth of the baby occurs outside the womb (i.e., the brain size of a newborn triples between birth and four years, and at maturity the brain is four times larger than at birth); (d) the gestation period is comparatively reduced (when compared with and adjusted to that of other primates): it is estimated that a human baby would require 21 months in utero and at birth would weigh 25 pounds.

B. Behavioral adaptations:the mother-infant interaction, as described earlier, results in emotional bonding that ensures that the mother wants to take care of a helpless and demanding (if adorable) infant for not just weeks and months but years. Special characteristics of the interaction reinforce her brain's neural networks and neurochemistry for affiliation by (a) the mother's use of special alterations to the vocal, visual, and gestural affinitive signals that she directs to her infant: formalization (composed, simplified), repetition, exaggeration, elaboration, and manipulation of expectation) as well as by prolonging and temporally coordinating them; (b) affiliative hormones such as oxytocin, present in the brains of all mammals, are released during patterned, dynamic, rhythmic activity with infants (as with adults); (c) feelings of trust and confidence are further effects of coordination and entrainment with other humans.

3 Why Should the Baby Talk Interaction Ever Have Evolved?

As mentioned earlier, babies are born wanting (and able to entice) their caretakers to engage in the peculiar altered vocal, visual, and gestural signals that comprise baby talk. In addition to establishing the just-described physiological and emotional attunement called "bonding" that promotes maternal care of helpless infants, there are numerous other benefits to babies as they develop.

- *Emotion recognition and regulation.* Through interaction with a caretaker, babies learn to discriminate different emotions as they are expressed vocally and visually; they gain acquaintance with their own shifting levels of excitation and positive or negative feelings, thereby developing some degree of self-regulation of these; they learn that through their own responses and signals they can regulate the other person's stimulation.
- *Cognitive practice.* By anticipating, babies "hypothesize" what will come next and learn how to evaluate discrepancies from the expected; they test and perfect their expectations or predictions.
- *Social practice*. The interaction first acquaints infants with back-and-forth, give-and-take socializing, the rudiments of their prospective lives as social beings where their behavior calls forth reciprocal responses in another; they learn that others presuppose, require, and react in turn to their (the baby's) responses.
- *Language learning*. Babytalk prepares the way for a baby's being able to produce and understand the prototypical and meaningful sounds of the language it will eventually speak.
- *Learning of culture.* In the interaction, different cultures instill their own norms of proper behavior, whether demonstrative, restrained, and so forth.

The developmental importance of these benefits of bonding is clearly illustrated by the well-known "unintended experiment" revealed by some overcrowded Eastern European orphanages after the break-up of the Soviet Union. The babies' physical needs were haphazardly addressed, but just as serious was that there was essentially no face-to-face interactive play with their attendants. After later adoption by Western European and North American parents, it was noted that the children had difficulties interacting socially with their adoptive families and in school with other students, as well as with their own emotional regulation and control.

4 Conclusion: A Hypothesis of the Origin of Human Musical Behavior

4.1 Exaptation and ritualization

My hypothesis of the origin of human musicality requires the incorporation of two more theoretical subjects that were too complex to describe in my conference presentation, although I will describe them briefly in this article. First, it is important to emphasize that mother-infant interaction *is not itself music, but it provides the ingredients that became musicality in an evolutionary process called* ***exaptation***. This term describes an evolved feature that has become adaptive in one context (as when feathers evolved for warmth) but later evolved further in a different adaptive context. In the case of feathers, they went on to enable flight as well as, in some cases, becoming colored and patterned for male courtship or territorial display. In the case of musicality, the ancestral adaptive feature is mother-infant interaction whose components evolved, as I have described, to instill and enable the emotional bond that facilitated better care by mothers of increasingly immature and helpless infants.

I propose that the origin of these bonding components of mother-infant interaction was earlier social signals of voice, face, and body that in human adults today (and presumably in the ancestral past) are universally used to communicate friendliness and a willingness to cooperate. These are so common that we may not even be consciously aware of them until they are described by psychologists of emotional expressions;: they subliminally convey positive, nonaggressive intentions such as smiling; a receptive face with open eyes and mouth; an agreeable voice; and unthreatening movements. These everyday communicative signals in our remote ancestors were available to be exapted by mothers with babies when altered by the five "operations" described above—again, formalization (organization or simplification), repetition, exaggeration, elaboration, and manipulation of expectation (or surprise). The operations gave the friendly adult signals an evolutionarily new emphatic effect when solicited by and directed to infants: attracting their attention, sustaining their interest, and creating and manipulating their emotions.

Such a transformation of communicative signals in animals—especially birds, but also many species of fish, reptiles, and even insects—was identified by twentieth-century ethologists and called ***ritualization***, the second theoretical subject. Although some researchers have claimed that humans have no ritualized behaviors, it can be argued that individuals in groups everywhere observe conventions of greeting and parting, displaying status, showing submission, and so forth, in culturally different ways. However, I claim that in mothers' interaction with their infants, the five behavioral components or operations, and their effects, can be considered as an adaptive ritualized behavior, albeit with some cultural variants.

4.2 Ritual

The ethological term "ritualization" was derived from the widely observed human practice of rituals, which take place at times of existential uncertainty and transitions from one state to another. They reflect a universal desire for supernatural help or protection, with the hope of obtaining goods or deflecting evils. Rituals are important occasions in which groups spend a great amount of time, energy, and costly goods in order to demonstrate to their "spirits" or deities that they really care about the matter at hand. The costliness and effort of the ritual is commensurate with (and thus demonstrates) the group's need and worthiness.

In ritual ceremonies ordinary speech and sounds become singing, chanting, or playing instruments; ordinary movements become dancing. Activities such as decorating one's surroundings, wearing costumes and masks, performing stories, and making marks are similarly transformations of everyday behaviors. They attract attention, sustain interest, and create and shape emotion, using the same "operations" that in ancestral mothers and infants were adaptive solutions to the consequences of the obstetric dilemma. The same adaptive neural connections and hormones that promoted bonding between mother-infant pairs became, in ritual contexts, exaptations that, as music and the other arts, bonded members of a social group in confidence and unity. In other words, the multimodal behaviors of voice, face, and movement that we identify today as "music" (or "musicality") arose and were developed in ritual practices that bonded small human bands together in their dangerous and unpredictable world, thereby contributing to their survival and reproductive success.

It is important to recognize that rituals, fundamentally, are collections of arts and would not exist without them. The arts are essential to ritual ceremonies. They discharge the neurochemicals that bind members of a group together in confidence and unity.

This article, like my talk, has presented a whirlwind tour over many complex subjects that have to do with human-human "interaction," the theme of this conference. I leave my readers to wonder how many of the "operations" of musical behaviors foreshadowed in mother-infant interaction still obtain in human-computer interactions. If we examine these compositions. I daresay that we will generally find the exapted operations of (once again) formalization (organization or simplification), repetition, exaggeration, elaboration, and manipulation of expectation (surprise) and their emotional effects of attracting attention, sustaining interest, and creating and manipulating emotions.

References

[1] 2022 Lament and liminality: A musical prototype of early artmaking. In Jonathan L. Friedmann, ed., Music in Human Experience: Interdisciplinary Perspectives on a Musical Species Cambridge: Cambridge Scholars Press, pp. 161-186.

[2] 2021 Ancestral human mother-infant interaction was an adaptation that gave rise to music and dance. Commentary to P. E. Savage, et al., "Music as a coevolved system for social bonding." Behavioral and Brain Sciences 44, p. 68. DOI: https://doi.org/10.1017/S0140525X20001144

[3] 2018 Homo musicus: Are humans biologically predisposed to be musical? In Dusan Bogdanovich and Xavier Bouvier, eds., Tradition and Synthesis: Multiple Modernities for Composers and Performers. Lévis Québec: Doberman-Yppen, pp.17-28.

[4] 2014 Homo musicus: ¿Estamos biológicamente predispuestos para ser musicales? In Silvia Español (ed.), Psicología de la música y del desarrollo Una exploración interdisciplinaría sobre la musicalidad humana. Buenos Aires: Paidós.

[5] 2012 The earliest narratives were musical, Research Studies in Music Education 34(1) 3–14.

[6] 2010 Tracking musical chills, In Alice Andrews & Joseph Carroll, eds., The Evolutionary Review: Art, Science, Culture. 1(1). State University of New York Press, 64-68.

[7] 2009 Root, leaf, blossom, or bole: Concerning the origin and adaptive function of music. In Stephen Malloch & Colwyn Trevarthen, eds., Communicative Musicality. Oxford: Oxford University Press, 17-30.

[8] 2009 Bodies swayed to music: The temporal arts as integral to ceremonial ritual. In Stephen Malloch & Colwyn Trevarthen, eds., Communicative Musicality. Oxford: Oxford University Press, 533-544.

[9] 2008 If music is the food of love, what about survival and reproductive success? Musicae Scientiae Special Issue: 169-195. (2010) French translation: Si la musique est nourriture d'amour, qu'en est il de la survie et du success reproductif? In Irène Deliège, Oliver Vitouche, & Olivia Ladining, eds., Musique et Evolution, Wavre, Belgium: Mardaga, 65-80.

[10] 2010 French translation: Si la musique est nourriture d'amour, qu'en est il de la survie et du success reproductif? In Irène Deliège, Oliver Vitouche, Olivia Ladining, eds., Musique et Evolution, Wavre, Belgium: Mardaga, 65-80.

[11] 2006 Ritual and ritualization: Musical means of conveying and shaping emotion in humans and other animals. In Steven Brown & Ulrik Volgsten, eds., Music and Manipulation: On the Social Uses and Social Control of Music. Oxford: Berghahn, 31-57.

[12] 2000 Antecedents of the temporal arts in early mother-infant interaction. In Nils Wallin, Björn Merker & Steven Brown, eds., The Origins of Music. Cambridge, MA: MIT Press, pp. 389-410.

[13] 2007 Natalie Angier, The dance of evolution, or how art got its start, New York Times Science Section (29 November).

[14] 2010 Bruce Bower, Birth of the beat: Music's roots may lie in melodic exchanges between mothers and babies, (in Special Issue: A Mind for Music, Cover Title: "Infancy's Symphony: How Musical Instincts Help Mom and Baby Connect), Science News 178:4, 18-23.

[15] 2015 June Countryman, Martha Gabriel, and Katherine Thompson, "Children's spontaneous vocalisations during play: Aesthetic dimensions." Music Education Research (19 pages). /doi/full/10.1080/14613808.2015.1019440

Music as Social Interaction

Nick Crossley

University of Manchester, UK

Abstract

In this paper I present a sociological perspective upon interaction and music. Specifically, I argue that music is social interaction. I begin by making the basic case for this claim. I then embellish the claim by considering the role of conventions and resources in musical interaction. I conclude with a reflection on the broader networks of interaction involved in music worlds.

1 Music as Social Interaction

We sometimes refer to music as if it were an object but it isn't. Music is an activity or, better still, a form of social interaction [Becker, 1982], [Small, 1989]. This contention is nicely captured in Small's (1989) suggestion that 'music' be considered a verb ('to music'), taking the present participle, 'musicking' (see also [Crossley, 2022]). It can be elaborated by way of a more common definition of music, posited by Blacking [Blacking, 1973] amongst others, as 'humanly organised sound'. In contrast to the vibrations of air which become sound within the human auditory system, and which exist independently of human beings, sound itself is an intentional phenomenon which only exists in the experience of beings who hear it. Sound must be heard in order to exist (as sound). If music is sound, therefore, it only exists for and in virtue of audiences who hear it. Moreover, importantly, hearing is the product of interaction between a sentient, listening organism and generative events in the immediate environment of that organism which give rise to the aforementioned vibrations; namely, in the case of music, the activities of performers, either live or recorded. Music exists within the interaction between performers and listeners. Furthermore, whilst, in the case of music, sounds are organised by performers and, in many cases, by composers before them, listening too plays a role in the organisation of sonic materials. As [Merleau-Ponty, 1962] says of the gaze, listening, informed by demands arising from the activities of the organism, interrogates its environment, ordering and patterning the materials it encounters in an effort to make sense of that environment. It 'groups' discrete tones, for example, such that the listener hears melodies [Husserl, 1964] and, indeed, rhythms. There are no melodies or rhythms in the absence of listeners capable of arranging what they hear in these ways. If music is 'humanly organised sound', therefore, it is equally, necessarily, social interaction.

The performer and listener roles involved in these musical interactions are ordinarily played by different people, such that music is a straightforward case of social interaction. In some cases, however, including solitary practice, the same individual plays both roles – performer and listener. This may seem to contradict the argument that music is social interaction but I have suggested elsewhere that it is better considered a special case of it ([Crossley, 2020], see also [Small, 1989]). This is not the place to rehearse this argument in detail. Suffice it to say that, as G.H. Mead [Mead, 1967] argued for human reflexivity more generally, when players take up the listener role in relation to their own performances, reflecting upon and judging what they are doing, they draw upon the perspectives of others (both individuals and collectives to which they belong) with which they are familiar, effectively bringing those perspectives to bear in a simulated conversation. They anticipate what others might have to say about their tone or timing, for example, and act upon these 'responses'. In addition, in many cases they play and thereby interpret pieces written by others, using equipment designed, made, sold and sometimes maintained by yet others in an elaborate division of labour. They may be alone in the room but their activities depend upon and engage with those of multiple others.

As these latter comments suggest, the interaction between performer and listener is only one dyad in what is, in effect, a potentially complex network. In addition to the interaction between performers and a composer (typically mediated by way of a written score), for example, performers interact with one another, seeking to find and maintain a common groove, adjusting to one another's mistakes and stimulating and encouraging one another [Monson, 1996]. Likewise audiences interact. This is clear in the live situation, where they dance together, collectively form mosh pits and support one another's crowd surfing. Together they enact rituals and generate what Durkheim [Durkheim, 1915] called 'collective effervescence', all of which shapes their listening and thus the music which they hear. In addition, the formation of audience subcultures, listening publics and fan clubs, alongside more mundane (though often animated) conversations between enthusiasts and critics all enter into musicking, framing and informing the listening process in ways which influence what is heard, the meaning it takes on and higher-order interpretations of it.

To these interactions we must add the input of what Becker [Becker, 1982] calls 'support personnel'; that is, producers, promoters, sound engineers, roadies etc. Whether it involves live performance, recording or playback of recordings, musicking is an event (sandwiched between openings and closings) whose possibility rests upon multiple roles. Some, such as sound engineers, contribute more directly to what is heard than others (e.g. box office staff and record store assistants) but, to take the example of live musicking, events must be publicised and tickets both sold and collected if they are to succeed both financially and as experiences. This makes promoters and ticket office staff, amongst others, indispensable.

2 Conventions and Resources

Considered thus, music is not only interaction but, as Becker suggests, collective action involving a vast network. I return to the idea of networks. Firstly, however, I want to explore musical interaction further by way of two of its key elements: conventions and resources.

Any attempt at social interaction must negotiate what Lewis [Lewis, 1969] calls 'coordination problems'. There are often many different ways in which participants could achieve their goals (individual and/or collective), none more obvious or better than the others. Participants could choose any of these ways but if they are to succeed in achieving their goals it is important that they agree in what they do. It does not matter, for example, whether we drive on the left or right-hand side of the road but it is important that road users agree on one or the other. Given the many variables involved in even quite simple interactions and the multiple parties who come and go in some cases it is clearly impossible for such agreements to be struck anew in each case. Rather, past agreements tend to become institutionalised in the form of conventions [Lewis, 1969], [Becker, 1982]. By orienting to conventions individuals know how they should act in particular situations, form expectations about the actions of others in those situations, and anticipate the expectations that others have of them.

This is important in relation to musicking, which, according to Susan McClary [McClary, 2001], is convention 'all the way down'. As we know from historical and cross-cultural comparisons, every aspect of musicking, from tonal intervals (e.g. the distance between Bb and B) and scales, through notation and forms (e.g. the sonata and twelve-bar blues), to audience behaviour and ticket sales is a matter of convention. This facilitates all of the above interactions and goes right to the very heart of musical meaning and the emotions it stirs. In a very influential argument, for example, Leonard Meyer [Meyer, 1956] argues that the emotional effect of musicking is achieved by way of the manipulation of conventions (which, in turn, drives the evolution of those conventions) (see also [Huron, 2007]). Audiences orient to the conventions employed by composers and performers when listening, he argues. This, in part, enables them to order their experience and make sense of what they hear. Pieces that are too conventional and therefore too predictable are often experienced as boring. To make a piece moving (in different ways) composers and performers must play with audience expectations by deviating from conventions, in particular by delaying resolutions or finding alternative routes to them. However, such deviation presupposes conventions which both performers and listeners orient to, even as the former deviate from them.

