Dr Ivan Zavada, Dale Keaveny

The Xenophone, an electroacoustic representation of intercultural communication trends on social media

The University of Sydney / Sydney Conservatorium of Music

ivan.zavada@sydney.edu.au / dale.keaveny@live.com.au

Abstract

It is proposed in this paper that technology such as internet-based social networks can be an effective form of virtual dynamic controller for the purpose of composition. Since the field of music technology is in itself interdisciplinary, it is only natural to integrate various other forms of technology to develop new tangible instruments. The conceptual and aesthetic value of integrating elements from other disciplines can therefore significantly alter traditional artistic approaches taken by composers to use, integrate, and modify existing technologies to forge new musical ideas. Most often these approaches include synthesis, signal processing, parametric and algorithmic control methods to generate musical/sonic material. A further goal with this project is to examine the nature of craft versus creativity focusing on the compositional experience itself, in this case, a technology driven exercise created by the new socio-cultural contexts through social media networks rather than the music technology itself to generate new musical works. As an immediate corollary, the question arises whether these new ideas stem from existing creative constructs or from new compositional methodologies influenced by external systems such as social media networks. The idea behind the Xenophone project is to integrate the interdisciplinary aspects of musical creativity in a broader sense to reveal the mechanisms of creativity through technology. Is there not a connection between global technological evolution and human creativity?

Introduction

Vast digital communication networks allow for the rapid and effective dissemination of ideas and concepts, which can significantly impact how people interact in a global context. This integrated relationship with technology has the potential to directly influence global trends in human communication, inclusive of possible interactions with musical thinking and creativity.

The principle of causation can be applied as a form of reverse engineering to explore possible existing taxonomies of sound objects for analytical purposes and further our understanding of sound based composition as a representation of human thought/activity. With that premise in mind, the aim of this research is to represent intercultural communication trends by sampling the global communication pulse at a given time and create a virtual sonic experience to generate new musical content based on contemporary human thought and activity.

To discuss this framework legibly, we've consolidated the broader implications of transformative media to illustrate their relevance to the musicological field. This helps avoid misrepresenting the deep-set sociological and psychological connotations of social-media,

reinforcing the fact that any inclusion of topics rooted in these areas are used purely for the purpose of generating music.

Research Goals

The Xenophone Music Generator is used as a case study to take advantage of readily accessible communications data pulled from social media networks to influence procedurally generated harmonies, melodies and sonic material. This is accomplished by sampling communication activity on social networking services such as Twitter for example, where users post and interact with short messages of up to 140 characters to communicate an idea or a thought. The generated music attempts to mirror the mood and intent of the sourced Twitter posts in real time, informing art through its interpretation of global discourse.

Framework

Given the prototypical nature of the Xenophone framework and application, the implemented data is limited to accessible media we propose is most malleable. The multitude of neurological patterns that make up human expression must be abstracted into broader categories to adequately disseminate legible structure. The resultant data forms can be categorized as follows:
   A. Textual
   B. Social
   C. Geographical

Textual data is the primary data analysed, providing emotional expression to be converted into sound objects. It defines user terminology and helps consolidate emotionally irrelevant text [ie 'it', 'and', 'with'].

Social data helps provide context to the textual expression, narrowing the possible definitions of a given phrase so as to uncover user perspective. This data can also link perspectives and social trends to help solidify generated sonic motifs.

Geographical data reinforces both textual and social data, providing evidence of the physical space social discourse is rooted within. It brings tangible distance to the events and concepts users discuss, providing the means to deduce historical context to inform instrumentation and timbre.

Data categorized in this format can be used to generate sufficient expression to be translated into a sonic medium.

Textual data

Discussing text representation of the emotional spectrum is a key factor in understanding how sonic media can be generated using data from other areas, as it is the fundamental means through which our application processes multidisciplinary information.

```
Tree('S', [('At', 'IN'), ('eight', 'CD'), ("o'clock", 'JJ'),
          ('on', 'IN'), ('Thursday', 'NNP'), ('morning', 'NN'),
       Tree('PERSON', [('Arthur', 'NNP')]),
          ('did', 'VBD'), ("n't", 'RB'), ('feel', 'VB'),
          ('very', 'RB'), ('good', 'JJ'), ('.', '.')])
```

