

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

KYIV NATIONAL LINGUISTIC UNIVERSITY

**Professor O.M. Morokhovsky Chair of English Philology, Translation and
Philosophy of Language**

Master's diploma paper

Academic vocational programme Modern philological studies (English and another
foreign language): linguistics and translation studies

Speciality 035 Philology, specialisation 035.04 Germanic languages and literatures
(translation included), English major.

**STYLISTICS OF 'IMAGINE DRAGONS' POP ROCK BAND'S
DISCOURSE: A MULTIMODAL PERSPECTIVE**

KHRYSTYNA BLASHCHUK

Group MLa 51-18

Department of Germanic Philology

Research Adviser

OLENA S. MARINA

Ass. Prof., DrSc. (Philology)

Kyiv 2019

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ КИЇВСЬКИЙ
НАЦІОНАЛЬНИЙ ЛІНГВІСТИЧНИЙ УНІВЕРСИТЕТ**

**Кафедра англійської філології, перекладу і філософії мови імені
професора О.М. Мороховського**

Дипломна робота магістра

за освітньо-професійною програмою Сучасні філологічні студії (англійська мова і друга іноземна мова): лінгвістика та перекладознавство
зі спеціальності 035 Філологія спеціалізації 035.04 Германські мови і літератури (переклад включно), перша – англійська

**на тему: «ЛІНГВОСТИЛІСТИКА ДИСКУРСУ ПОП-РОК ГРУПИ
“IMAGINE DRAGONS” У МУЛЬТИМОДАЛЬНОМУ ВИМІРІ»**

Допущено до захисту

«__» _____ року

Студентку групи МЛа 51-18

факультету германської філології

денної форми навчання

Блащук Христину Любомирівну

(ПІБ студента)

Завідувачка кафедри

англійської філології,

перекладу і

філософії мови імені

професора О.М. Мороховського

_____ *Воробйова О.П.*

(підпис)

(під)

Науковий керівник:

доц., д.ф.н. Маріна Олена Сергіївна

(науковий ступінь, вчене звання, ПІБ)

Національна шкала _____

Кількість балів _____

Оцінка ЄКТС _____

TABLE OF CONTENTS

| | |
|---|-----------|
| INTRODUCTION | 4 |
| CHAPTER ONE. THEORETICAL FRAMEWORK FOR THE RESEARCH OF MUSICAL DISCOURSE FROM THE STANDPOINT OF MULTIMODALITY..... | 8 |
| 1.1. Multimodality as a phenomenon | 8 |
| 1.2. Multimodal analysis: Opportunities and challenges | 14 |
| Conclusions to Chapter One | 15 |
| CHAPTER TWO. MUSICAL DISCOURSE: ANALYTICAL TOOLS..... | 17 |
| 2.1. Structure of musical discourse | 17 |
| 2.2. Music videos as fragments of musical discourse | 22 |
| 2.3. Analysis of music and lyrics in musical discourse..... | 28 |
| 2.4 Analysis of visual elements in musical discourse | 32 |
| Conclusions to Chapter Two | 34 |
| CHAPTER THREE. MULTIMODAL ANALYSIS OF 'IMAGINE DRAGONS' MUSIC VIDEOS..... | 36 |
| 3.1 “Believer” (2017): structure and interaction of modes | 36 |
| 3.2 “Radioactive” (2012): Apocalyptic motives and allegory | 41 |
| 3.3. Visual symbolism in “Shots” (2015)..... | 45 |
| 3.4. Pictorial and conceptual metaphors in “I bet my life” (2014)..... | 50 |
| 3.5. Hidden movies in “On top of the world” (2013)..... | 55 |
| Conclusions to Chapter Three | 59 |
| GENERAL CONCLUSIONS | 62 |
| RÉSUMÉ | 65 |
| REFERENCES | 67 |

INTRODUCTION

In present-day life multimodality is associated not only with the history and interpretation of written texts, but also with different media aspects. It is a complex notion embracing multiplicity of modes of human perception. At that, multimodal tools are widely applied to video and audio analysis to decode senses generated by multimodal forms. This approach is at work in music stylistic analysis, especially when it comes to popular music videos.

Multimodal music discourse of *Imagine Dragons* popular pop-rock band is in the focus of this paper. In particular, we analyze the band's song lyrics and music videos. There is an opinion that it is not solely a powerful message in song lyrics that prompts *Imagine Dragons*' success, but it is also a coherent idea expressed in lyrics, sound, and video.

Furthermore, a great emphasis is put not only on identifying various modes in music videos and explaining their interplay, but also on considering their contribution to the general idea of the video, as well as to its message.

Multimodal perspective, being relatively new in linguistics and stylistics in particular, presents a set of problematic questions being analyzed by scholars of the past few decades (Halliday, 1978; Baldry and O'Halloran, 2004; Jewitt, 2009; Machin, 2007; Kress and van Leeuwen, 2006; O'Toole, 2011 [1994]; Forceville and Urios-Aparisi, 2009; Currie, 2014). Musical discourse and its structure was studied by Blacking (1982), Scollon (2004), Blommaert (2005), Gumperz (2001), Gillet, Essid, and Richard (2007) Poh & Kittler (2010).