Some conventions are common across the multiple types of music found in a given national or even international society. Others, however, serve to mark out distinct musical styles and the different 'music worlds' in which they are enjoyed. The differences between reggae, heavy metal and soul, for example, can be specified in terms of the musical conventions they typically involve.

I return to music worlds but before I do I want to consider the importance of resources to musicking and the role which musical interactions play in the exchange, pooling and combination of those resources. Musicking involves a variety of resources including time, energy, skill, equipment and often dedicated spaces, as well as the money necessary to purchase some of these. Single individuals typically lack the resources necessary for even modest musical projects, on their own, and this is often amongst their incentives for entering into interactions and relations with others. Performers, skilled in and equipped with one instrument, seek out performers skilled in and equipped with other instruments, as well as individuals with managerial skills and relevant contacts, who may negotiate with promoters, venue owners and record labels on their behalf etc. Different musical projects require different resources and access to resources therefore shapes their opportunities.

In previous work I have developed this idea by reference to 'critical mass'; that is, the number of participants required to allow an event to be staged and to succeed [Crossley, 2015], [Crossley, 2020]. Critical mass is not necessarily entirely a matter of tangible re-

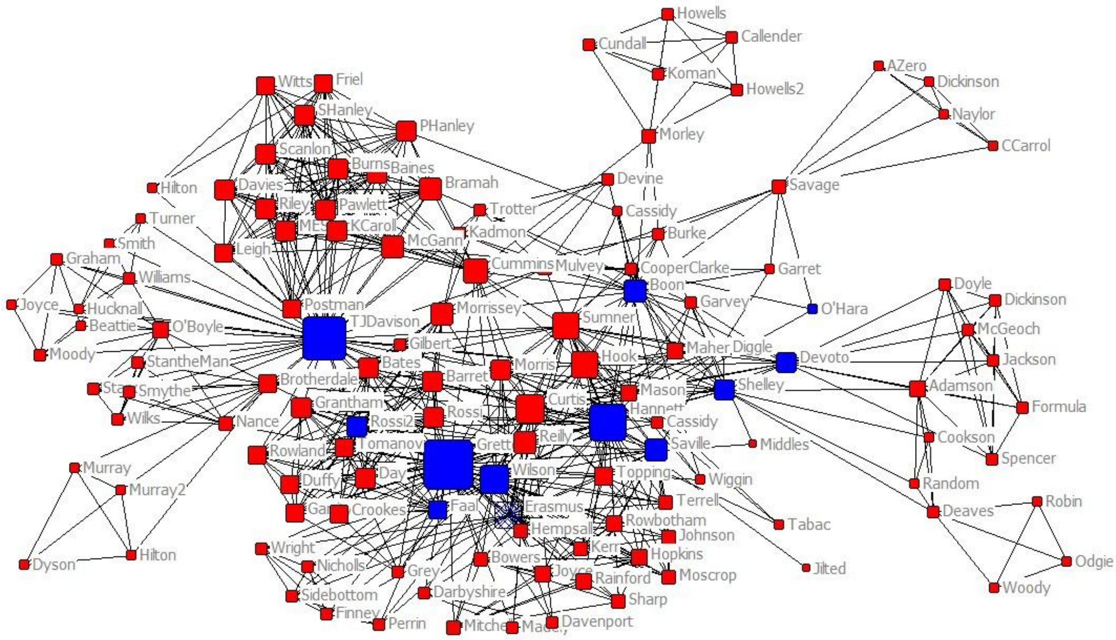


Figure 1: *The Manchester Punk/Post-Punk World, 1980*

sources. For example, the atmosphere and collective effervescence which contribute to the success of a gig – and of the course definitions of success may vary - depends in some part upon audience numbers. Beyond this, however, any musical event requires that a number of roles are performed, which, in turn, requires a sufficient number of suitably skilled people to fill those roles (one person may play several roles but the need for simultaneous contributions places limitations upon this). Furthermore, the financial incentives that may be required to fill some roles adds to costs (e.g. for venue hire, transporting and hiring equipment etc.) that must be recouped by charging audiences for admission, setting a further threshold for critical mass: audience members will generally have an upper limit to what they are prepared to pay for any given event, which determines the audience size required to break even.

3 Worlds as Networks

A critical mass of willing participants is a necessary prerequisite for a sustainable 'music world'. The classical concert world, the jazz, punk and folk worlds (and/or any of their local manifestations, e.g. the Milan jazz world, the Liverpool punk world, Italian opera) all require sufficient participants to do their particular forms of musicking, both live and recorded, and (in the absence of the wealthy patrons and/or government grants which more highbrow worlds might attract) a sufficient number of paying customers to meet the costs which those forms incur. However, a critical mass is more than an aggregate of individuals. It involves interaction and relations between participants, forming a network with identifiable (and variable) properties. These networks may be huge but smaller ones can be visualised as graphs. In Figure One I visualise the network of performers and support personnel involved in Manchester's punk and post-punk world, as of January 1980 (see [Crossley, 2015] for further elaboration). Individual participants are represented by small squares ('vertices') and relations between them, which in this case involve co-participation on a musical project, are visualised by lines which connect them ('edges').

To emphasise an earlier point I have coloured the vertices of participants who played a support/organisational role in this world (sometimes alongside a performer role) blue, and those who did not red. I also sized vertices according to the number of connections each has (its 'degree'). This very clearly shows that organisers are typically better connected in music worlds than performers (their average degrees are 24 and 9 respectively in this case), underlining their significance in musicking of this type.

Networks are crucial for the movement of resources through a world. However, they play a variety of other roles too. They are crucial for the coordination of activities (e.g. getting everybody involved in an event, including audiences in the case of gigs, to the right place at the right time). They cultivate trust, mutual support and cooperation, which, in turn, allows their participants to act in ways that may not otherwise be possible (e.g. borrowing equipment and learning new 'tricks') [Crossley, 2015]. They simultaneously cultivate competition, which encourages practice and innovation amongst artists, who seek to stand out from the rest. They facilitate recruitment, both of band members and audiences. And they create an environment in which styles (both musical and sartorial) which might be perceived as deviant and discouraged in wider society can be cultivated (ibid.).

4 Conclusions

In this paper I have a sketched a brief case for considering music to be a form of social interaction. I began by noting that the common definition of music as 'humanly organised sound' presupposes on-going interaction between performing and listening (even if both roles are played by the same person in some cases). I then considered the many wider roles often involved in musicking and the various interactions that occur both within and between sets of incumbents of these various roles. Embellishing this further, I first considered the 'coordination problems' posed by all attempts at interaction and the role of convention in helping participants to resolve these problems. I then considered the role of resources in musicking and the pooling, exchange and combination of resources in musical interactions. Finally, I considered the wider networks of interaction which comprise 'music worlds' and which in various ways enable musical interaction.

References

[Becker, 1982] Becker, H. (1982). *Art Worlds*. Berkeley, University of California Press.

[Blacking, 1973] Blacking, J. (1973). *How Musical is Man?* Seattle, University of Washington Press.

[Crossley, 2015] Crossley, N. (2015). *Networks of Sound, Style and Subversion: the Punk and Post-Punks Musical Worlds of Manchester*. London, Liverpool and Sheffield 1976-1980, Manchester University Press.

[Crossley, 2020] Crossley, N. (2020). *Connecting Sounds; the social life of music*. Manchester, Manchester University Press.

[Crossley, 2022] Crossley, N. (2022). *Musicking to Music Worlds*. Music Review Annual 3, 1-24, https://musicresearchannual.org/crossley-musicking-to-music-worlds-3/.

[Durkheim, 1915] Durkheim, E. (1915). *Elementary Forms of the Religious Life*. New York, Free Press.

[Huron, 2007] Huron, D. (2007). *Sweet Anticipation*. Cambridge, MIT Press.

[Husserl, 1964] Husserl, E. (1964). *The Phenomenology of Internal Time Consciousness*. The Hague, Kluwer.

[Lewis, 1969] Lewis, D. (1969). *Convention*. Cambridge, Harvard University Press.

[McClary, 2001] McClary, S. (2001). *Conventional Wisdom*. University of California Press.

[Mead, 1967] Mead, G. (1967). *Mind, Self and Society*. Chicago, Chicago University Press.

[Merleau-Ponty, 1962] Merleau-Ponty, M. (1962). *The Phenomenology of Perception*. London, Routledge.

[Meyer, 1956] Meyer, L. (1956). *Emotion and Meaning in Music*. Chicago, Chicago University Press.

[Monson, 1996] Monson, I. (1996). *Saying Something*. Chicago, Chicago University Press.

[Small, 1989] Small, C. (1989). *Musicking*. Middletown, Wesleyan University Press.

Ecology and Semiotics of Sounds in an Interacting World

Almo Farina

University of Urbino, Italy

Abstract

In this paper, we discuss the central role of sounds as semiotic vehicles to integrate and connect different compartments of natural and human domains.

We present the new discipline of ecoacoustics, which investigates in a non-invasive way the ecological role of sounds. This discipline is an example of epistemological integration with general ecology and, in particular, with landscape ecology.

Then, we will bring into question the ecosemiotic property of sounds and the possibility of using sounds as an efficient tool to activate interactions between the natural world and the hybrid world created by humanity.

Finally, we will explore new technological advancements that allow us to make virtual trips around the world, giving an extraordinary opportunity to people to explore remote areas thanks to a net of on-stream acoustic sensors, and to participate in the development of citizen science.

1 Introduction

The environment is an extraordinary collection of organisms and geophysical, biological, ecological and semiotic processes tuned by astronomical and climatic constraints. Most of our scientific knowledge is based on the functioning of the physical world driven by processes like climate dynamics, soil geochemistry and population and community dynamics [Odum, 1971]. However, despite the incredible acceleration of ecological knowledge that has occurred in the last decades, the semiotic domain that pertains communication, signification and environmental coding remains poorly investigated.

Visual, olfactory, acoustic and tactile senses are at the basis of this semiotic domain [Farina, 2022]. The information enclosed in this semiological "universe" is dominated by visual cues such as the colour patterns of bird feathers or the design of butterfly wings. These signals are utilized to maintain social cohesion in groups, populations, guilds and communities and in reproductive behaviour.

The semiology of sounds is part of this ecosemiotic universe that until now has provided only a minor amount of information. However, studies on natural and human sounds have recently experienced a promising spring opening, a grand vista on a new unexplored world of acoustic signs [Farina, 2014]; [Farina & Gage, 2017]. The advent of bioacoustics [Fletcher, 2014] and the recent development of ecoacoustics have shed new light on the complexity and diversity of bio-ecological processes.

A multitude of organisms produce sounds to activate and maintain intra- and inter-specific communication, navigate across the landscape and track resources. Passive/active sounds are extraordinary vehicles of information and meaning through natural and human-modified habitats. Although they have not received much attention, sounds are dominant phenomena in several environments. For instance, wind and rain are penetrative and persistent sounds that often mask or alter the majority of biological sounds. The spectacular falls of the Niagara River wouldn't be so majestic without the deafening sound of the river making a 99-meter leap. Bird choruses in the morning twilight in temperate and boreal forests or the insects' and frogs' daily choruses in tropical forests are examples of the degree of the dominant effects of this sonic phenomenon.

The roles of sounds and their application to the everyday life encouraged Stuart Gage, a pioneer in this field, to conceive that sounds are a powerful tool to listen the Earth's beat. Bernie Krause, composer and field recordist, considers sounds to be a celebration of the magnificence of the biodiversity offered by organisms as their choruses operate like "a natural orchestra" [Krause, 2012]. A decrease in the abundance of individuals in a population or a reduction in the number of species in a community reduces the quantity of sounds released. Several decades ago, the reduction of biological sounds was used by Rachel Carson to denounce the Earth's degradation [Brumm, 2009] as consequence of the abuse of pesticides were important causes of bird decline after the second world war. In particular, DDT and its relatives alter birds' calcium metabolism which results in thin eggshells.

In the field of biodiversity conservation, sounds are extraordinary testimonials of thousands of undescribed species that are at risk of extinction before humanity can classify them for the first time. "Fragment of extinction" is a wonderful program carried out by David Monacchi, composer and music professor at the Giacomo Rossini Conservatory in Pesaro, Italy. This scholar has tested innovative recording technologies based on multichannel devices in primeval forests of Borneo, Ecuador and Central Africa. The acoustic files collected in this way return a multidimensional sound used for a more realistic reproduction in a dedicated exhibition theatre, "the sonosfera" [Monacchi, 2016].

Natural sounds and the sounds produced by machines have a contact in hybrid landscapes where human intrusion intercepts the natural processes, reducing the level of spontaneous functionality of the environment. The diffusion of hybrid landscapes, characterized by a blend of undisturbed and modified habitats, poses a great challenge to the conservation of natural habitats, their dynamic processes and biodiversity.

2 Ecoacoustics: a short presentation

The ecological study of sound has been carried out in recent years by a new branch of ecology: Ecoacoustics [Sueur & Farina, 2015]. After a world congress in 2014 in Paris, scientists attributed the name "Ecoacustics" to the study of sounds under an ecological perspective. Ecoacoustics is the result of the integration of other disciplines like physical acoustics, bioacoustics, ecology, animal behaviour, biotremology, landscape ecology and ecosemiotics [Farina, 2022]. Figure 1 depicts a scheme of the sources, epistemological references, tools and applications offered by ecoacoustics.

Ecoacoustics has been recognized as an umbrella discipline that offers a scientific and cultural "niche" to several aspect of the sonic domain. Soundscape ecology, a term used before 2014 to indicate the study of sounds across a landscape, was recognized as a branch of ecoacoustics. Scholars were aware that the epistemological and semantic aspect of the study of environmental sounds was of primary importance to offer a clear vision of this subject.

Sounds in ecoacoustics are processed according to two major acoustic fundamentals: Frequency and Amplitude. The quantitative analysis of frequencies (represented by a spectrogram) is possible

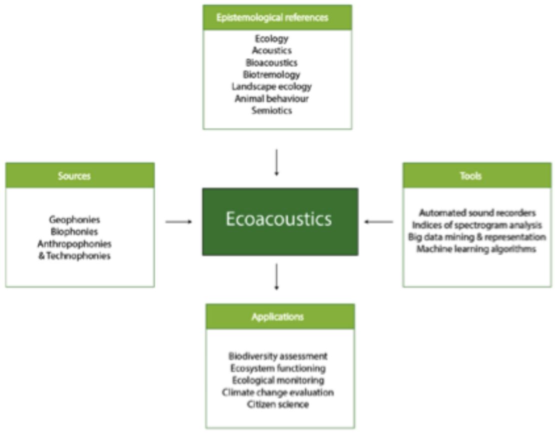


Figure 1: *The epistemological domain and its competences [Farina, 2022].*

thanks to indices like the Acoustic Complexity Index [Pieretti et al., 2011], the Normalized Difference Soundscape Index NDSI index [Gage & Axel, 2014]; [Kasten et al., 2012], [Harris et al., 2016]] and their implementation. There are more then 60 indices available to ecologists to dissect sound and to return quantities and qualities [Sueur et al., 2014]; [Gasc et al., 2015]].

3 The ecological and ecoacoustic paradigms

Before exploring the level of similarity between landscape ecology and soundscape ecology, we make a connection between ecoacoustics and some ecological paradigms like habitat, community and niche.

Ecological habitat and acoustic habitat

The habitat is a geographical area that guarantees most resources for a species [Farina, 2012]. In ecoacoustics, the acoustic habitat is a geographical area characterized by sounds that are perceived as favourable to a species. For instance, some species avoid the urban soundscape but others find this soundscape at least neutral [Mullet et al., 2017]. The acoustic habitat represents the acoustic resources necessary to a species to settle down and find other resources.

Ecological community and acoustic community

The ecological community consists of aggregates of groups of various species occupying the same geographical area at the same time. In ecoacoustics, the acoustic community is defined as an aggregation of species that produce sounds by using internal or extra-body sound-producing tools. An acoustic community is composed of only soniferous species and has an ephemeral life, regulated by solar dynamics [Farina et al. in prep.]. For instance, an acoustic community active during morning twilight is substituted by another after sunrise. The same may happen at evening twilight when bird choruses are substituted by nocturnal insect choruses. Acoustic communities also change during a season and across different regions.