Figure 1: Natural Language Processing categorization

The framework utilizes the grammar and syntax of the twitter posts (figure 1), analysing embellishments such as capitalized adjectives or expletives to inform musical parameters. This will create more consonant or dissonant musical structures based on the syntactical/semiological analysis.
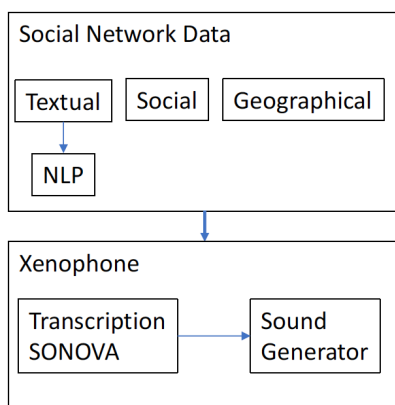
Figure 2: Xenophone framework

Natural Language Processing (NLP) has proven itself as an integral means of dynamically categorizing terabytes of text data common within social media conglomerates. The means of digitally interpreting the fuzzy logic of human communication to a practical degree is a valuable asset in forming an unbiased account of human emotional activity, regardless of the medium. The Xenophone uses NLP at the core of its analysis, breaking down text objects into expressions with the purpose of categorizing them by emotional resonance (figure 2). Tonality is achieved by assigning a text object to a particular realm of modality coupled with mood, providing enough leeway to generate a wide array of representative melodico-rhythmic and sonic content. Our primary concern is the recognition/categorisation of the expressions themselves, their sonification and transformation as well as their symbolic interpretation.

This process summarizes the conversion mechanism upon which our framework is dependent; serving as a technical metaphor we define as multidisciplinary translation. When discussing the relevance of text objects to the human emotional spectrum, we reinforce the concept that any converted expression, once represented in a digital format can be repurposed as direct representation in other forms of media. That being acknowledged, this section of the paper will illustrate how emotional data is disseminated into a sonic format.

Social data

The inherent democracy at the heart of social media – the commodity of interest informing relevance to the community at large provides enough real time data to the Xenophone with relatively little direct user input. At any given time, active conversations about contextually relevant events are pushed to the forefront of social media labelled as 'trending' topics, providing unprecedented access to individual reactions to history on a mass scale. From a

musicological standpoint, pairing directly representative music with such access would provide future historians with the means to illustrate emotional dissemination of relevant subjects with startling accuracy.

> I am sure that the time will come when the composer, after he has graphically realized his score, will see this score automatically put on a machine which will faithfully transmit the musical content to the listener. As frequencies and new rhythms will have to be indicated on the score, our actual notation will be inadequate. The new notation will probably be seismographic. And here it is curious to note that at the beginning of two eras, the Mediaeval primitive and our own primitive era (for we are at a new primitive stage in music today) we are faced with an identical problem: the problem of finding graphic symbols for the transposition of the composer's thought into sound.[1]

From this perspective, the Xenophone acts as a sort of seismograph measuring tremors in social discourse, giving it the potential to form culturally significant artworks to be reintegrated into popular events. Musical accounts of historical events are generally composed in reaction to the primary occurrence or are popular resonant compositions at the time of the event. Either way there are multiple forms of abstraction indicative of individual perspective. This range of expression can be benefited by the inclusion of scalable communal composition, enabling the 'individual' to incorporate global perspective in addition to personal bias. This may provide insight into the propagation of creative ideas as they stem from relevant points in history.

The sphere of online interaction has become a gateway to categorizing emotional expression. The paths through which the contemporary digital participant can link their emotional viewpoint to a wider context, be it an event, phenomenon, concept or pre-existing information (i.e. a text, a musical score, etc.), an interdisciplinary advantage to research interactivity within social media.

The digitization of text media and word processing has already shown digital understanding of human syntax, providing the literature community with real-time error detection, an action achieved only by building a link between the digital community and the literature community. Both computational and linguistic syntax utilize logical rule sets as user guidelines and consider context as a staple of problem solving. Researchers in both fields regard the need for an external analysis to help in resolving issues such as syntactical errors and data recognition. While this intent is a practical one there is room to build a connection between contextual analysis and intent, with the eventual goal of building a wider study of digital communication.

Geographical data

Musical parameters such as instrumentation and spectral qualities of sounds will also be influenced by regions linked to the discussion; therefore providing geo-cultural variation following the evolution of online discussions. Users of differing background and political association discussing a certain topic would therefore influence the system to produce a musical mood based on their opinion of that topic. In essence this will provide a new platform for the musical representation of meaning that can be used in conjunction with existing spectromorphological analysis methods to link compositional musical gestures with human behaviors and thoughts.