There is also a wide range of applications and software that are of a great use in collecting and interpreting multimodal data. They (MSAF (Music Structure Analysis Framework), MIR (music information retrieval), libROSA) facilitate music structure analysis by means of combining different computation features, algorithms, datasets, and metrics.

The topicality of this paper is justified by the interest in multimodal studies and their application to stylistic musical discourse. *Imagine Dragons* band's discourse is a great field for multimodal research as the group is famous and

recognizable all over the world. Their songs and music videos are on lips of millions around the globe. Therefore, the research of pop rock band`s discourse from a multimodal perspective is a great example of how the modes interact and what effect this interaction produces upon the audience.

Thus, it is significant to analyze the band`s stylistics and provide a distinct as well as structured scheme of modes interaction in their videos. Diversity of forms, techniques, and methods actively used in multimodality presupposes the application of a significant number of approaches in order to integrate them into a general structure of present-day research.

The object of the paper is stylistics of *Imagine Dragons* pop rock band`s discourse.

The subject-matter of the thesis is multimodal specificity of stylistics of *Imagine Dragons* pop rock band`s discourse. The paper also aims to reveal analytical frameworks of musical discourse, as well as the most widely spread techniques and principles of music video analysis.

The aim of the paper is to determine multimodal features of *Imagine Dragons`* pop rock band`s discourse.

The aim is interconnected with the following **tasks** to resolve:

- ✓ to define the notion of multimodality for the research of musical discourse;
- ✓ to discover opportunities and challenges of multimodal analysis;
- ✓ to reveal the structure and elements of musical discourse;
- ✓ to outline music videos as fragments of musical discourse;
- ✓ to analyze music and lyrics in *Imagine Dragons'* musical discourse;
- ✓ to examine visual elements in *Imagine Dragons'* musical discourse;
- ✓ to elaborate on music structure and interaction of modes in “*Believer*” (2017);
- ✓ to outline apocalyptic motives and allegory in the music video “*Radioactive*” (2012);
- ✓ to determine visual symbolism in “*Shots*” (2015);
- ✓ to identify pictorial and conceptual metaphors in “*I bet my life*” (2014);

✓ to infer hidden movies in “*On top of the world*” (2013).

The methodology used is based on the following **methods**:

- the method of linguistic interpretation to determine functions of stylistic means in a musical video;
- multimodal stylistic analysis to examine a structure of the musical video and defining the interaction of modes in it;
- the method of music discourse analysis to determine interaction between discourse and occurrence of syntactic structures in all aspects;
- ELAN analysis.

Theoretical value of the paper lies in defining the structure of musical discourse according as multimodal, as well as in characterizing the interaction of all modes in music videos, which constitute *Imagine Dragons*` musical discourse and outline its stylistics.

Practical value of the paper lies in possibility of applying multimodal tools and music analysis frameworks to *Imagine Dragons*` music videos. The research results can be used while studying multimodal stylistics, as well as for further research in course and diploma papers.

Structurally, the master thesis consists of Introduction, three Chapters, Conclusions to chapters, General Conclusions, and References.

The first Chapter ‘**Theoretical Framework for the Research of Musical Discourse from the standpoint of multimodality**’ defines multimodality as a phenomenon in musical discourse. The Chapter discusses the opportunities and challenges of multimodal analysis.

The second Chapter ‘**Musical Discourse: Analytical tools**’ elaborates on a structure of musical discourse. The Chapter defines music videos as fragments of musical discourse. This Chapter also discovers the role of music and lyrics, as well as visual elements in musical discourse.

The third Chapter ‘**A Multimodal Analysis of ‘*Imagine Dragons*’ Music Videos**’ analyzes “*Believer*” (2017), “*Radioactive*” (2012), “*Shots*” (2015), “*I bet my life*” (2014), “*On top of the world*” (2013).

General Conclusions present a consistent summary of the thesis with an outline of main points analyzed.

CHAPTER ONE. THEORETICAL FRAMEWORK FOR THE RESEARCH OF MUSICAL DISCOURSE FROM THE STANDPOINT OF MULTIMODALITY

1.1. Multimodality as a phenomenon

In present-day stylistic research, a multimodal perspective appears to be one of the most efficient. Multimodal stylistics aims to enlarge a number of perception modes to which stylistic analysis can be applied. In other words, a toolkit of multimodal stylistics can be used to analyze not solely a printed text, but also pictures, music, moving images, types, color, etc., which also participate in meaning-making (Gibbons, 2010; Nørgaard, 2010).

Within multimodal studies, the phenomenon of multimodality lies on four milestones:

- communication requires the interaction of different modes, hence it is multimodal;
- analysis of any type of text loses its adequacy in case it focuses on verbal mode only since it cannot convey meaning fully;
- each mode of perception has its features that tend to meet certain communication needs;
- every mode of communication contributes greatly to meaning making in the text (García, Flores, & Spotti, 2016).