Ecological niche and ecoacoustic niche The ecological niche is a central paradigm in ecology because it conjugates evolutive and adaptive processes with environmental constraints. This concept has attracted generations of ecologists for more than a century, feeding a scientific debate headed by Joseph Grinnell, Charles Sutherland Elton and G. Evelyn Hutchinson. For Grinnell, a niche is the habitat in which a species lives [Grinnell, 1917]. For Charles Sutherland Elton the niche is the place where an animal is in relation to food and enemies [Elton, 2001]. For Hutchinson, niche is a "n-dimensional hypervolume" where the dimensions are realized by the environmental condition and resources [Hutchinson, 1957]. In ecoacoustics, the acoustic niche is the frequency repartition that distinguishes one species from another and may be considered one axis of the ecological niche. This hypothesis has been introduced by Bernie Krause [Krause, 1987]; [Krause, 1993] to explain the contemporary presence of singing species that use different frequencies to avoid an acoustic overlap. Today a growing mass of evidence confirms this hypothesis, especially in frogs, insects, birds and bats (e.g., [Sinsch et al., 2012]; [Schmidt & Balakrishnan, 2015]; [Gomes et al., 2021]).

4 Landscape ecology and soundscape ecology: an example of epistemological interaction

Landscape ecology is a branch of ecology and has a main goal of investigating the geographical dimension of ecosystems. It was developed in Central Europe at the beginning of the last century and achieved popularity in North America at the end of the nineteen-eighties thanks the great contribution of botanist Richard Forman [Forman & Godron, 1986] and agronomist Frank Golley, founder of the journal Landscape Ecology, just to cite two of the most influential figures. In Europe during the same period, thanks to the contribution of ecologists like Zev Naveh and Wolfgang Haber, landscape ecology developed a more cultural integrated approach linking the spatial pattern of the environment with cultural stratification, identity and human values attributed to landscapes [Naveh, 1990]; [Haber, 2004]. In synthesis, landscape ecology focuses on the spatial distribution of vegetation, the movement of animals across an environmental matrix and the influence of human transformation of the landscape [Wiens, 2002]. Because the spatial distribution of sounds across a region is influenced by the characteristics and dynamics of the landscape, landscape ecology has been a source of inspiration to develop new ecoacoustic theories and experiment on novel indicators to assess and monitor the environment [Pijanowski et al., 2011a]; [Fuller et al., 2015]. The relationship between landscape attributes and soundscapes is the object of interesting studies on behavioural ecology [Laiolo & Tella, 2005]; [Briefer et al., 2010].

Sounds are so important in the natural world that we have no hesitation to call the acoustic universe around us a "sound-scape", using a substantial modification of the more popular term "land-scape". According to the source, a soundscape is the result of the overlap of geophonies (e.g., running water, rain, wind, sea waves), biophonies (e.g., bird song, insects, fishes calls), anthropophonies (e.g., human voice) and technophonies (e.g., movement of cars, trains, airplanes, music) [Pijanowski et al., 2011a]; [Pijanowski et al., 2011b].

Sounds are present in every type of freshwater, marine and terrestrial environment. Every environment/habitat/landscape has a specific sonic ambience that concurs to create a strict relationship with different abiotic and biotic components of the ecosystems [Tonolla et al., 2010]; [Staaterman et al., 2013]; [Gottesman et al., 2020]. Soundscape ecology, as a branch of ecoacoustics, investigates the distribution of sounds across the landscape. Thus, the first point in common between landscape ecology and soundscape ecology is the geographic dimension in which sonic processes occur. Sounds spread across the spatial structure of a landscape and are influenced by such a configuration with a strong impact on the behaviour of species [Laiolo & Tella, 2005]; [Briefer et al., 2010]. Landscape ecology and soundscape ecology are examples of semantic and epistemological interactions. To better understand the epistemological link between landscape ecology and soundscape ecology, it is useful to examine separately the patterns and the processes that occur across a landscape. According to the American school [Forman, 1995], a landscape is composed of a mosaic of spatial units (patches) with homogeneous characters. Composition, size, shape and spatial distribution are the major attributes of these units that are the result of alternation between vegetation, bare soil, water and human settlements.

The spatial distribution of different sounds creates homogeneous units like landscape patches that we call "sonotopes" [Farina, 2014] composed of the combination of geophonies, biophonies and technophonies. When we consider only the distribution of biophonies, we call such aggregations "soundtopes" [Farina, 2014] (Figure 3). Soundtopes are the geographical representation of the acoustic communities. In landscape ecology ecotones are areas of connection between two or more different patches. In soundscape ecology the areas of contact between two or more sonotopes/soundtopes are called sonotones and

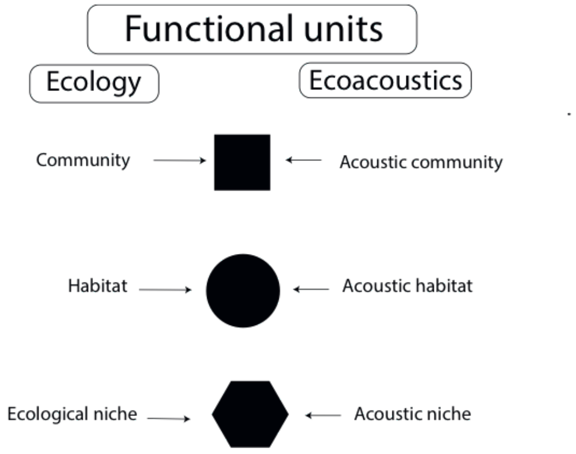


Figure 2: *Functional units in ecology and ecoacoustics.*

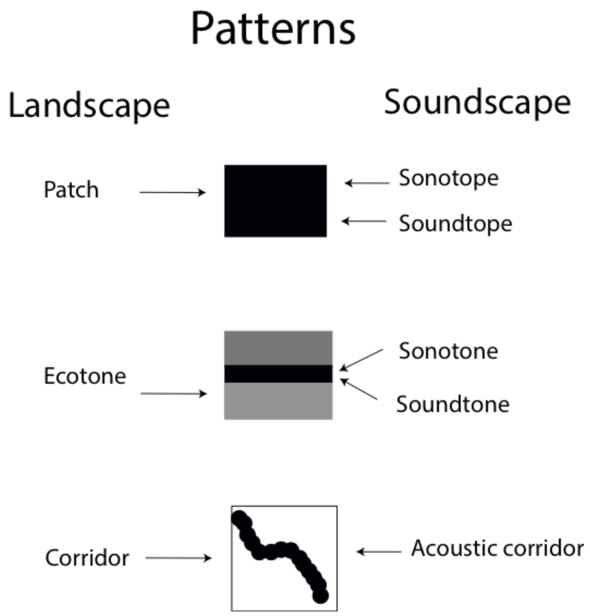


Figure 3: *Comparisons between landscape and soundscape structural entities.*

soundtones, respectively.

Disturbance, fragmentation, connectivity, connectedness and corridors are further elements in common between landscapes and soundscapes.

Disturbance

Disturbances such as a wildfire, a flood or the spray of pesticides are common processes in landscapes, originating new patterns and processes [Rykiel Jr, 1985]; [Rundel, 1998]. In ecoacoustics, disturbance is represented by the competition between sounds that have the same frequencies. For instance, technophonies with high amplitudes that commonly occur in urban areas have frequencies that overlap the same frequencies of biological origin. In this case the masking of a dominant sound (e.g., car traffic) at a higher amplitude over weaker sounds originated by biophonies is commonly known as noise and produces changes in the acoustic performance of soniferous species. Noise represents one of the most popular themes in ecoacoustic research [Erbe et al., 2018]; [Fletcher, 2012]; [Brumm, 2013].

Fragmentation

Fragmentation is a process by which homogenous areas are transformed into isolated fragments of different sizes and levels of separation according to the severity of the natural disturbances (e.g., wildfires, tornados, tree gaps) or the level of human intervention (deforestation and urbanization) [Lindenmayer & Fischer, 2013]. Fragmentation has consequences for several organisms that move across the landscape and that perceive the spatial configuration at different level of connectivity. In ecoacoustics this process has direct consequences with the separation between intraspecific acoustic sources. For instance, the isolation of a singing bird from another conspecific singer or from eavesdropping reduces or prevents intraspecific communication. Acoustic fragmentation may also be observed at the scale of an acoustic community. In this case, acoustic fragmentation is noticed as a high value of dissimilarity between two sampled positions [Farina et al. in prep.].

Connectivity and connectedness

The habitat is part of the environment that a species perceives as favourable/suitable for finding all the necessary resources to stay alive. Connectivity measures the way in which organisms maintain a functional contact with their habitats. When part of the necessary resources are located in another geographical area a species must travel to track those resources across a hostile environment. The amount of travel and the friction suffered while crossing such unfavourable environments to move from a part of a habitat to another is measured as the level of connectivity and remains species-specific. Connectedness, on the contrary, is the physical distance between two patches that have the same characters, and this variable is independent of the species [Baudry & Merriam, 1998]. These two concepts find some difficulties in being translated into the ecoacoustic domain. For instance, we can assume that the presence of a physical obstacle like a ridge, barriers like a row of dense, tall vegetation or a noisy stream may prevent the propagation of sound. When an area is favourable for several individuals of the same species, acoustic connectivity is potentially high. Landscape connectivity has been demonstrated as influencing acoustic communities and the diversity of the acoustic community [Desjonquères et al., 2018].

Corridors

In fragmented habitats, corridors represent a strategy to connect isolated patches of habitat offering an amount of favourable habitat sufficient to guarantee the transit of species from one portion of habitat to another (e.g., [Xu et al., 2019]). Ecoacoustic corridors are represented by a portion of an acoustic community represented only by one or a few species that are sufficient to guide other conspecifics.

5 The anthropological importance of soundscapes

The study of the relationships between natural and anthropogenic sounds is gaining a growing interest worldwide [Schafer, 1994]; [Samuels et al., 2010]. Sound analysis is fundamental to investigate biodiversity and the human impact on terrestrial and aquatic systems. Sounds are sources of information on population and community dynamics. The phenological phases of sounds, called sonophases [Mullet et al. in prep.], may be used as a proxy of the long-term climate change or simply to document the effect of environmental manipulation by human action. Sounds are characterized by a high capacity to integrate different epistemological approaches and have been sources of inspiration by our ancestors, a way to reduce the philosophical distance between the real world and the spiritual world beyond. Sounds have been used in sacred ceremonies in all the religions and historic periods. The use of instruments like drums (universally used by all the ancient people) or the bullroarer (used by Australian aborigines) that mimics the sound of nature have encouraged this cultural approach. It is important to spend more words on this subject that is investigated by archeo-acoustics, a branch of acoustics that tries to connect space, the position of ancient monuments and sacred sites as resonant boards [Díaz-Andreu et al., 2022]. For instance, the location of a sacred/dangerous place close waterfalls in the Virginia Blue Ridge Mountains is an example of utilization of the sound of a waterfall as a sensory experience. The location of visual motifs in Upper Paleolithic caves and acoustic responses seem correlated, confirming that sound could have influenced the behaviour of Paleolithic societies [Fazenda et al., 2017].

Modern anthropogenic sounds contribute to the identity of a place and the disappearance of characteristic sounds and may reduce the identification and recognition of a place. The sounds become an integral part of every landscape and may be considered true landmarks. Sounds are recognized in the same way as mountain peaks or other scenic environments. In particular, in rural landscapes the sounds of nature are mixed with the sounds of human activities, creating a unicum that becomes an important component of the local culture. The decennial land abandonment of rural areas in most of the world contributes to the loss of this intangible resources. For instance, on Gomera Island the whistle of residents represents a language based on sound modulation [Classe, 1957]; [Lindblad,]. This "language" developed in an environment with a steep morphology, rich in vertical walls separated by deep canyons. The great difficulty of crossing the territory has produced the whistles as a semiotic tool of human communication.

Education from sounds

Our ancestors utilized passive sounds to localize prey and prevent predator attacks in the savanna and forest in which they were wandering in search of prey, seeds and fruits. In modern societies this attitude has been forgotten but not erased in our genetic heritage and the rediscovery of sounds eliciting a great number of dormant sensations. According to the ecosemiotics narrative, natural sounds may be considered optional resources according the GTR [Farina, 2012] that become necessary to achieve some spiritual and psychological goals. The human well-being is also passing through the presence of natural sounds. The recent history of honeybee therapy that uses the sounds produced by thousands of wings flapping (buzzing) from the interior of a hive teaches us about the importance of this sound in alternative medicine and opens new modalities to recover from several diseases. If sound is important for the scientific advancement, in the same way it is a relevant subject for a permanent education of "common" people. Today, new technologies of digital communication based on the transmission of information from satellite internet allows deployment of microphones in remote regions of the Earth to hear live-streaming of the environmental sounds to obtain a useful and educational tool available to every category of people.

For instance, the "Locus Sonus Stream Project [1] offers a worldwide network of "open mikes" that permanently stream local soundscapes to a dedicated server. The resulting live audio is used in a large variety of artistic projects.

A recent project (December 2022) of the Fivizzano Commune of Northern Tuscany, EcoSoundscape[2] offers the soundscape from four mikes that stream all day from different rural landscapes. The website provides detailed information about the locations, the species and also offers a "scientific diary" in which relevant acoustic events are published (Figure 4). This project will receive a further implementation with the automatic identification of the species by applying techniques of machine learning that allow the automatic identification of species. The original files are stored in a temporary memory located in the device and can be downloaded on demand for scientific data processing and analysis. The creation of a net of such devices in a region will allow us to monitor changes in animal phenology, biodiversity composition and dynamics and assess the level of human intrusions, just to list some of the potential services offered by this technology.

Recently, thanks to the new algorithms of artificial intelligence (AI) are available platforms that identify bird species in real time and are grow-

[1] https://locusonus.org/wiki/index.php?page=Locustream.en
[2] https://Ecosoundscape.it

Figure 4: *A SET™ (Lunilettronik.it), a station used in the Eco-Soundscape project ® for the International Institute of Ecoacoustics.*

ing projects in so-called citizen science that may be defined as public participation in scientific projects [Eitzel et al., 2017]. For instance, BirdNET is a citizen science platform launched by the K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology[3]. BirdNET is also a software that is used to analyse large collections of audio. This application can identify 3,000 of the world's most common species and the list will be enlarged in the future.

An active role from citizen science is realized by Merlin, a platform launched at the same laboratory, that can transmit a record from a smartphone to a server that will be associated with the GPS position of the smartphone in the locality in which the contact is established. In this way, a great amount of acoustic information can be harvested around the world, creating an open data bank that can be accessible to scientific processing and evaluation.

6 Concluding comments: Sound as an ecosemiotic barrier to Earth's degradation

Sounds, conceived as a tangible resource actively tracked by soniferous species to satisfy functions like reproduction, territorial patrolling, socialization, safety assessment, navigation and habitat selection, are also efficient vehicles of eco-semiotic interactions at the scale of individuals, populations and communities. Many of the vital traits of organisms are guided and shaped by emitted and received sounds. Sounds create a network of dynamic interactions that converge into the semiotic agency of the "acoustic community". Sound is probably the least expensive ecosemiotic vehicle to transfer information between organisms. Sound is a dynamic and independent vector according to who is hearing it and complete information is associated to it. Its ecosemiotic nature is confirmed and represents the best way to track biodiversity or to assess environmental quality without moving through a hostile environment or disturbing the source. In conclusion, while biodiversity is currently decreasing at an alarming rate and species change their habitats or suffer from phenological mismatches, natural sounds represent the most accessible tool to investigate environmental quality.

[3] https://birdnet.cornell.edu/

References

[Baudry & Merriam, 1998] Baudry, J. & Merriam, H. (1998). *Connectivity and connectedness: functional versus structural patterns in landscapes*. In: Schreiber, K.-F. (Ed.), Connectivity in Landscape Ecology. Proceedings of the Second International Seminar of the International association for Landscape Ecology, Munster, 1987, vol. 29. Munstersche Geographische Arbeiten, pp. 23–28.

[Briefer et al., 2010] Briefer, E., Osiejuk, T., Rybak, F., & Aubin, T. (2010). *Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach*. Journal of Theoretical Biology, 262(1), 151–164.

[Brumm, 2009] Brumm, H. E. (2009). *Silent Spring*. 1962.

[Brumm, 2013] Brumm, H. E. (2013). *Animal Communication and Noise (Vol. 2)*. Springer Science Business Media.

[Classe, 1957] Classe, A. (1957). *The whistled language of La Gomera*. Scientific American, 196(4), 111–124.

[Desjonquères et al., 2018] Desjonquères, C., Rybak, F., Castella, E., Llusia, D., & Sueur, J. (2018). *Acoustic communities reflect lateral hydrological connectivity in a riverine floodplain similar to macroinvertebrate communities*. Scientific Reports, 8(1), 1–11.