---

1 Edgard Varèse, "The Liberation of Sound" [excerpts from lectures by Varèse, compiled and edited with footnotes by Chou Wen-Chung], *Perspectives of New Music*, 5(1), 1966, p. 12, http://music.arts.uci.edu/dobrian/CMC2009/Liberation.pdf (last accessed 01/18).

Most importantly, the musical content emerging from the data sonification itself may provide some leads towards a better understanding of electroacoustic music analysis and spectromorphological patterns of construction within larger forms of musical expression.

This prototype builds upon the idea of using spectromorphological analysis vocabulary for the choice of sound material and creation in compositional activity[2]. The proposed system also reorients the concept of 'searched objects'[3] for re-composition/re-instrumentation purposes within the realm of 'searched meaning' and symbolic association through sound, while integrating the unpredictable and perhaps reactive nature of network media interaction. The main advantage of the system as a compositional tool is that it does not depend on a fixed medium compositional paradigm, and as the information database evolves over time, it will develop into a more innovative platform to generate new compositional material based on current communication trends in social media, with further applications in other fields such as real-time interaction/sonification for online gaming platforms.

The data shared by these media interactions can be recycled for the purposes of generating a visual feed representing the Xenophone analytics. The resulting assets can be expanded to formulate a complete audiovisual representation of global discussion fit for stage production or live installations, as it is intended to work autonomously after a single hashtag search. The resulting symbolic expression provides a unique view of cultural trends with the added complexities of multi-parametric musical interpretations. Twitter's access to vast archives of digitized human interaction can provide the grounds for in depth artistic representation of the global community. The framework therefore intends to simplify the monolithic bank of user data integral to the function of modern social media, and replicate the simplified data in the form of a schematized musical mood profile. While its function here is linked specifically to Twitter, the concept behind its construction is fluid and adaptable to other sources of data and human communication and leads toward a richer and wider capacity for interactive communication and expansion of the musical realm of human expression.

From symbols to metaphors: semiological implications

As social media has become widespread, emoticons have played a significant role in communication through technology. They offer another range of 'tone' and feeling through text. We theorize that in the process of bridging linguistic and musical expression, artists can materialise intent through the transcription process more clearly. By emphasising the emotional core of a potential music composition through contextual wording, which can be more directly translated into neighboring forms of dictation, a composer may find a wider range of tools to express experience through the organization of sound. The transcribed sonic result becomes a concrete metaphor through which the original story intent is transformed and recontextualised, in a new stream of sonic memes through time. This puts us in a good position to make a link with the following premise Wishart:

> Using concrete metaphors, (rather than text), we are not 'telling a story' in the usual sense, but unfolding structures and relationships in time; ideally we should not think of the two aspects of

---

[2] Manuela Blackburn, "The Visual Sound-Shapes of Spectromorphology: an illustrative guide to composition", *Organised Sound*, 16(01), 2011, p. 12.

[3] Ricardo Climent, "Applications of Typomorphology in *Acute*; Scoring the Ideal and its Mirror", in *Proceedings of SMC08 – 5th Sound and Music Computing Conference*, Berlin, 2008, http://smc.afim-asso.org/smc08/images/proceedings/session5_number4_paper19.pdf (last accessed 01/18).

the sound landscape – the sonic and the metaphorical – as different, but as complementary aspects of the unfolding structure.[4]

As a corollary idea, if we are to describe the transformational process from a musical score to the performance, the score indicates no tangible depiction of mood, intent and meaning. In fact, the music may be imagined performatively in the person's mind[5]. We are provided with musical information and performance instructions indicating the order through which each musical event is added, yet the purpose of the organization is not expressed until the actual performance, where human perception may alter some musical characteristics. At that point during the process, artistic interpretation may contain performance variations such as nuances in tempo, phrasing, character and dynamics. As a result, the implied metaphorical meaning can drastically change in character. "Representations of musical abstractions have been a double-edged sword: a representation allows us to create, disseminate, and preserve a musical abstraction; yet, these representations may lead to interpretations that are inconsistent with the composer's intent […]."[6]

Within the realm of traditional notation, the four elements of music are purely structural, propagated around the ideal of replicating performance rather than artistic intent.