[Díaz-Andreu et al., 2022] Díaz-Andreu, M., Jiménez Pasalodos, R., Rozwadowski, A., Álvarez Morales, L., Miklashevich, E., & Santos da Rosa, N. (2022). *The Soundscapes of the Lower*

Chuya River Area, Russian Altai: Ethnographic Sources, Indigenous Ontologies and the Archaeoacoustics of Rock Art Sites. Journal of Archaeological Method and Theory, 1–28.

[Eitzel et al., 2017] Eitzel, M., Cappadonna, J., Santos-Lang, C., Duerr, R., West, S., Virapongse, A., & Jiang, Q. (2017). *Citizen science terminology matters: Exploring key terms*. Citizen science: Theory and practice, 1–20.

[Elton, 2001] Elton, C. (2001). *Animal Ecology*. University of Chicago Press. p. 64. ISBN 978-0226206394.

[Erbe et al., 2018] Erbe, C., Dunlop, R., & Dolman, S. (2018). *Effects of noise on marine mammals*. Effects of anthropogenic noise on animals, 277–309.

[Farina, 2012] Farina, A. (2012). *A biosemiotic perspective of the resource criterion: Toward a general theory of resources*. Biosemiotics, 5, 17–32.

[Farina, 2014] Farina, A. (2014). *Soundscape ecology*. Springer, Dordrecht.

[Farina, 2022] Farina, A. (2022). *Principles and Methods in Landscape Ecology: An Agenda for the Second Millennium*. In Principles and Methods in Landscape Ecology: An Agenda for the Second Millennium (pp. 1–42). Cham: Springer International Publishing.

[Farina & Gage, 2017] Farina, A. & Gage, S. E. (2017). *Ecoacoustics: The ecological role of sounds*. John Wiley Sons.

[Fazenda et al., 2017] Fazenda, B., Scarre, C., Till, R., Pasalodos, R., Guerra, M., Tejedor, C., ..., & Foulds, F. (2017). *Cave acoustics in prehistory: Exploring the association of Palaeolithic visual motifs and acoustic response*. The Journal of the Acoustical Society of America, 142(3), 1332–1349.

[Fletcher, 2012] Fletcher, J. E. (2012). *Effects of Noise on Wildlife*. Elsevier.

[Fletcher, 2014] Fletcher, N. (2014). *Animal bioacoustics*. Springer handbook of acoustics, 821-841.

[Forman, 1995] Forman, R. (1995). *Land mosaics*. Cambridge University Press, Cambridge, US.

[Forman & Godron, 1986] Forman, R. & Godron, M. (1986). *Landscape Ecology*. Wiley Sons Canada.

[Fuller et al., 2015] Fuller, S., Axel, A. C., Tucker, D., & Gage, S. H. (2015). *Connecting soundscape to landscape: Which acoustic index best describes landscape configuration?* Ecological indicators, 58, 207–215.

[Gage & Axel, 2014] Gage, S. & Axel, A. (2014). *Visualization of temporal change in soundscape power of a Michigan lake habitat over a 4-year period*. Ecological Informatics, 21, 100–109.

[Gasc et al., 2015] Gasc, A., Pavoine, S., Lellouch, L., Grandcolas, P., & Sueur, J. (2015). *Acoustic indices for biodiversity assessments: Analyses of bias based on simulated bird assemblages and recommendations for field surveys*. Biological Conservation, 191, 306–312.

[Gomes et al., 2021] Gomes, D., Toth, C., Cole, H., Francis, C., & Barber, J. (2021). *Phantom rivers filter birds and bats by acoustic niche*. Nature Communications, 12(1), 3029.

[Gottesman et al., 2020] Gottesman, B., Francomano, D., Zhao, Z., Bellisario, K., Ghadiri, M., Broadhead, T., ..., & Pijanowski, B. (2020). *Acoustic monitoring reveals diversity and surprising dynamics in tropical freshwater soundscapes*. Freshwater Biology, 65(1), 117–132.

[Grinnell, 1917] Grinnell, J. (1917). *The niche-relationships of the California Thrasher*. The Auk, 34(4), 427–433.

[Haber, 2004] Haber, W. (2004). *Landscape ecology as a bridge from ecosystems to human ecology*. Ecological research, 19, 99–106.

[Harris et al., 2016] Harris, S., Shears, N., & Radford, C. (2016). *Ecoacoustic indices as proxies for biodiversity on temperate reefs*. Methods in Ecology and Evolution, 7(6), 713–724.

[Hutchinson, 1957] Hutchinson, G. (1957). *Concluding remarks*. Cold Spring Harbor Symp, 22, 415–427.

[Kasten et al., 2012] Kasten, E., Gage, S., Fox, J., & Joo, W. (2012). *The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology*. Ecological Informatics, 12, 50–67.

[Krause, 1987] Krause, B. (1987). *The niche hypothesis: How animals taught us to dance and sing*. Whole Earth Review, 57(57), 14–16.

[Krause, 1993] Krause, B. (1993). *The niche hypothesis: a virtual symphony of animal sounds, the origins of musical expression and the health of habitats*. The Soundscape Newsletter, 6, 6–10.

[Krause, 2012] Krause, B. (2012). *The great animal orchestra: finding the origins of music in the world's wild places*. Hachette UK.

[Laiolo & Tella, 2005] Laiolo, P. & Tella, J. (2005). *Habitat fragmentation affects culture transmission: patterns of song matching in Dupont's lark*. Journal of Applied Ecology, 42(6), 1183–1193.

[Lindblad,] Lindblad, P. *On the whistle language of Gomera*. Working papers/Lund University, Department of Linguistics and Phonetics, 9.

[Lindenmayer & Fischer, 2013] Lindenmayer, D. & Fischer, J. (2013). *Habitat fragmentation and landscape change: an ecological and conservation synthesis*. Island Press.

[Monacchi, 2016] Monacchi, D. (2016). *A philosophy of eco-acoustics in the interdisciplinary project fragments of extinction*. Environmental sound artists, 159–168.

[Mullet et al., 2017] Mullet, T., A, F., & SH, G. (2017). *The Acoustic Habitat Hypothesis: an ecoacoustics perspective on species habitat selection*. Biosemiotics, 10, 319–336.

[Naveh, 1990] Naveh, Z. (1990). *Landscape ecology as a bridge between bio-ecology and human ecology. Cultural aspects of landscape*. PUDOC, Wageningen, 45–58.

[Odum, 1971] Odum, E. (1971). *Fundamentals of Ecology*. Saunders, Philadelphia.

[Pieretti et al., 2011] Pieretti, N., Farina, A., & Morri, D. (2011). *A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI)*. Ecological indicators, 11(3), 868–873.

[Pijanowski et al., 2011a] Pijanowski, B., Farina, A., Gage, S., Dumyahn, S., & Krause, B. (2011a). *What is soundscape ecology? An introduction and overview of an emerging new science*. Landscape ecology, 26, 1213–1232.

[Pijanowski et al., 2011b] Pijanowski, B., Villanueva-Rivera, L., Dumyahn, S., Farina, A., Krause, B., Napoletano, B., ..., & Pieretti, N. (2011b). *Soundscape ecology: the science of sound in the landscape*. BioScience, 61(3), 203–216.

[Rundel, 1998] Rundel, P. (1998). *Landscape disturbance in Mediterranean-type ecosystems: an overview*. Landscape Disturbance and Biodiversity in Mediterranean-Type Ecosystems, 3–22.

[Rykiel Jr, 1985] Rykiel Jr, E. (1985). *Towards a definition of ecological disturbance*. Australian Journal of Ecology, 10(3), 361–365.

[Samuels et al., 2010] Samuels, D., Meintjes, L., Ochoa, A., & Porcello, T. (2010). *Soundscapes: Toward a sounded anthropology*. Annual Review of Anthropology, 39, 329–345.

[Schafer, 1994] Schafer, R. (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT, Destiny.

[Schmidt & Balakrishnan, 2015] Schmidt, A. & Balakrishnan, R. (2015). *Ecology of acoustic signalling and the problem of masking interference in insects*. Journal of Comparative Physiology A, 201, 133–142.

[Sinsch et al., 2012] Sinsch, U., Lümkemann, K., Rosar, K., Schwarz, C., & Dehling, M. (2012). *Acoustic niche partitioning in an anuran community inhabiting an Afromontane wetland (Butare, Rwanda)*. African Zoology, 47(1), 60–73.

[Staaterman et al., 2013] Staaterman, E., Rice, A., Mann, D., & Paris, C. (2013). *Soundscapes from a Tropical Eastern Pacific reef and a Caribbean Sea reef*. Coral Reefs, 32, 553¬–557.

[Sueur & Farina, 2015] Sueur, J. & Farina, A. (2015). *Ecoacoustics: the ecological investigation and interpretation of environmental sound*. Biosemiotics, 8, 493–502.

[Sueur et al., 2014] Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2014). *Acoustic indices for biodiversity assessment and landscape investigation*. Acta Acustica united with Acustica, 100(4), 772–781.

[Tonolla et al., 2010] Tonolla, D., Acuña, V., Lorang, M., Heutschi, K., & Tockner, K. (2010). *A field-based investigation to examine underwater soundscapes of five common river habitats*. Hydrological Processes, 24(22), 3146–3156.

[Wiens, 2002] Wiens, J. (2002). *Central concepts and issues of landscape ecology*. Applying Landscape Ecology in Biological Conservation, 3–21.

[Xu et al., 2019] Xu, H., Plieninger, T., & Primdahl, J. (2019). *A systematic comparison of cultural and ecological landscape corridors in Europe*. Land, 8(3), 41.

Interaction between Past and Future: the Algorithmic Composition Library FDSDB_XXth_CT

Fabio De Sanctis De Benedictis

Conservatory P. Mascagni, Livorno, Italy

Abstract

According to Meyer, the musical style gradually evolves through the normalization of deviations and the creation of new ones. This raises the question of continuity and interaction between the music of the past and the future, in particular for the music of the last century and the current one. The writer's algorithmic composition library, FDSDB_XXth_CT,[1] for OpenMusic, devoted to compositional techniques of the twentieth century and beyond, can fit into this panorama, contributing to the continuity between past, present and future. After a brief description of general concepts, we move on to describe the overall structure of the library menus, the functions introduced in the updating phases, and those planned for the next release, exemplifying some patches. The compositional techniques globally refer to composers such as Schönberg, Berg, Webern, Babbitt, Boulez, Carter, Donatoni, Ligeti, Maderna, Manzoni, Messiaen, Nono, Stockhausen, Xenakis, and soon also Harvey and Eimert. The possible dangers inherent in an unaware use of these tools are also highlighted during the paper, and the benefits and applications are described in the conclusions, both for composition and for teaching, as well as possible uses for electronic composition.

1 Introduction

Meyer defines style as a system of relations between sounds, characterized by exact probabilistic relationships.[2] The most frequent procedures establish the normative components, the less frequent ones represent the deviations. Precisely because they are rarer, deviations create the expectation of solving a problem posed by their very occurrence, and thus produce an affective reaction, and a series of interactions – anticipations and feedbacks – between what the listener expects and what will actually occur in the unfolding of the music. The stylistic evolution occurs when the most experimental composers, impelled by a more fervent expression, use a larger number of deviations than the average, deviations that, in the long run, stabilize and become the normative procedures of the next stylistic stage. In Meyer's perspective, mainly aimed at the tonal system, we find an evolution that is essentially continuous. How can we reconcile this position with twentieth century music and current compositional production? Or again: what place can the compositional techniques of the recent past occupy in today's composition? The writer's algorithmic composition library, FDSDB_XXth_CT, devoted to compositional techniques of the twentieth century and beyond, proceeds from the reversibility of musical analysis and compositional techniques,[3] formalizing in algorithms a series of compositional processes derived from one's own music analysis or those of others, finally arriving at the realization

[1] FDSDB stands for the initial letters of my first name and my two surnames, XXth refers to the last century, CT stands for "Composition Techniques". You can download it from the following page: https://sites.google.com/site/fdsdbmascagni/code/fdsdb_xxth_ct-per-open-music-1.

[2] [Meyer, 1992], p. 77.

[3] [Bent, 1990], p. 2.

of a library for OpenMusic.[4] Compositional techniques, musical analysis and aesthetics according to Bent are in fact elements distributed along an axis. In this sense, the path from the finished work to its constituent cores is substantially the reverse in musical analysis and composition. That is, just as the analyst can come to define the main cores of a work, so the composer, starting from a few basic elements, can arrive at the finished work.

The library object of this contribution can have different purposes and applications, which will be illustrated at the end, but at the same time it could also arouse a certain compositional automatism, reducing the compositional techniques of the past to passive instruments, without a real awareness of personal compositional and aesthetic processes. Moreover, this danger is always present when we place ourselves in front of assisted musical composition and information technology tools in general.[5] However, we would like to make Donatoni's statement our own: «I am almost certain I can share the opinion according to which composing cannot be taught if with this expression one wishes to consider the inventive act considered in its entirety. [...] It may happen, however, that the dormant inventiveness of a young person needs any discipline in order to wake up. [...] It is within the framework of this discipline that [...] the means [...] that make up any compositional process can be taken into consideration.»[6]

Computer-assisted musical composition can then be considered a tool of rapid and precise development of the musical material and of the composer's intentions, and also a mean of preserving and transmitting a heritage of compositional techniques of the past, as well as a source of stimulus for "l'inventiva assopita", towards new solutions and personal inventions. Trying, we hope, to create an interactive bridge between past and present, towards the future, in order to recover that continuity which we mentioned earlier in relation to Meyer.

In this paper, after brief notions of a general nature, the structure of the library object of this dissertation is presented, illustrating some of the updates made with respect to what has already been described elsewhere,[7] and previewing those of the next release, to end with some conclusive considerations.

2 The assisted composition

Computer-assisted composition, or algorithmic composition, is the musical composition that uses algorithms, that is a series of unambiguous, detailed and finite instructions, which can be interpreted and executed with speed and accuracy by a computer program or by an automaton. The approach to this discipline can be constructive or declarative, that is the computer program can be used for the production of musical material, or the actual creation of a composition that responds to constraints imposed by the user.[8] Almost all the functions of the library object of this paper follow a constructive approach. However, by combining some functions together it is possible to arrive at the determination of complete compositions or compositional fragments, with a very similar approach to the declarative one, although without the aid of constraint programming.

3 The library FDSDB_XXth_CT

A library is a piece of software that adds new functionalities to an existing computer program. FDSDB_XXth_CT presents numerous functions related to the composition techniques of composers of the last century and beyond. With a few exceptions, it appears that extensive work has never been done in this area,[9] at least as far as we know. Its main menu is divided into: Dodecaphony, Serial-postserial music, PCST (Pitch-Class Set Theory), Rhythm, Pitch, Utilities. Each item is divided into submenus which include the techniques of composers who have used that particular system: under Dodecaphony we have the items Schoenberg, Berg, Webern, Babbitt; about serial and post-serial music we find Babbitt, Boulez, Carter, Donatoni, Ligeti, Maderna, Manzoni, Messiaen, Nono, Stockhausen, Xenakis; PCST includes functions for composing with pitch-class sets, mainly referring to the theoretical work of Robert Morris;[10] the Rhythm menu includes generic development tools for the duration parameter; under Pitch we have other tools for the pitch parameter; finally Utilities menu contains utilities to facilitate the work with the functions of this library.

The addressed compositional techniques are many, and each function is accompanied by a specific inline documentation, as well as specific examples that can be showed by selecting the function and

[4] OpenMusic is an algorithmic composition software developed and maintained at Ircam: http://repmus.ircam.fr/openmusic/home.

[5] See also what is discussed finally in [De Sanctis De Benedictis, 2018].

[6] [Donatoni, 1982], pp. 83-84. Translation by the author.

[7] We again refer to [De Sanctis De Benedictis, 2018].

[8] About constraint programming we refer to [Truchet – Assayag, 2011]. About the concepts of constructive and declarative approach to algorithmic composition we refer to [Agon – Assayag – Bresson 2006] e [Bresson – Agon – Assayag 2008].

[9] We can remember an algorithmic formalization of the compositional process of *Achorripsis* by Xenakis, made by Michail Malt in Open Music, and libraries like OMTristan or Esquisse that contain functions relative to spectral composition techniques.

[10] [Morris 1987, 1991, 2001].

pressing the "t" key on the computer keyboard. Tutorials are also available. The documentation and the examples show the bibliographic source from which the compositional technique was taken, subsequently formalized and verified in an OpenMusic patch, finally coded in LISP language to create the library function.
In a previous article we have already discussed version 1.0 of the library,[11] here, in order to avoid unnecessary repetitions, we would like to illustrate the steps taken to upgrade to version 1.1 and 1.2, as well as to mention some of the features that will soon be implemented in the next update. In Figure 1 the scheme of the main menu of the library and the sub-menu related to Messiaen's compositional techniques.

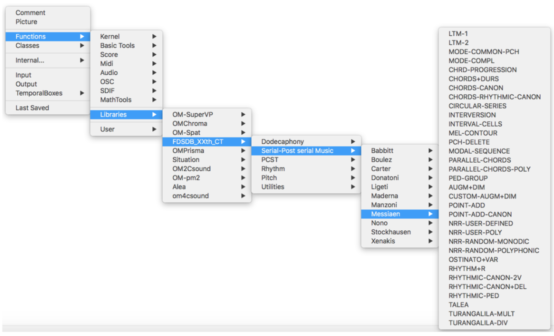


Figure 1: Library menu structure

4 Updates to version 1.1 and 1.2

In version 1.1, the work set up for a Masterclass held during the academic year 2020-2021 at the Sassari Conservatory – together with my friend and colleague Francesco Maria Paradiso, and focused on the figure of Messiaen – mainly merged. The writer's role concerned the implementation in OpenMusic of the compositional techniques described and the explanation of the mechanism of the involved functions. Some of those were already present in the first version of the library, this occasion gave the starting point and the possibility to expand them, to include almost all the compositional processes described in *Tecnica del mio linguaggio musicale*.[12] In Figure 2, for example, we find a patch relating to the automatic development of a rhythmic pattern similar to the 14th-century *Talea*.

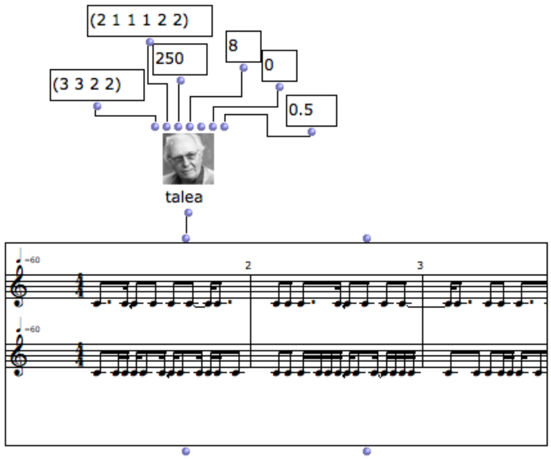


Figure 2: The talea function.

The function receives the following parameters as input:

- proportions of the duration of the first rhythmic module → (3 3 2 2)
- proportions of the duration of the second rhythmic module → (2 1 1 1 2 2)
- rhythmic unit of measure in milliseconds, here 250, equivalent to the sixteenth note with the quarter note set to 60
- last four parameters related to rhythmic quantization

In Figure 3 we can see the complete result in musical notation. The convenience of this function, and more generally of computer assisted composition, consists in the fact that to obtain other results we have just to change the initial parameters, without having to rewrite the whole algorithm from scratch.

Figure 3: The result of the previous function.

In the next update is already foreseen a function that allows to use not only two modules, as reported in the examples in the Messiaen treatise, but also more than two. In Figure 4 the preparation and verification patch for the subsequent addition of this function to the library.

[11] [De Sanctis De Benedictis, 2018].
[12] [Messiaen, 1944-1999].

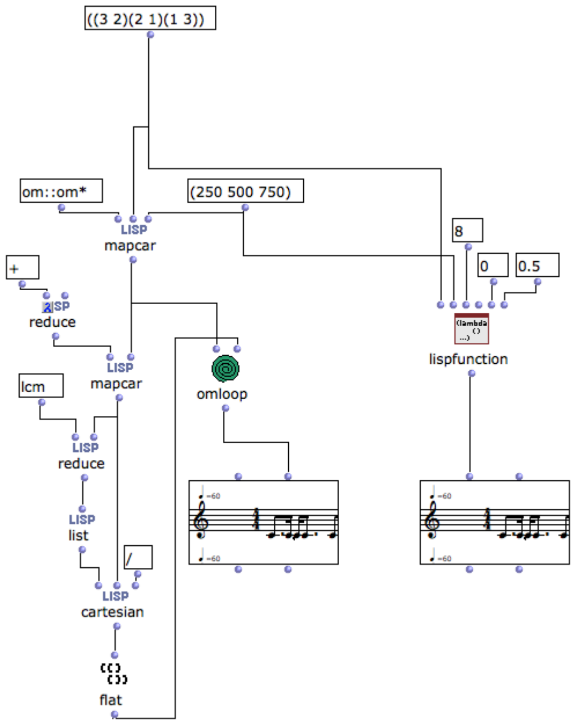
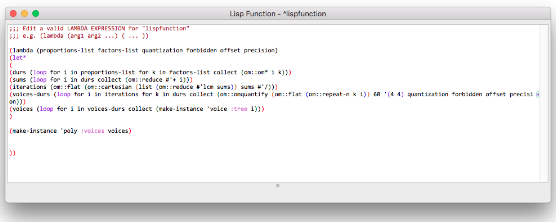


Figure 4: The patch relating to the next implementation of the talea-n function.

This is not the place to explain in detail all the functions used and the logic of the algorithm. We can simply say that in the left part of the figure we have the implementation of the algorithm through the native functions of OpenMusic or of the LISP language in general.[13] The function in the right part of the figure, denominated "*lispfunction*", contains the LISP code built on the model of the objects used in the previous algorithm. It is essentially the coded version of the "visual" algorithm, so to speak, of the left part of the figure. It will be this code, with a few necessary additions, to be used for the realization of the specific object of the library, *talea-n*. The resulting advantage consists in having a single object instead of ten or more functions of the starting algorithm, thus being able to develop more complex combinations with greater immediacy and in less visual space.
In Figure 5 we report the contents of *lispfunction*.



Lisp Function - ^lispfunction

```
;;; Edit a valid LAMBDA EXPRESSION for "lispfunction"
;;; e.g. (lambda (arg1 arg2 ...) ( ... ))

(lambda (proportions-list factors-list quantization forbidden offset precision)
(let*
(
(durs (loop for i in proportions-list for k in factors-list collect (om::om* i k)))
(sums (loop for i in durs collect (om::reduce #'+ i)))
(iterations (om::flat (om::cartesian (list (om::reduce #'lcm sums)) sums #'/)))
(voices-durs (loop for i in iterations for k in durs collect (om::omquantify (om::flat (om::repeat-n k i)) 60 '(4 4) quantization forbidden offset precisi »
on)))
(voices (loop for i in voices-durs collect (make-instance 'voice :tree i)))
)

(make-instance 'poly :voices voices)


))
```

Figure 5: The LISP code contained in the lispfunction object.

Figure 6 instead shows the musical result of the previous algorithm, which will be identical both evaluated through the functions of OpenMusic and through the LISP code.

Figure 6: Result of Figures 4 and 5.

The input parameters are, in order:

- the list of proportions lists, one for each voice → ((3 2) (2 1) (1 3))
- the list of multiplying factors for each voice → (250 500 750), i.e. sixteenth note, eighth note, dotted eighth note
- the rhythmic quantization parameters in the last four inputs

In the update to version 1.1 Messiaen's compositional techniques have been enriched to the point of having a situation like the following:

- limited transposition mode theory, including functions for detecting common notes between two modes and complementary sets
- various types of chord sequences such as chord canons, rhythm canons, rhythm pedal, modal successions, etc.
- various ways of managing regular and irregular rhythmic augmentation and diminution, including the addition of the point
- rhythmic canons of different types, by augmentation, by adding a point, by retrogradation, etc.
- different techniques related to non-retrogradable rhythms, including the creation of random rhythmic polyphonies or determined by user parameters
- various forms of melodic permutation, use of interval cells, circular series, repetition of a sequence of notes with progressive elimination of pitches from the end

In version 1.2 they have been implemented: quarter tone series, series with specific interval relations, as in *Klang* by Stockhausen, Babitt's time points system, the twelve-tone matrix in music notation, techniques derived from *Lumen* by Donatoni, tendency masks, techniques taken from *Canto Sospeso* by Luigi Nono, inversion in space of absolute pitches (P-Space), random asynchronous

[13] LISP is the programming language on which OpenMusic is based.

rhythms. In the next version, probably the number 1.2.1, in addition to the cited *talea-n*, they will be introduced techniques taken from:

- Jonathan Harvey (*Quartetto n. 1*, *Ritual Melodies*)
- an additional version of Ligeti's *Lux Aeterna* polyphony
- various forms of development of a twelve-tone series, desumed from Eimert and Berg

Similar work is expected in the future for some analytical methodologies.

5 Conclusions

About the possible purposes of this library we have already mentioned the dissemination, preservation and enrichment of the compositional techniques of the past, with a view to continuity with the future, encouraging an effective interaction between past, present and future. An additional advantage in adopting a this library, if you choose to use the described techniques, consists in freeing the composer from the tiring work of development of the compositional material, allowing him to focus more on results and processes, on music. A non-negligible aspect is then the use in teaching, being able to immediately show examples of even complex processes, listening to the sound result by means of the MIDI implementation possibilities given by the software. A further particular, and often fruitful, use may consist in combining these tools with audio synthesis and spatialization, using other OpenMusic libraries such as OMChroma, om2csound, om4csound, Modalys, OM-Chant, OM-pm2, OM-Spat, OM-SuperVP, OMPrisma, obtaining potentially unprecedented results.

References

[Agon – Assayag – Bresson 2006] Agon C. – Assayag F. – Bresson J. (2006). *The OM Composer's Book. Volume One*. Delatour/Ircam.

[Bent, 1990] Bent I. (1990). *Analisi musicale*, EDT.

[Bresson – Agon – Assayag 2008] Bresson J. – Agon C. – Assayag F. (2008). *The OM Composer's Book. Volume Two*. Delatour/Ircam.

[De Sanctis De Benedictis, 2018] De Sanctis De Benedictis F. (2018). FDSDB_XXth_CT: una libreria di composizione algoritmica per Open Music e PWGL dedicata alle tecniche compositive del Novecento. In *XXII Colloquio di Informatica Musicale, Atti della Conferenza* (pp. 122-127). Udine, Italia.

[Donatoni, 1982] Donatoni F. (1982). Processo e figura. In Donatoni F. *Il sigaro di Armando* (pp. 83-86). Spirali edizioni.

[Meyer, 1992] Meyer, B.L. (1992), *Emozione e significato nella musica.* Il Mulino.

[Messiaen, 1944-1999] Messiaen O. (1944-1999). *Tecnica del mio linguaggio musicale*. Durand. (Italian translation by Lucia Ronchetti).

[Morris, 1987] Morris R.D. (1987). *Composition with Pitch-Classes: A Theory of Compositional Design*. Yale University Press.

[Morris, 1991] Morris R.D. (1991). *Class Notes for Atonal Music Theory*. Frog Peak Music.

[Morris, 2001] Morris R.D. (2001). *Class Notes for Advanced Atonal Music Theory*. Frog Peak Music.

[Truchet – Assayag, 2011] Truchet, C. – Assayag, G. (2011), *Constraint programming in music*, Wiley.

The ASMA Tool-Suite: Augmenting Singing Instruction of Elementary School Students

Natalia Kotsani[1], Giorgos Dedousis[1], Evangelos Angelakis[1], Areti Andreopoulou[1] and Anastasia Georgaki[1]

[1] *Laboratory of Music Acoustics and Technology (LabMAT), National and Kapodistrian University of Athens, Greece*

Abstract

The ASMA project (Assistance for Singing and Music Aesthetics) studies the social and aesthetic importance of vocal training in elementary schools, under the prism of contemporary teaching methods and the informed use of assistive educational tools. The aim of this project is the development of an interactive educational tool-suite supporting vocal instruction and correct singing practices in the Greek elementary school music courses. This work presents the revised ASMA Tool-Suite in its entity, after the necessary revisions that followed its original testing. The ASMA tool-suite consists of (a) Vowel Maps and Phonetograms, (b) Vocal Quality Tools, (c) Rhythmic and Melodic Tools, and (d) Polyphonic Tools, offering a wide range of vocal analysis and tuning tools, as well as vocal warm-up and training exercises. The Tool-Suite is accompanied by a "Guidebook of Proper Vocal Practice and Instruction for Elementary School Teachers", providing teachers with a fundamental scientific, aesthetic, empirical, and technological background, necessary to better incorporate singing in the elementary school classroom.

1 Introduction

The act of singing is a very powerful tool in the hands of music teachers as it is a natural way of introducing students to, otherwise rather complex, music concepts, such as melody, rhythm, dynamics, form etc, while also promoting the children's well-being [Welch et al., 2014]. Nevertheless, vocal instruction in Greek elementary schools is challenging for several reasons, including but not limited to i) the reduced time music instruction occupies in the weekly curriculum (once a week), ii) the lack of specialized classrooms, acoustically treated and properly equipped for music lessons, and iii) the fact that the majority of music teachers does not have the proper expertise to approach children vocal instruction effectively [Sotiropoulou Zormpala et al., 2015]. Based on the above observations, it becomes evident that music instruction in Greek elementary schools needs to be reformed through the use of a modern educational model, employing contemporary teaching practices and educational tools which will tailor music instruction to needs of the students [Stavropoulou et al., 2014].

Over the past two decades there has been an increase in the number of academic or commercial software tools enriching music instruction and providing opportunities for personalized learning experiences to students. There exists evidence showing that such tools can widen the students' vocal range and help them become conscious of their voices [Fuchs et al., 2009]. These tools monitor a wide variety of parameters, such as pitch / tonal accuracy, rhythm precision, formant analysis, spectral qualities, or even electroglottograph (EGG) indications, in an attempt to affect the students' singing voice development and musicianship, in general [Tu, 2020]. Examples of such tools include but are not limited to Voce Vista[1] and AIRS (Advancing Interdisciplinary Research in Singing) Test Battery of Singing Skills [Gudmundsdottir & Cohen, 2015]. For a comprehensive overview of such software, including both scientific prototypes and commercial products, see [Andreopoulou et al., 2021].

ASMA (Assistance for Singing and Music Aesthetics)[2], is a 2-year research project, funded by the H.F.R.I. (Hellenic Foundation for Research and Inno-

[1] https://www.vocevista.com/
[2] https://asma.music.uoa.gr

vation), which studies the social and aesthetic importance of vocal training in elementary schools, under the prism of contemporary teaching methods and the informed use of assistive educational tools. Within this scope, ASMA has focused on the development of an interactive educational tool-suite supporting vocal instruction and correct singing practices in elementary school music courses. Prototype versions of selected tools from the project have been published in [Andreopoulou et al., 2021, Angelakis et al., 2021, Kotsani et al., 2021]. This work will present the ASMA Tool-Suite in its entity, after the necessary revisions that followed the Alpha testing period, during which, constructive feedback was collected from members of the extended project team, the external project advisors, and selected user groups.

2 The ASMA Datasets

Three distinct periods of audio sample collection from students and adults speech and singing voices, choir, or solo recordings, resulted in the creation of a rich dataset with annotated audio samples. More specifically, the first two periods consisted of solo student voices, recorded and used for the design and implementation of the analysis and assessment tools of solo voices, and the third included recordings of selected children choirs, which was used for the design and implementation of tools assessing accuracy and coordination in polyphonic settings. An additional dataset of warm-up vocal exercises, melodic and rhythmic exercises, and scales in various tuning systems was compiled for the implementation and enrichment of the Melodic and Rhythmic Tools.

3 The ASMA Tools

The initial stage of the ASMA project included the creation of a "Guidebook of Proper Vocal Practice and Instruction for Elementary School Teachers", which accompanies the proposed manual and digital tools listed within the next subsections. This guidebook is intended to assist teachers, by providing them with the fundamental scientific, aesthetic, empirical, and technological background, necessary to better incorporate singing in the elementary school classroom. The guidebook objective is two-fold: i) to help teachers acquire a more concise knowledge and use of their own voice, and ii) to equip them with information, knowledge, educational strategies, and music tools, enhancing and enriching the singing instruction of their young pupils. This edification of the Elementary School Music teachers is intended to make the music lesson at school more efficient, enjoyable, motivating, and engaging, as well as safer both for the children's and the teachers' voices. To this end, the guidebook provides i) a thorough overview of the vocal instrument (anatomy, physiology, acoustics) through a practical approach, ii) edification on pertinent learning theories, psychology matters, methodological approaches, and advice on vocal health, and also contains iii) an overview of vocal training basics (exercises, repertory suggestions, etc.). The Guidebook is supplemented by a database of appropriate music material for use in the classroom.