> Since the eclipse of the oral tradition with the advent of the scribes, we have struggled to bridge the abyss of lost meaning between representation and intent. Just as the written word alone does not exclusively impart meaning, musical representation does not solely communicate compositional intent.[7]

This perhaps stems from processes intrinsically linked to music composition, with intangible qualities like inspiration and philosophical direction taking precedence over notational expression, instrumentation, orchestration, sound design and transformation. As the sound of acousmatic music is abstracted from causality, meaning is also abstracted from textual information through sonic interpretation/translation, whilst keeping the metaphorical relationship through dynamic categorization. The complex means through which a composer communicates ethereal qualities into the tangible world serves as an obstacle in accurately transcribing/receiving music. Spectromorphology, therefore, has the ability to bridge the gap between metaphorical substance and the organization of sound objects.

Through the presentation of text and its social context, it is also possible to become interested by the figure/intent of the author's text where one can 'guess' what is not explicitly implied by the text itself, but integrated in the literary code that allows for a wider interpretation of implicit meaning (Barthes).

Here we refer to Barthe's five codes which are interwoven into the meaning of a text/narrative:
1. Hermeneutic/Enigma
2. Proairetic/Action
3. Semantic
4. Symbolic
5. Referential/Cultural

---

[4] Trevor Wishart, "Sounds Symbols and landscapes", in *The language of electroacoustic music*, Simon Emmerson (ed), London, Macmillan Press, 1986, pp. 55.

[5] Margaret Kartomi, "Concepts, Terminology and Methodology in Music Performativity Research", *Musicology Australia*, 36(2), 2014, pp. 189-208.

[6] Mary Simoni (ed), *Analytical methods of electroacoustic music*, New York, Routledge, 2006, p. 5.

[7] *Ibid.*, p. 1.

Where Hermeneutic code refers to mystery within a text, and clues are provided with no clear answers, Proairetic provides sequential elements of action and suspense, Semantic refers to the additional layer of meaning within the text, Symbolic adds elements of contrast and emphasis to the actual meaning and Referential refers to external body of knowledge (scientific, historical, political, cultural). In extension to implicit code, text cannot be reduced to a pure function of communication, and a very interesting parallel can be made with Kristeva's definitions of the distinction between semiotic and symbolic, in other words *genotext* and *phenotext*, "designating the genotext in a text requires pointing out the transfers of drive energy that can be detected in phonematic devices (such as the accumulation and repetition of phonemes or rhyme) and melodic devices (such as intonation or rhythm) [...]."[8]

The genotext is therefore not a linguistic device; Kristeva describes this as a 'process' which tends to articulate structures that are ephemeral. In addition to the tendency for reiteration, repetition of drive factors (i.e. which the author refers to drive charges) within social constraints that allow for the stabilization of their resulting surrounding structures.

> The phenotext on the other hand is constantly split up and divided, and is irreducible to the semiotic process that works through the genotext. The phenotext is a structure (which can be generated, in generative grammar's sense); it obeys rules of communication and presupposes a subject of enunciation and an addressee.[9]

If these two terms – genotext and phenotext – could be translated into a metalanguage that would convey the difference between them, one might say that the genotext is a matter of topology, whereas the phenotext is one of algebra. This distinction may be illustrated by a particular signifying system.

In his analysis of Barthe's preoccupation of the distinct roles of performer and listener, Michael Szekely proposes

> "How do we, as listeners, become players of the music?" Collective improvisation comes to mind here, not so much as a preference concerning a particular music genre, but rather because of the dynamics it engenders, for both performer and listener. The ideal, at least, in this kind of music emphasizes active, spontaneous engagement, the shifting of attention, constructive meaning-creation and the blurring of individual and collective roles.[10]

The social network communicative platform becomes the new stage for both the production and performance of a meta-musical discourse where both genotext and phenotext interact to form a symbolic and metaphoric representation of the developing and transformative social interaction.

In the Xenophone system, the transposition is not a literal process but rather an interpretive system allowing some flexibility to create sonic works that unfold through time, according to characteristics set by the mood of the social media networks and therefor act as musical/sonic signatures/emojis presenting a new perspective on the current intercultural trends. The resulting sequence of symbolic representations can provide information based on recurring patterns within a different medium, that of spectral and dynamically evolving musical structures. The framework therefore extends the Barthesian notion of 'grain of the voice' as

---

[8] Julia Kristeva, *The Kristeva Reader*, Toril Moi (ed), New York, Columbia University Press, 1986, p. 120.