3.1 Phonetic Maps and Phonetograms

This set of tools targets basic characteristics and qualities in speaking and singing, the assessment of which is fundamental in any attempt to characterize an individual's voice. The *Phonetic Map tool* and the *Formant Range Profile (FRP) tool*, take as input a student's voice pronouncing the Greek vowels, and outputs a visualization of the user's vowel pronunciation. When the spoken vowels are used as input, the tool extracts the first two formants (F_1 by F_2) from the audio files and outputs an F_1 by F_2 static plot (Phonetic Map). When the sung vowels are used as input, the user can choose to see how the first two formants change over time. The Phonetic Map and the FRP tools, help music teachers identify the way children articulate each vowel when speaking and how it is positioned in the articulatory space. In addition, the FRP tool can be used for studying the homogeneity in a student's singing voice timbre among the various voice registers or among other students' voices in a choir (see Section 3.4. The output of FRP tool can be seen in Figure 1.

The Vocal Range Profile (VRP) tool, otherwise known as a *Phonetogram*, takes as input a variety of students' vocal recordings gradually reaching their vocal range extremes. The more information is given as input, the denser a student's Phonetogram is. The Phonetogram tool helps music teachers determine the vocal range of their students, as it visualises the calculated pitch (fundamental frequency) by its dynamic level (sound intensity).

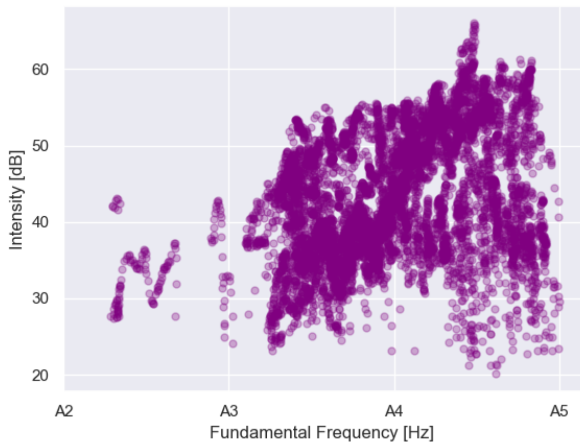


Figure 1: *The Output of the Phonetogram Tool*

3.2 Vocal Quality Tools

For the analysis of the students' vocal quality, two novel tools for quantifying breathiness and nasalitation in singing were implemented [Angelakis et al., 2021, Kotsani et al., 2021]. The *Breathiness tool* quantifies the breathiness characteristic of the students' singing voice using as a parameter the Smoothed Cepstral Peak Prominence (CPPS). In addition, using multiple linear regression, a new index (CDH) for the quantification of the breathiness characteristic is being introduced, which is based on audio samples and EGG signal results from the CPPS, DOQ, and HOQ parameters.

The *Nasality tool* quantifies the nasality characteristic of the students' voice based on the formant central frequencies and bandwidths. Both tools take as input an audio sample of a student's voice pronouncing the 'A' vowel, and output a rating of the nasality and breathiness quality of their voice.

3.3 Rhythmic and Melodic Accuracy Tools

This set of tools, which is developed in Max/MSP, targets the improvement of the students' tonal and rhythmic accuracy, through practice exercises, tuning tasks, and real-time qualitative and quantitative feedback.

The *Digital Warm Up Tool* (Figure 2) loads midi files of melodic and rhythmic exercises in a variety of scales, tempos, and levels of difficulty to be used by the students for properly and effectively warming up their voices prior to singing. A database of such exercises (Section 2) is already pre-loaded in the tools, but the collection can be further expanded by new, user-defined entries. All exercises have recommended presets for the tempo and the optimal range of transpositions. Nevertheless, users can also choose to change those to their preferred settings. The tool operates in two modes: "student-view", which displays only the score of the exercise to be sang, and "teacher's view", which includes the melodic line accompaniment of each exercise with a pre-count bar, as well as real-time visual feedback of the student's voice spectral characteristics and their tonal accuracy. Both modes offer the same parametrization option to the users. The students can either just read the score and sing or can choose to listen to the exercises in real-time via head-phones.

The *Beat Tracking Tool* operates in two modes: as a regular metronome and as a rhythmic trainer. In its first functionality, the tool operates as a regular metronome, which the students can use to maintain a steady paste as they sing or train. Users can parametrize the frequency of the metronome tone, as well as its speed. In its second functionality, the tool allows users to load rhythmic exercises, either from a wide variety of presets (Section 2) or user defined, and helps them train their rhythmic accuracy by providing them visual feedback of their deviation from the target onset times. In the "teacher's mode" the tool also displays statistics of the overall accuracy of the students as they practice a certain rhythmic pattern.

The *Tuner Tool*, operates in two modes too: as a regular tuner and as a musical interval trainer. As a regular tuner, it lets users set a reference tone in Hz, and monitor their tonal accuracy and deviation from the target tone, as they attempt to sing that reference. Users have control over the frequency, timbre, and duration of the reference tone. Deviations from the target are displayed on the tools graphically (color coded) as well as numerically in cents. As a musical interval trainer the tool is designed to respond to various tuning systems, including but not limited to "European-equal tempered", "European-modal", "Byzantine-modal", "Ancient Greek", as well as custom tunings defined by the users. Users are supposed to select a tuning system, as well as a certain interval or melodic phrase from a collection of presets, and attempt to sing it, while maintaining the interval characteristics of that system. Onve again, deviations from the proper tuning are displayed on the tool graphically as well as numerically.

Both the *Beat-Tracking* as well as the *Tuner* tools, can function either as stand-alone applications or in conjunction with the the *Digital Warm-up Tool* providing additional real-time feedback to the users singing melodic or rhythmic practice exercises.

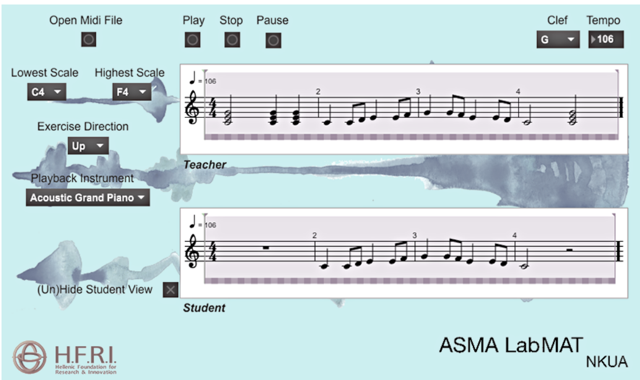


Figure 2: *The ASMA Tool-Suite Digital Warm-up Tool*

3.4 Coordination and accuracy in polyphony

This set of tools is designed to assist teachers monitor the singing accuracy and enhance the coordination of students singing in large ensembles. More specifically, the *Vocal Homogeneity Tool* is programmed to help teachers quantify the vocal homogenity of a small ensemble or choir, based on the students phonetic map and vocal data. With the help of simple vocal measurements as described in Sections 3.1 and 3.2 teachers can identify where the vocal properties of each student lie in comparison to the rest of the ensemble and help them solve individual issues in a targeted manner.

The *Tonal Accuracy in Polyphony Tool* uses auditory spatialization in order to surround users with virtual choir-related auditory information, to help them train in more realistic situations. It assists vocal tuning when singing in a choir, a setting which is normally polyphonic. In such a setting, the position of a choir member within a large ensemble is directly related to what they hear around them, and as a result to what cues they may use in order to be in tune and coordination with the others. With this tool, the aim is to virtually place the users, who will wear headphones, in various spots within a choir, helping them realize the changes in auditory cues they may perceive, and, in a way, training them for a smoother integration with the choir.

4 Future Work

The ASMA Tool-Suite opens up new possibilities in singing education in Greek public schools, offering teachers the additional tools to assess student voices, and, more importantly, to engage students in exploring the foundations of artistic expression using their voice. For the purpose of developing ASMA digital tools, the human voice has been regarded as a system consisting of separate, albeit interconnected, physiological mechanisms. However, the practical-empirical scope, under which the accompanying Guidebook presents the necessary background for a healthier and more efficient vocal production, attempts an approach involving the whole individual through kinesthetic learning and cognitive functions that promote voice perception and awareness [Leman, 2007]. This is also the direction toward which the ASMA tools will evolve in the future. In its current state the ASMA Tool-Suite has been re-designed after having completed its Alpha testing by the extended team of the project, external project advisors, and selected users. Currently, we are in the final stage of the project, during which the tools are being extensively tested by music teachers. A publication of the evaluation results will follow along with a public release of the tools. Future extensions of the ASMA Tool-Suite will include (i) the transfer of all tools to a web-based platform, (ii) gamification possibilities, to assist teachers engage students in using the tools in the classroom, and (iii) more options for parametrization.

5 Acknowledgments

The research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the “First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant” (Project Number: HFRI_FM17_TΔ E_3832).

The authors would like to thank Prof. Iakovos Steinhaouer, Dr. Sofia Stavropoulou and Kostas Katsantonis, for their valuable contribution to this work.

References

[Andreopoulou et al., 2021] Andreopoulou, A., Kotsani, N., Dedousis, G., & Georgaki, A. (2021). Evaluating the vocal characteristics of elementary school students: Basic assessment tools and methodology. In *Interaction Design and Children* (pp. 216–223).

[Angelakis et al., 2021] Angelakis, E., Kotsani, N., & Georgaki, A. (2021). Towards a singing voice multi-sensor analysis tool: System design, and assessment based on vocal breathiness. *Sensors*, 21(23), 8006.

[Fuchs et al., 2009] Fuchs, M., Meuret, S., Thiel, S., Täschner, R., Dietz, A., & Gelbrich, G. (2009). Influence of Singing Activity, Age, and Sex on Voice Performance Parameters, on Subjects' Perception and Use of Their Voice in Childhood and Adolescence. *Journal of Voice*, 23(2), 182–189.

[Gudmundsdottir & Cohen, 2015] Gudmundsdottir, H. R. & Cohen, A. J. (2015). Advancing interdisciplinary research in singing through the AIRS Test Battery of Singing Skills. *Musicae Scientiae*, 19(3), 234–237. Publisher: SAGE Publications Ltd.

[Kotsani et al., 2021] Kotsani, N., Angelakis, E., & Georgaki, A. (2021). Evaluating the nasalisation of the singing voice. *Models and Analysis of Vocal Emissions for Biomedical Applications: 12th International Workshop*, (pp. 119).

[Leman, 2007] Leman, M. (2007). *Embodied music cognition and mediation technology*. MIT press.

[Sotiropoulou Zormpala et al., 2015] Sotiropoulou Zormpala, M., Trouli, K., & Linardakis, M. (2015). Arts education offered by Greek Universities to future pre-school and primary school teachers. *Preschool and Primary Education*, 3(1), 34–52. Publisher: Laboratory of Pedagogical Research & Applications.

[Stavropoulou et al., 2014] Stavropoulou, S., Georgaki, A., & Moschos, F. (2014). The effectiveness of visual feedback singing vocal technology in greek elementary school. *ICMC-SMC 2014*, (pp. 1786–1792).

[Tu, 2020] Tu, C. M. (2020). Correlations among Music Aptitude, Singing Voice Development, and Singing Accuracy Achievement in Young Children. In F. A. Russo, B. Ilari, & A. J. Cohen (Eds.), *The*

Routledge Companion to Interdisciplinary Studies in Singing, Volume I: Development, volume 1 (pp. 520). Routledge, 1st edition.

[Welch et al., 2014] Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., & Sarazin, M. (2014). Singing and social inclusion. *Frontiers in psychology*, 5, 803.

CrazySquare: an Interactive Tool

Federica Caruso[1], Tania Di Mascio[1], Marco Pennese[2] and Sara Peretti[3]

[1] *DISIM, University of L'Aquila, Italy*
[2] *Scuola Media G. Pascoli, Rieti, Italy*
[3] *Centre of Excellence DEWS, University of L'Aquila, Italy*

Abstract

In a world in which a large amount of learning interactive tools deal with the musical abilities, without really supporting a high level of music education, CrazySquare, inspired by the Gordon's theory, guarantees the acquisition of the learning guitar objectives that students should learn at the end of Italian Middle Schools. Initially implemented as a paper-and-pencil game, supporting the study of rhythm and its representation, currently, CrazySquare is an interactive tool for learning guitar. It is designed according to a multidisciplinary Technology-Enhanced Learning User-Centered Design methodology, which allows to adequately integrate the technology into teaching activities.

This paper, after presenting the current prototype of the interactive tool, briefly reports the conducted expert-based evaluation and presents the ongoing user-based evaluation protocol.

1 Introduction

Millions of people listen to music every single day. As a consequence, learning to play musical instruments is something that many people have interest in. Moreover, it is recognized by the literature that musical learning (music education and learning to play musical instruments) enhances a wide range of cognitive and affecting functions, such as language and social cognition [Miendlarzewska & Trost, 2014]. Nevertheless, learning to play musical instruments is not an intuitive and automatic task; in fact, many people consider this a challenging task that requires a lot of time and adequate guidance [Miendlarzewska & Trost, 2014]. This is especially true for the guitar that it is one of the most popular musical instrument in Italian Middle Schools [Caruso et al., 2020].

Existing tools for learning guitar (pure commercial products, scientific literature outcomes, and research prototypes become commercial products) adopt either one of or both the two approaches: (1) a more experience-oriented approach supporting the instrumental practice and (2) a more lesson-oriented approach including music theory (e.g., ear-training) and "how to play" modules. Most popular solutions follow the approach (1). Commercially based solutions are generally Software (SW) applications assisting the users during the execution of exercises with the guitar, providing real-time feedback of their performance, thus improving the self-evaluation procedures. Research-based prototype solutions are mainly based on augmented reality technology. All the solutions following the approach (2) are SW applications implementing lessons trough video-tutorials made by some teachers or musicians who explain in detail how the learners have to properly perform practical exercises. Solutions mixing the two approaches are commercial SW applications that not only provide lessons through video-tutorials but also interactively implement them, proposing "question and answers" modules. For a detailed discussion see Table 1.

Unfortunately, as discussed in [Beck, 2017], existing tools are often not adequately evaluated (especially in the case of commercial ones) and do not offer the high level of music education that is guaranteed, at school, by the expertise of teachers following consolidated music-oriented pedagogical approaches (e.g., [Gordon, 2007, Jaques-Dalcroze, 1921]). In particular, Gordon proposed a music learning theory based on the concept of *audiation* [Gordon, 1989].

Furthermore, [Konecki, 2014] observed that although research shows the positive impact of interactive tools in music education, most of the existing ones have limited effect due to the lack of individualization of the learning process. Indeed, these issues are also highlighted by the [Decree law no. 201, 1999] that organized the music education in Italian Middle Schools. This decree explicitly describes the musical education skills that students should achieve at the end of Middle School (melody, harmony, rhythm, timbre, dynamic, agogic, and instrumental skills). Furthermore, the adequately controlled adoption of tools made available by modern technologies is defined to be methodologically effective. Although the Decree Law points out that one of the main objectives of music education is to guarantee musical literacy, (1) no specific guidelines are provided in terms of which pedagogical approaches and interactive learning tools should be applied to achieve the cognitive and instrumental skills and (2) no specific guidelines on the interactive learning tools could be used to achieve these skills.

According to new technologies that incorporates concepts coming from games to boost students' motivations and interests [Denis & Jouvelot, 2005], CrazySquare aims at filling these lacks, by designing a game-based interactive tool for music education, inspired by one of the most well-consolidated pedagogical theories to music education (i.e., the Gordon's Music Learning Theory (MLT) [Gordon, 2007]).

The tool is described in Section 2 both from the the learning model and the interaction tool viewpoints. Section 3 reports the conducted expert-based evaluation and presents the ongoing user-based evaluation protocol. Finally, in Section 4, conclusion and future works are described.

2 CrazySquare

CrazySquare was conceived as a teachers' ally in music teaching activities in Italian Middle school, hence utilized by students in the 10-14 age range. Currently, CrazySquare is dedicated to the guitar since it is one of the most popular musical instrument in Middle Schools. To provide a "way" to effectively integrate this technology into teaching activities, CrazySquare is underpinned by the Technology-Enhanced Learning oriented User-Centered Design (TEL-UCD) approach [Di Mascio et al., 2016] that guarantees the compliance of CrazySquare as an interactive learning tool with Gordon's MLT. The TEL-UCD methodology expands the traditional iterative UCD approach to emphasise the necessity of designing in parallel both a learning intervention and the interactive tool, realising it in a context of mutual dependency, along with the choice of a learning theory.

After the identification of learner needs, the specification of the context of use and the identification and specification of psycho-pedagogical assessment strategy is necessary, followed by the specification of requirements, the definition of the learning model, and the creation of the learning material. Additionally, the evaluation of the prototyping solutions is mandatory, both from the learning model and the usability viewpoints. At the end, the realisation of the tool solution, consists of the realization of both interactive tool and the learning intervention.

2.1 The CrazySquare learning Model

According with the TEL-UCD, we regarded the specific teaching method adopted, within the Italian Middle School, as a possible instantiation of a more general model capturing its specifics, as well as its principles, from consolidated music-oriented pedagogical approaches, starting with Gordon's MLT. In [Gordon, 1989], the author underlines the distinction between *music aptitude* and *music achievement*, where the former mainly represents one's potential to learn to *audiate* whereas the latter mainly represents what one has learnt to *audiate*. An adequate interactive tool supports students with low music aptitude as well as those have high music aptitude, the former should not become frustrated by the difficulty of proposed exercises, whereas the latter should not become bored by the simplicity of proposed exercises. The CrazySquare Musical Skill Learning Model, namely MuS-LM, hence embed the two concepts of achievement and aptitude, introducing gamification elements to ameliorate the learning intervention and to making it more motivating and engaging [Margoudi et al., 2016].

In particular, students have to achieve a set up of (interdependent) *musical skills*:

- **R - Rhythm skill**: Perceive and maintain the pulsation for predefined Beats Per Minute (BPM) value;
- **RS - Rhythm Symbols skill**: Recognize and execute by reading a sequence of rhythmic symbols;
- **MN - Musical Notes skill**: Play with the instrument musical notes, articulating them through a reading of rhythmic symbols;
- **CC - Chord Change Speed skill**: Execute change chords at different speeds.

These skills are in turn achieved by acquiring competencies of two levels of difficulty: base and advanced (from now denoted as 1 and 2). The learning intervention is then represented by the competencies' list to be acquired in a certain order, that is: (R^1, R^2, RS^1, MN^1, CC^1, RS^2, MN^2, CC^2). Thus, such competencies will be acquired by doing sets of homogeneous learning exercises; the number of these to do is not fix: students with low music aptitude have to perform the set of mandatory one, while students with high music aptitude can perform also the set of optional ones. After a specific number of learning exercises, all students relax themselves by playing relaxing mini-games (a type for each musical skill), which serve precisely as rewarding mechanism, with the dual aim of lowering the cognitive load, preventing its overload, while entertaining and engaging students with riddles and quizzes. Mini-games are not the only gamification elements; other elements (e.g., levels, stars, rewards, unlockable content) are introduced in the CrazySquare tool to improve the interactivity of the tool itself.

Specifically, taking advantage of the game-based approach, each learning exercise is itself a game level and, consequently, a specific set of learning exercises address a specific musical skill. As better a students perform learning exercises (i.e., game levels), as

Table 1: *Analysis of existing interactive tools for learning guitar (C - Commercial product, R - Research outcome) against the main requirements of CrazySquare: (U1) Mastery of playing guitar; (U2) Adaptive and motivating learning intervention.*

Approach	Type	Name	Application	Learning Intervention	U1	U2
(1)	C	SmartMusic [SmartMusic, 2022]	Web application	The system provides different exercises (i.e., music sheets) to be executed with the guitar by the student. The system assists the students during the execution of exercises, providing also real-time feedback on their performance.	Yes	-
		RockSmith [Rocksmith, 2022]	Music videogame for computer, PlayStation, and Xbox	The player plugs in virtually any electric guitar to the device and plays the songs included in the game. The system assists the player during the execution of songs, providing also real-time feedback on their performance	Yes	Yes
	R	GuitarSolo [Seol et al., 2016]	AR-based system including an Arduino Mega2560 as the main controller board, a LED guitar fretboard for fingering display, and a smartphone application to select the songs to play.	The system allows students to practice guitar, without a music sheet, by lightening the LEDs on the corresponding positions at the guitar fretboard. The augmented guitar is able to identify that the user is playing correctly by using the sensing information and then providing real-time feedback.	Yes	-
		GuitAR [Löchtefeld et al., 2011]	AR-based system including a light-projector system to be mounted on the guitar fretboard (or in front of the player) and a mobile application.	The system allows students to practice guitar, without a music sheet, by projecting lights on the corresponding positions at the guitar fretboard. The system is able to identify that the user is playing correctly by using the audio heard from the mobile device running the application and then providing real-time feedback.	Yes	-
	Both	Novaxe [Novaxe.com, 2022]	Web application	The system provides different exercises (i.e., guitar tablature) to be executed with the guitar by the student. The system assists the students during the execution of exercises, providing also real-time feedback on their performance.	Yes	-
(2)	C	GuitarTricks [GuitarTricks, 2022]	Web application providing video lessons for learning guitar	The system offers different video lessons.	Yes	
		FenderPlay [FenderPlay, 2022]	Mobile application	The system offers different video lessons.	Yes	-
	Both	Rockway [Rockway, 2022]	Web application	The system offers different video lessons.	Yes	-
(1)+(2)	C	MelodiQ [MelodiQ, 2022]	Android mobile application	The system provides a series of video lessons including also an interactive practice activity. The system is able to provide real-time feedback on the player performance.	Yes	Yes
		Yousician [Yousician, 2022]	Cross-platform application	The system provides video lessons including also an interactive practice activity. The system is able to provide real-time feedback on the player performance.	Yes	Yes

higher the rewards student obtains in terms of gained stars.

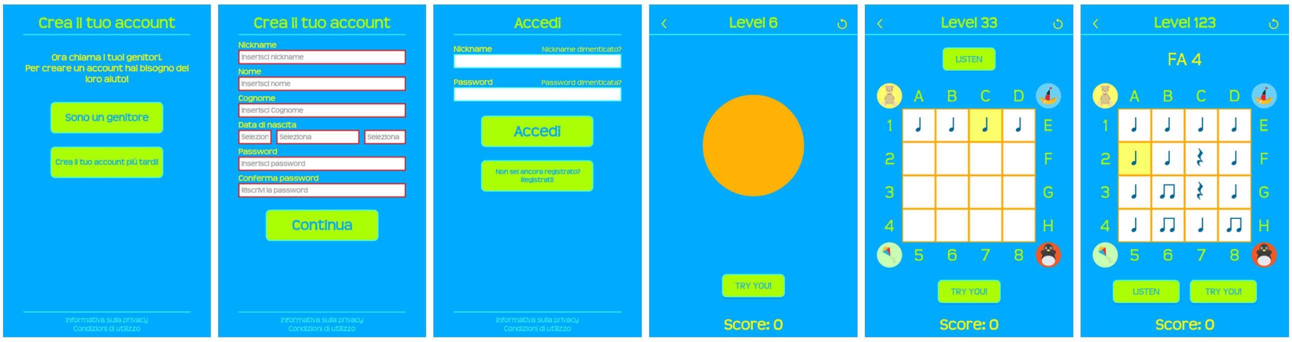
2.2 The CrazySquare interactive tool

CrazySquare was implemented based on the so described MUS-LM founding model as an Android mobile application using Android Studio 3.2 and exploiting several libraries (e.g., Tarsos-DSP and MIDI library). Its Grafical User Interface (GUI) is compliant with the main guidelines and suggestions concerning preadolescents' GUI design [Nielsen, 2019]; for example the used language is clear, simple, and close to teen slang (see Fig. 1).

Each student need a personal account to play. Then, the system provides both sign-up (see Fig. 1a and Fig. 1b) or sign-in (see Fig. 1c) options. For the sign-up, the system provides a two-step procedure, also involving the student's parents (as indicated by [EU-GDPR – Art. 8, 2016]).

With its personal account, the student can start a learning intervention according to the MuS-LM, which is articulated into incremental difficulty game levels (i.e, learning exercises). For example, game levels addressing R deal with the learning of predefined rhythms, that are easier in R^1 (i.e., 70, 75, and 65 BPM) and more advanced in R^2. The student has to perceive and keep the rhythm by executing a number of musical beats clapping their hands, vocalizing, or pitching a guitar string. As a hint, the rhythm is visualized by a bouncing circle placed at the center of the GUI provided by this kind learning exercise (see Fig. 1d); in addition, there is also an acoustic metronome to help student keeping the rhythm. The game levels addressing RS, MN, and CC are similar to each other; the student has to recognize and execute by reading a sequence of rhythmic symbols following a specific BPM that is displayed through a visual metronome. These symbols are presented within a four-by-four square grid (that inspired the crazy square name). The visual metronome is shown as a yellow highlighter of the rhythmic symbol to play (see Fig. 1e and Fig. 1f). In particular, within the game levels dealing with RS, the student performs the execution of rhythmic symbols (basic in RS^1, advanced in RS^2) by clapping hands, with the voice, or by pitching a guitar's string. Instead, within the game levels addressing MN, the student has to articulate the execution by playing, with the guitar, predefined musical notes (basic in MN^1, advanced in MN^2). The game levels relating to CC differ from those provided by MN only on the number of rhythmic symbols to be played with the same musical note. In fact, within MN's game levels, the change of chord happens every four symbols; instead, within CC^1's game levels the change happens every two symbols, and in CC^1's game levels the musical note changes at every symbol.

During the execution of the task provided by each game level, the system assists the student via different visual/acoustic hints as well as evaluates



(a) (b) (c) (d) (e) (f)

Figure 1: *Some screenshots of the current CrazySquare mobile application. (a-b) Sign-up, (c) Sign-in, (d) GUI of R's game levels, (e) GUI of RS's game levels, (f) GUI of MN's game levels*

if he/her is performing well or not. Coherently, at the end of each game level, CrazySquare shows the student the number of stars gained according to his/her performance. Once all the mandatory game levels related to a specific competence were passed, the system allows the student to continue along the learning intervention accordingly to the MuS-LM, i.e., the student can decide if play optional game levels provided by this competence or continue along the learning intervention, as indicated in the MuS-LM, by playing the game levels related to the following competence. As indicated by the MuS-LM, CrazySquare regularly intersperses the learning intervention with four types of relaxing mini-games, according to the musical skills being learned. For example, whether the student is playing with R's game levels, the system proposed a relaxing mini-games consisting of watching two different videos of people playing music (without sound) and the student has to choose which one of the two plays with the fastest or the slowest rhythm by tapping the corresponding button on the screen. As another example, during the learning intervention related to RS, CrazySquare proposes a riddle, i.e., the system plays a sequence of four rhythmic symbols and the student has to find the corresponding sequence on a row, a column, or a diagonal of the crazy square and select his/her answer by tapping it on the display. A video demo and the android Application Package (APK) of the current prototype of CrazySquare are available at `https://drive.google.com/drive/folders/1sSw-QTZZrdnUDBHwM0ssD4YOY39JKhHT?usp=sharing`.

3 The CrazySquare Evaluation

The evaluation of the CrazySquare interactive tool is currently articulated in one expert-based evaluation (already conducted, see for details [Caruso et al., 2020]) and several user-based evaluations, to be conducted. The expert-based evaluation aim was assessing the adequacy of learning model and the usability of preliminary prototypes. Experts (i.e., music teachers, guitar teachers, musical pedagogues, involved within the project) evaluated the MuS-LM and the correlated learning material consisting of 574 learning exercises (350 mandatory and 224 optional) and about 100 relaxing mini-games, while experts of visual design evaluated the GUIs prototypes using heuristic evaluation and expert review. A number of issues came up and successively fixed (e.g., the choice of how to display the visual metronome in the GUI).

The user-based evaluation aim is evaluating if the CrazySquare interactive tool is better/worse/equal than its paper-and-pencil version, in terms of students' performance. A between-group (true) experimental design is to be arranged and conducted. Its protocol will be administrated during regular school activities:

1. Students are randomly divided into two groups, (G_a and G_b), matched for age and sex.
2. At the beginning of the school year, before starting the learning intervention, all students (of both groups) perform two tests that provide the psycho-pedagogical assessment strategy: (1) an entrance test to assess the students' starting musical skills; and (2) the Intermediate Measures of Music Audiation (IMMA) [Gordon, 2007].
3. During the school year, G_a adopts the paper-and-pencil version of CrazySquare while G_b the CrazySquare interactive tool.
4. At the end of the school year, both groups perform again both tests.
5. The statistical analysis to evaluate the differences between G_a and G_b in terms of students' performance, musical aptitude, and musical achievement levels, has to be carried out.

4 Conclusion and Future works

In this paper, we introduced and discussed the current status of CrazySquare that aims to provide a game-based interactive tool to support the teaching/learning of guitar musical skills by preadolesences attending Italian Middle Schools, inspired by Gordon's MLT. A multidisciplinary research team was involved within each step of the CrazySquare project.

The strength of the CrazySquare, unlike the available interactive tools for music education, is that it aims to guarantee the acquisition of the learning objectives that students should learn at the end of Italian Middle Schools. In addition, being designed following multidisciplinary TEL-UCD methodology, it could be easily integrated into classroom practices as a support during the musical teaching activities and at home, as a support of homework activities.

During the last school year, we conducted a (pilot) controlled experiment involving a small group of first-year students (i.e., 6 students, 4 males and 2 females) attending the "*Giovanni Pascoli*" Middle School in Rieti (Italy). However, due to the small sample size and the limited continuity of the experimental activities due to COVID-19 related issues, the so far data collected is not enough to draw accurate conclusions. In the next future, another user-based evaluation with a large sample of students attending Italian Middle Schools will be conducted.

References

[Beck, 2017] Beck, D. (2017). Literature review: Music technology. *Education Kierstin Bible University of Arkansas ETEC 5203-Foundations of Educational Technology*.

[Caruso et al., 2020] Caruso, F., Peretti, S., Corbacchini, L., Centofanti, C., & D'Errico, A. (2020). The crazysquare project: A technological pedagogical content knowledge solution. In *CSEDU (1)* (pp. 655–662).

[Decree law no. 201, 1999] Decree law no. 201 (1999). Ministerial decree august 6, 1999, n. 201, on the subject of *Reconditioning and arrangement of experimental courses in music in middle school*.

[Denis & Jouvelot, 2005] Denis, G. & Jouvelot, P. (2005). Motivation-driven educational game design: applying best practices to music education. In *Int. Conf. on Advances in computer entertainment technology* (pp. 462–465).

[Di Mascio et al., 2016] Di Mascio, T., Gennari, R., Melonio, A., & Tarantino, L. (2016). Supporting children in mastering temporal relations of stories: The terence learning approach. *Int. J. Distance Educ. Technol.*, 14(1), 44–63.

[EU-GDPR – Art. 8, 2016] EU-GDPR – Art. 8 (2016). Eu general data protection regulation 2016/679 (GDPR), art. 8 *Condition applicable to child's consent in relation to Information Society Services*.

[FenderPlay, 2022] FenderPlay (2022). https://www.fender.com/play. Accessed: 2022-03-01.

[Gordon, 1989] Gordon, E. (1989). Audiation, music learning theory, music aptitude, and creativity. In *Suncoast music education forum on creativity*, volume 75.

[Gordon, 2007] Gordon, E. (2007). *Learning sequences in music: A contemporary music learning theory*. Gia Publications.

[GuitarTricks, 2022] GuitarTricks (2022). https://www.guitartricks.com/.

[Jaques-Dalcroze, 1921] Jaques-Dalcroze, E. (1921). *Rhythm, music and education*. GP Putnam's Sons.

[Konecki, 2014] Konecki, M. (2014). Learning to play musical instruments through dynamically generated lessons in real-time based on adaptive learning system. In *25th Central European Conf. on Information and Intelligent Systems* (pp. 124–129).

[Löchtefeld et al., 2011] Löchtefeld, M., Gehring, S., Jung, R., & Krüger, A. (2011). guitar: supporting guitar learning through mobile projection. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1447–1452).: ACM.

[Margoudi et al., 2016] Margoudi, M., Oliveira, M., & Waddell, G. (2016). Game-based learning of musical instruments: A review and recommendations. In *European Conf. on Games Based Learning* (pp. 426).

[MelodiQ, 2022] MelodiQ (2022). https://play.google.com/store/apps/details?id=com.ultimateguitar.assessment&-%20hl=en.

[Miendlarzewska & Trost, 2014] Miendlarzewska, E. A. & Trost, W. J. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Front. Neurosci.*, 7, 279.

[Nielsen, 2019] Nielsen, J. (2019). UX Design for Teenagers, 3rd edition.

[Novaxe.com, 2022] Novaxe.com (2022). Novaxe: Your platform for learning music. https://www.novaxe.com/. Accessed: 2022-03-01.

[Rocksmith, 2022] Rocksmith (2022). https://www.ubisoft.com/en-gb/game/rocksmith/plus.

[Rockway, 2022] Rockway (2022). https://www.rockway.fi.