[9] *Ibid.*, p. 121.

[10] Michael Szekely, "Gesture, Pulsion, Grain: Barthes' Musical Semiology", *Contemporary Aesthetics*, 4, 2006, https://quod.lib.umich.edu/c/ca/7523862.0004.005/--gesture-pulsion-grain-barthes-musical-semiology?rgn=main;view=fulltext (last accessed 01/18).

the place where a language finds a voice, assuming a dual posture: to be language and to be music[11].

Music Generation

Spectromorphological categorization utilizes symbolic depiction of sonic expression to generate stronger links between artistic and cognitive intent. While other forms of categorizing human expression are also effective within their individual focus, we propose that these areas of study have developed in parallel and are yet to be linked. We have utilized modern data collection methodology typically associated with linguistic and text based categorization to breach the gap between spectromorphological analysis and linguistic analysis – Natural Language Processing, effectively introducing a valid link between literal expression and compositional intent.
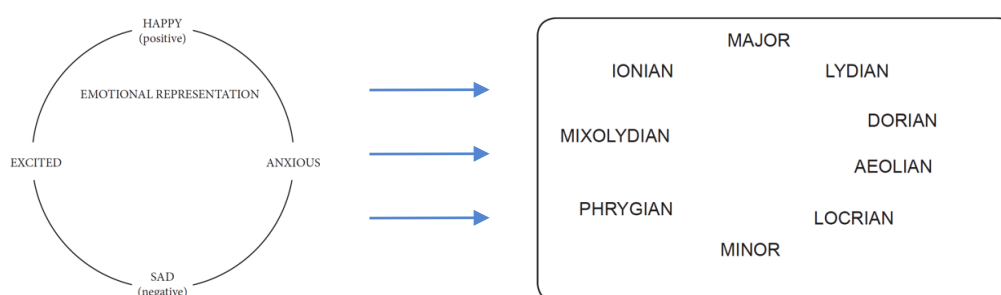

Figure 3: Combination of emotion with musical modes

We programmed the Xenophone trying to include the least amount of creative bias possible, attempting to represent emotion through musical modes rather than composed melodies in the traditional sense (figure 3). A wide spectrum of harmonic/inharmonic content can be generated through the combination of contextual modes, with effects providing grounds for microtonal expression informed by the same data (i.e. Twitter post, textual expression). We are trying to make the system as autonomous as possible for the prototype, however, the links between modal representation and the aforementioned emotional representations are currently arbitrary and therefore not proven.

> Technologies, however, are not only sound and image transmitters, but they also collaborate in the construction of new ways of listening, looking, feeling and knowing the Other, changing sensibilities, modes of perception and ways of knowing and listening to the world and life[12].

---

[11] Simone Luci Pereira, "About the possibility to listen to the Other: voice, world music, interculturality", *E-Compos* [on line], Brazil, 2012, p. 10, http://www.compos.org.br/seer/index.php/e-compos/article/viewFile/791/590 (last accessed 01/18).
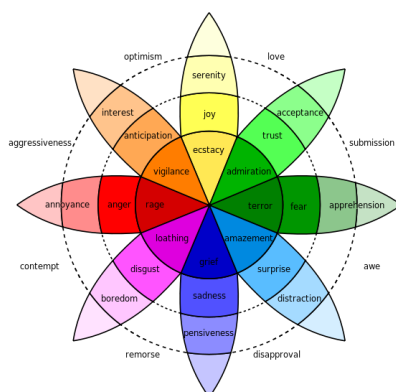[12] *Ibid.*, p. 4.

Figure 4: Plutchik's wheel of emotions

The assignment of emotional context (figure 4) to each chosen mode is the area of this framework's development least likely to exclude human bias, as the placement of modes on the emotional spectrum impacts the likelihood of particular melodies being generated by more common online textual expressions. If, for example, the Dorian mode was assigned to wording more associated with happiness rather than the Lydian mode, status updates accounting for over 25% of Xenophone generated melodies would sound completely different.

Figure 5: Xenophone framework implementing the Sonova font created by Andreas Hedman based on the work of Lasse Thoresen and Peter Tornquist

On the other hand, the rotation of the emotional representation wheel/clock potentially exposes relative distance or intervallic relativity between uncategorized emotional states/objects. We are therefore defining a parallel between emotional states and sound objects based on their spectromorphological qualities; essentially creating a new synesthetic relationship between two different forms of human expression. With this in mind, possible future iterations of this framework could use neural networking to deduce modal association with more accuracy over time, further removing creative bias from its internal functions. This framework is dependent on constant streams of real time social data analytics, used in a modern context to predict, inform and control the flow of information shared between users of any given social network.