[Seol et al., 2016] Seol, S., Shin, Y., & Lee, K. (2016). Learning guitar with an embedded system. *Contemporary Engineering Sciences*, 9(12), 553–560.

[SmartMusic, 2022] SmartMusic (2022). https://www.smartmusic.com/.

[Yousician, 2022] Yousician (2022). https://yousician.com/.

Towards an Integration of OMR Models in Pattern Improvisation

Francesco Ardan Dal Rì[1, 2] and Nicola Conci[1]

[1] *Department of Information Engineering and Computer Science - DISI, University of Trento, Italy*
[2] *Conservatory F.A. Bonporti, Trento, Italy*

Abstract

In this paper, we present a prototype for an interactive musical system, which allows the user to manage musical patterns in real-time using images of printed music in traditional Western notation. Monophonic scores are interpreted by a pre-trained Optical Music Recognition model, converted into numeric sequences, and sent to a pattern generator/synthesizer. This prototype constitutes a novel proposal, integrating deep learning and computer vision techniques into music making and promoting interactivity across disciplines.

1 Introduction

Musical composition through pattern manipulation has a long tradition, ranging from the earliest experiments in algorithmic music to the latest computer music practices [Magnusson & McLean, 2018]. In the last two decades, the emergence of live coding as a performative practice has positioned the use of musical patterns at a prominent role in musical creation. In this context, a great variety of libraries and dedicated environments have been presented, each characterized by a specific form of notation capable of controlling the evolution over time of these patterns (e.g [Kirkbride, 2016, Roberts & Kuchera-Morin, 2012, McLean & Wiggins, 2010]). Regarding the notation of melodic-rhythmic structures, however, the traditional notation - here referred to as Common Western Music Notation (CWMN) - is still particularly effective.

On the other hand, the recent advances in deep learning have allowed for the development of models able to automatically interpret the content of printed music, which have proven to be particularly effective in the case of monophonic scores (e.g [Liu et al., 2021, Castellanos et al., 2020]).

Our proposed prototype combines these two trends into a single system that uses an Optical Music Recognition (OMR) model to convert score images into patterns that can be played by a digital synthesizer. This paper is structured as follows: in Section 2 we present the relevant related work in the fields of pattern music, notations, and OMR; in Section 3 we briefly describe the implementations of the prototype; we then discuss the usage and limitations, and eventually conclude outlining possible future developments.

2 Background

In this section, we briefly introduce several relevant works and concepts from which we started to design our system.

2.1 Pattern music

In the musical context, the term *pattern* generally refers to a well-recognizable melodic-rhythmic cell, which repeats itself over time. In the last century, examples of compositions through the creation, superimposition, and permutation of patterns can be found both in the field of instrumental music (significant examples are *In C* by Terry Riley or *Clapping Music* by Steve Reich), and in that of computer music (for example in the works of Grossi [Mori, 2015] or Spiegel [Spiegel, 1981]). Although the use of musical patterns has been explored since the first experiences of algorithmic music, in the last two decades their use has become particularly prominent in the context of live coding, a performative practice in which the performer improvises music on the computer by compiling instructions in real-time in the form of code [Magnusson & McLean, 2018]. Most of the dedicated libraries and environments (e.g. TidalCycles [McLean & Wiggins, 2010], Gibber [Roberts & Kuchera-Morin, 2012], FoxDot

[Kirkbride, 2016], ect) allow the user to generate patterns that are synchronized and repeated cyclically. While these improvisations are generally performed from scratch, another popular approach involves specifying a series of patterns in advance and simply playing them in a given order during the performance. Magnusson defines this approach as "weak coding" [Magnusson, 2014a].

At this stage, our system allows for a similar compositional approach, in which the user can pre-define a palette of patterns and play them to his or her liking.

2.2 Notations

In Section 2.1, we have seen that there are several environments that allow users to manage musical patterns in real time. Each of them requires the use of a specific form of notation. Various ways of encoding music have been proposed: some systems allow the user to specify a melodic-rhythmic pattern through pairs of numerical values, respectively pitch (note or frequency) and duration (absolute or relative), as in SuperCollider [McCartney, 2002]; others implement specific *mini-notations* (e.g. TidalCycles [McLean & Wiggins, 2010] or ixi lang [Magnusson, 2011]), or make use of graphical objects or visual feedback (e.g. Betablocker or Scheme Bricks [McLean et al., 2010]). These notations are often designed to be also applicable to control parameters, effects, functions, etc. However, they can result quite difficult to interpret, especially in the case of sequences of a certain length or density [Blackwell & Collins, 2005]. Indeed, writing instructions requires a certain abstraction, that is, to formalize musical thought into a symbolic-syntactic notation capable of being understood by the software interpreter. As for the notation of twelve-tone equal tempered pitch and metrical rhythm and duration musical cells, CWMN can prove to be more effective [Magnusson, 2014b], but its use is currently overlooked.

Therefore, we have decided to integrate it into our work.

2.3 Optical Music Recognition

OMR has been defined as "a field of research that investigates how to computationally read music notation in documents" [Calvo-Zaragoza et al., 2020]. Although it has been explored for over 50 years, research in this field has advanced considerably in the last decade thanks to the adoption of deep learning techniques (e.g. [Liu et al., 2021, Baró et al., 2019, Castellanos et al., 2020]). Nowadays, various commercial softwares exist that integrate OMR for score scanning (e.g. SharpEye [1], PhotoScore [2], SmartScore [3]). Nevertheless, OMR still involves a variety of non-trivial tasks to be addressed. In the first place, the set of semantic relationships between the various musical symbols - as in the case of polyphonic music, articulation marks, or irregular rhythms - can be particularly complex. Furthermore, the general lack of large labeled datasets with balanced classes makes it difficult to deal with the great variety of symbols in CWMN [Novotnỳ & Pokornỳ, 2015]. As shown by Shatri and Fazekas, most works in the OMR field are focused on the encoding of monophonic scores, in which neural networks are able to achieve excellent performance [Shatri & Fazekas, 2020].

Since pattern music is generally composed of a series of monophonic and quite simple musical structures, we have decided to include one of these neural networks in our system.

[1] http://www.visiv.co.uk/

[2] https://www.neuratron.com/photoscore.htm

[3] https://www.musitek.com/smartscore-pro.html

3 Implementation

The system proposed consists of three main modules: an OMR model, which converts the images into semantic notation, a translator/OSC parser module, which converts the semantic notation into messages suitable for the pattern generator, eventually sending them to the server via OSC, and a synthesizer module, which plays the notes as they are sequenced (Figure 1).

In this section, we briefly describe each module.

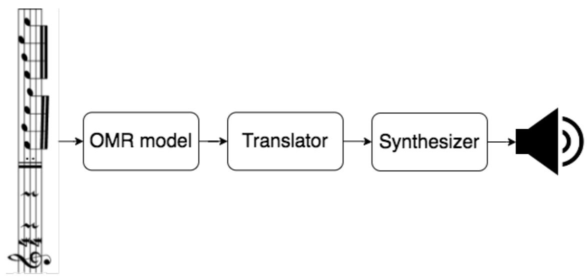


Figure 1: *Overall system pipeline*

3.1 OMR model

For this prototype, we decided to use a pre-trained, end-to-end, and open-source solution [Calvo-Zaragoza et al., 2017]. The model is trained on a series of monophonic musical incipits, namely the PrIMuS dataset [Calvo-Zaragoza & Rizo, 2018], and therefore proves to be sufficiently accurate for our needs. The model consists of a feed-forward Convolutional Recurrent Neural Network (CRNN), which extracts the individual features and converts them into discrete sequences of musical symbols. Further implementation details can be found in the original paper [Calvo-Zaragoza et al., 2017].

The model receives as input an image of a monophonic score on a single staff, and is sensitive to both clef and key signature. Consistently with the dataset used in the training phase by the authors of the original work, we consider inputs of a maximum of 4 bars in length. As output, we obtain a semantic encoding, which contains a list of all the musical symbols detected by the network in an extended format (Figure 2).

(a)

(b)

```
clef-G2, keySignature-DM, timeSignature-4/4,
rest-half, rest-eighth, note-A4_eighth,
note-B4_eighth, note-C#5_eighth, barline,
note-D5_quarter., rest-eighth, note-E5_sixteenth,
note-F#5_sixteenth, note-E5_sixteenth,
note-C#5_sixteenth, note-A4_quarter, barline
```

Figure 2: *An example of an input image (a) and the respective output in semantic notation (b)*

3.2 Translator

The translation module is responsible for converting the OMR model output into a format that the pattern generator for the synthesizer can understand, and for sending it through the Open Sound Control (OSC) protocol. The translation is performed by directly mapping a fixed vocabulary of durations and notes. The durations are considered to be relative to the quarter (e.g. quarter = 1, eighth = 0.5, etc.), while the notes are mapped to MIDI notes (e.g. C4 = 60, Db4 = 61, etc.). Rests are represented by the symbol "\" (Figure 3). We preferred to insert this translation module rather than retrain the model end-to-end for two main reasons: first, the retraining would have required the entire dataset to be completely reprocessed. Even if end-to-end retraining would be necessary for a possible definitive system, as specified in 5, at this stage we found semantic notation flexible enough to be easily translated into different types of notation.

3.3 Synthesizers

The synthesis engine, implemented in the popular open-source software SuperCollider (SC)[4], is composed of seven instruments: two additive synthesizers (characterized by a different type of envelope), two pitched single-oscillator (one octave up and one octave down, respectively), one subtractive, one FM and one producing pink noise (ideally simulating a simple drum). These instruments are recalled by the OSC messages sent by the translation module, while the related melodic-rhythmic patterns are recorded and sequenced on a quantized rhythmic grid. By default, a maximum of eight patterns can be played at the same time. In addition, the patterns are indexed: in this way, the user can activate/deactivate/overwrite specific patterns. In this first version, we have decided to keep the timbre characteristics of each instrument fixed: the user can therefore decide which instrument to play the input pattern.

[4] https://supercollider.github.io/

(b)

```
clef-G2, keySignature-DM, timeSignature-4/4,
rest-half, rest-eighth, note-A4_eighth,
note-B4_eighth, note-C#5_eighth, barline,
note-D5_quarter., rest-eighth, note-E5_sixteenth,
note-F#5_sixteenth, note-E5_sixteenth,
note-C#5_sixteenth, note-A4_quarter, barline
```

(c)

['\', '\', 57, 59, 61, 62, '\', 64, 66, 64, 61, 57]
[2, 0.5, 0.5, 0.5, 0.5, 1.5, 0.5, 0.25, 0.25, 0.25, 0.25, 1]

Figure 3: *An example of a semantic notation (b) translated into value arrays suitable for SuperCollider (c)*

4 Usage and limitations

As for the practical use of the system, it is first necessary to load a series of N images containing short musical sequences, which will be recalled with their index. These images constitute the palette of different musical sequences that can be converted into patterns. The user can input instructions consisting of four numbers, specifically representing: 1) image [1 - N]; 2) pattern index [0 - 7]; 3) instrument number [0 - 7]; 4) amplitude [0. - 1.]. The script is constantly waiting for instructions from the user, who at any time can decide to allocate a new pattern or overwrite an existing one. Note that the interpretation/translation system, on an 8GB commercial CPU, takes ~200ms to send the message to the SC server. Since the pattern generator is by default quantized to the quarter note and set to a tempo of 60 b.p.m., this latency has to be taken into account.

As for the limitations, the model used is currently not very robust in the presence of numerous ledger lines and particularly high musical densities. However, being sensitive to the keys, it is already possible to obtain a sufficient pitch range: in case the user needs more extended sounds, we have inserted two transposing instruments (see 3.3). As for the rhythmic density, we have decided not to consider durations lower than thirty-second. While we recognize this as a limitation, shorter durations are rarely used in pattern music. Finally, the available prototype does not allow direct manipulation of musical patterns or control over the parameters of the instruments.

5 Conclusions and future work

In this paper, we have presented a working prototype of a pattern music improvisation system starting from musical fragments in CWMN. We believe that the system proposed in this paper represents a contribution to developing further systems that can integrate Western musical notation in the context of interactive music production, constituting a possible creative stimulus, especially for more traditionally trained musicians, or a possible tool for music education. In addition, the prototype lends itself to various expansion possibilities, such as integration into existing systems/environments, both software and hardware, and the addition of additional interfaces/control possibilities. As a future work, the main objective will be to work on a new OMR model trained on fragments of handwritten scores in an end-to-end fashion, which also allows taking into account more complex musical symbols and relationships, in order to create an interactive system with more expressive possibilities with which the user can interact directly by writing music in real-time.

References

[Baró et al., 2019] Baró, A., Riba, P., Calvo-Zaragoza, J., & Fornés, A. (2019). From optical music recognition to handwritten music recognition: A baseline. *Pattern Recognition Letters*, 123, 1–8.

[Blackwell & Collins, 2005] Blackwell, A. F. & Collins, N. (2005). The programming language as a musical instrument. In *PPIG* (pp.$\tilde{1}$1).

[Calvo-Zaragoza & Rizo, 2018] Calvo-Zaragoza & Rizo, D. (2018). Camera-primus: Neural end-to-end optical music recognition on realistic monophonic scores. In *Ismir2018* (pp. 248–255).

[Calvo-Zaragoza et al., 2020] Calvo-Zaragoza, J., Jr, J. H., & Pacha, A. (2020). Understanding optical music recognition. *ACM Computing Surveys (CSUR)*, 53(4), 1–35.

[Calvo-Zaragoza et al., 2017] Calvo-Zaragoza, J., Valero-Mas, J. J., & Pertusa, A. (2017). End-to-end optical music recognition using neural networks. In *Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR* (pp. 23–27).

[Castellanos et al., 2020] Castellanos, F. J., Calvo-Zaragoza, J., & Inesta, J. M. (2020). A neural approach for full-page optical music recognition of mensural documents. In *ISMIR* (pp. 558–565).

[Kirkbride, 2016] Kirkbride, R. (2016). Foxdot: Live coding with python and supercollider. In *Proceedings of the International Conference on Live Interfaces* (pp. 194–198).

[Liu et al., 2021] Liu, A., Zhang, L., Mei, Y., Han, B., Cai, Z., Zhu, Z., & Xiao, J. (2021). Residual recurrent crnn for end-to-end optical music recognition on monophonic scores. In *Proceedings of the 2021 Workshop on Multi-Modal Pre-Training for Multimedia Understanding* (pp. 23–27).

[Magnusson, 2011] Magnusson, T. (2011). The ixi lang: A supercollider parasite for live coding. In *ICMC*.

[Magnusson, 2014a] Magnusson, T. (2014a). Herding cats: Observing live coding in the wild. *Computer Music Journal*, 38(1), 8–16.

[Magnusson, 2014b] Magnusson, T. (2014b). Scoring with code: Composing with algorithmic notation. *Organised Sound*, 19(3), 268–275.

[Magnusson & McLean, 2018] Magnusson, T. & McLean, A. (2018). Performing with patterns of time. In R. T. Dean (Ed.), *The Oxford Handbook of Algorithmic Music* chapter 14, (pp. 245–266). Oxford University Press.

[McCartney, 2002] McCartney, J. (2002). Rethinking the computer music language: Supercollider. *Computer Music Journal*, 26(4), 61–68.

[McLean et al., 2010] McLean, A., Griffiths, D., Collins, N., & Wiggins, G. (2010). Visualisation of live code. *Electronic Visualisation and the Arts (EVA 2010)*, (pp. 26–30).

[McLean & Wiggins, 2010] McLean, A. & Wiggins, G. (2010). Tidal–pattern language for the live coding of music. In *Proceedings of the 7th sound and music computing conference* (pp. 331–334).

[Mori, 2015] Mori, G. (2015). Pietro grossi's live coding. an early case of computer music performance. *ICLC2015 Proceedings*, (pp. 125–132).

[Novotnỳ & Pokornỳ, 2015] Novotnỳ, J. & Pokornỳ, J. (2015). Introduction to optical music recognition: Overview and practical challenges. In *DATESO* (pp. 65–76).

[Roberts & Kuchera-Morin, 2012] Roberts, C. & Kuchera-Morin, J. (2012). Gibber: Live coding audio in the browser. In *ICMC*, volume 11 (pp.$\tilde{6}$).

[Shatri & Fazekas, 2020] Shatri, E. & Fazekas, G. (2020). Optical music recognition: State of the art and major challenges. *arXiv preprint arXiv:2006.07885*.

[Spiegel, 1981] Spiegel, L. (1981). Manipulations of musical patterns. *Proceedings of the Symposium on Small Computers and the Arts*, (pp. 19–22).