Conclusion

We have shown in this paper, through the elaboration of a social media based music generator that spectromorphological representation can be an effective tool for the dynamic categorization of sound gestures and musical expression. It was also found that musical analysis can take a more metaphorical direction if coupled with ontologies such as natural language processing libraries. Sound based musical practice, from sonic art in the early

twentieth century, through to concurrent trends, practices and aesthetics in sound creation and reproduction today, has yet to embody a common representation system/toolset to clearly designate spectromorphological characteristics for compositional purposes.

We feel that exploring commonalities between other analytical systems may reinforce the value of existing representation platforms such as Sonova (figure 5), with the intent of synthesizing meaning and perception into the sonic equation. Vertical consolidation between different platforms and disciplines has the potential to enhance sonic and metaphorical articulation, in relation to creating form and structure over time as well as generating and recognizing micro-structures and sound objects with both analytical and creative purposes. Most importantly, the sonic medium remains an important and undeniable foundation facilitating human communication in/through electroacoustic music. We find ourselves at the crossroads of interdisciplinary human activity, through the multifunctional aspects of musical composition, where traditional musical paradigms can indeed co-exist with sound based art practice to convey, interpret and ultimately represent and codify human thought, expression and activity through sound. This experimental approach may provide insight on the foundations of creative activity through analytical strategies to enrich our knowledge of the musical phenomenon, its transformation and dissemination.

References

BARTHES Roland, *Le Plaisir du texte*, Paris, Éditions du Seuil, 1973.

BARTHES Roland, "Masson's Semiography", in *Responsibility of Forms: Critical Essays on Music, Art and representation*, translated from the French by Richard Howard, Berkeley and Los Angeles (CA), University of California Press, 1985, pp. 154-156.

BARTHES Roland, "The 5 codes", *SlideShare* [on line], https://www.slideshare.net/sarahlambe/barthes-5-narrative-codes-12063567/3 (last accessed 01/18).

BLACKBURN Manuela, "The Visual Sound-Shapes of Spectromorphology: an illustrative guide to composition", *Organised Sound*, 16(01), 2011, pp. 5-13.

CLIMENT Ricardo, "Applications of Typomorphology in *Acute*; Scoring the Ideal and its Mirror", in *Proceedings of SMC08 – 5th Sound and Music Computing Conference*, Berlin, 2008, http://smc.afim-asso.org/smc08/images/proceedings/session5_number4_paper19.pdf (last accessed 01/18).

WISHART Trevor, "Sounds Symbols and landscapes", in *The language of electroacoustic music*, Simon Emmerson (ed), London, Macmillan Press, 1986, pp. 41-60.

KARTOMI Margaret, "Concepts, Terminology and Methodology in Music Performativity Research", *Musicology Australia*, 36(2), 2014, pp. 189-208.

KRISTEVA Julia, *The Kristeva Reader*, Toril Moi (ed), New York, Columbia University Press, 1986.

PEREIRA Simone Luci, "About the possibility to listen to the Other: voice, world music, interculturality", *E-Compos* [on line], Brazil, 2012, http://www.compos.org.br/seer/index.php/e-compos/article/viewFile/791/590 (last accessed 01/18).

SIMONI Mary (ed), *Analytical methods of electroacoustic music*, New York, Routledge, 2006.

THORESEN Lasse and Andreas HEDMAN, "Spectromorphogical analysis of sound objects: an adaptation of Pierre Schaeffer's typomorphology", *Organised Sound*, 12(02), 2007, pp. 129-141.

VARÈSE Edgar, "The Liberation of Sound" [excerpts from lectures by Varèse, compiled and edited with footnotes by Chou Wen-Chung], *Perspectives of New Music*, 5(1), 1966, pp. 11-19, http://music.arts.uci.edu/dobrian/CMC2009/Liberation.pdf (last accessed 01/18).

SZEKELY Michael David, "Gesture, Pulsion, Grain: Barthes' Musical Semiology", *Contemporary Aesthetics* [on line], 4, 2006, https://quod.lib.umich.edu/c/ca/7523862.0004.005/--gesture-pulsion-grain-barthes-musical-semiology?rgn=main;view=fulltext (last accessed 01/18).