At that, multimodal relations are the clue to understanding communication. Multimodality facilitates a new approach to understanding and interpreting of the printed works since great emphasis is put not only on verbal mode, but also on all aspects of material realization of, for instance, a book, such as book cover, paper quality, font of the letters, pagination and, of course, smell (Simpson & Montgomery, 1995; McIntyre, 2008; Montoro, 2010).

Multimodal studies are associated with the semiotic theory. Halliday's (1978) theory of social semiotics sets the stage for further studies of semiotics as well as reviews of semiotic interaction, which is also called multimodal analysis (e.g.,

Jewitt, 2009; Machin, 2007; O'Halloran, 2011). In his research, Halliday points out that semiotic theory is applicable to various texts demonstrates strengths in that it is applicable to a wide range of texts. This fact enlarges the scope of semiotics and despite limitations on grammar and function of the text, so there is a possibility to analyze not only isolated sentences, but also to consider grammatical structures in a particular context. However, a high concentration on the structure of the text prevents the multimodal researcher from visual and aural analysis that usually go alongside with any text.

Although his approach grounds on semiotic theory, it contributes to contemporary multimodal studies and serves as a landmark for numerous works in this sphere. Stylistician encourages the opinion that the theory underpins the design (or `grammar`) of relations of semiotic resources as well as the funds themselves. Moreover, they are to fulfill four functions: to interpret our experience of the world (experiential), to build logical relations between an experience people acquire during their life (consistent), to present social relationships (interpersonal), and organize meanings into messages (textual). Hence, there appear multiple points important for the successful creation of meaning.

In *Multimodal Discourse: The Modes and Media of Contemporary Communication* (2001), Kress and van Leeuwen take the concept of mode from Halliday's (1978) distinction between writing and speech and expand it to all resources of representation. They exemplify their findings by texts, where writing combines with images, but they abandoned widely used in linguistics frames and terminology. Kress and van Leeuwen shift away from the assumption that every mode in the text has an individual task that contributes to the general structure, that creates meaning. Instead, they underpin universal semiotic principles that might occur in different modes. These concepts play a crucial role in decoding modes since with their help once can interpret music through action or images with the help of emotions.

What is more, van Leeuwen (2001) claims that individuals express meaning choosing the semiotic resources that at hand at a particular moment. Supporting

this approach, Kress (2001) introduces the notion of unity that encompasses all the signals that are significant for the meaning-making process. As the meaning-making process includes several signs and they contribute to the message addressees receive, this process should be analyzed from a multimodal perspective. Only after considering all the semiotic resources in particular content, the grounded conclusions can be made.

For instance, Baldry and Thibault (2006) have suggested a slightly new approach to transcription and analysis of multimodal texts shifting from a traditional concept of multimodality as "there are many other resources that can be used to create texts in addition to the spoken and written word." Their view is based on multimodal concordances capable of monitoring several items (e.g., sound, movement, color, or verbal input) interfering in different semiotic modes which can simultaneously be at work in the formats encompassed by the notion of the moving image. That is sonic, musical, pictorial or linguistic mode.

Stylisticians develop the various dimensions of language, aiming to demonstrate that this approach would be applicable to other sign systems as well. They state that multimodal analysis allow to conduct broader film, advert or verbal text analysis.

However, Baldry and Thibault's basic approach lies in underpinning stylistically marked elements from different modes that might affect the interpretation of the text. Sometimes, due to various reasons, this way of construing images might lead to mis- or overinterpretation of the given text.

Relatively new facets are at work in technology-based approaches to the multimodal moving pictures delineated by Baldry and O'Halloran (2004). Potential of their research is quite high due to a wide accessibility of technology among scholars as well as the community. Baldry developed Multimodal Corpus authoring (MCA) – a system operating as a multimodal concordancer based on XML and aimed to identify repetitive elements in films. O'Halloran's software is available as well since it bases on Adobe Premiere 6.0, which is also accessible online. It enables transparent matters that overlay on the original footage to that

extent that the text can be inserted, and different vectors and lines can be put on the footage as well.

This approach allows pointing out the semiotic choices and applies them to the image analysis alongside the data. Despite the fact the analysis is computer-based, it is at hand at collecting data and at distinction between the modes. The videos can be split on different components that can be considered and analyzed separately.

To some extent this feature limits the possibilities of multimodal analysis by splitting the video as the dimensions does not interfere and their perception is rather monomodal. At that, the software is of a great help for creating musical videos, but it requires a lot of skills and effort in case the video is to be analyzed.

One more software (ELAN) is a semi-automatic multimodal video analysis software (Twilhaar & Bogaerde, 2016). Despite the software was developed in 2000 by The Language Archive, it updates very often. The updates unable ELAN to meet the current trends and demands. As the tool is semi-automatic, it requires some skills on how to use the application.

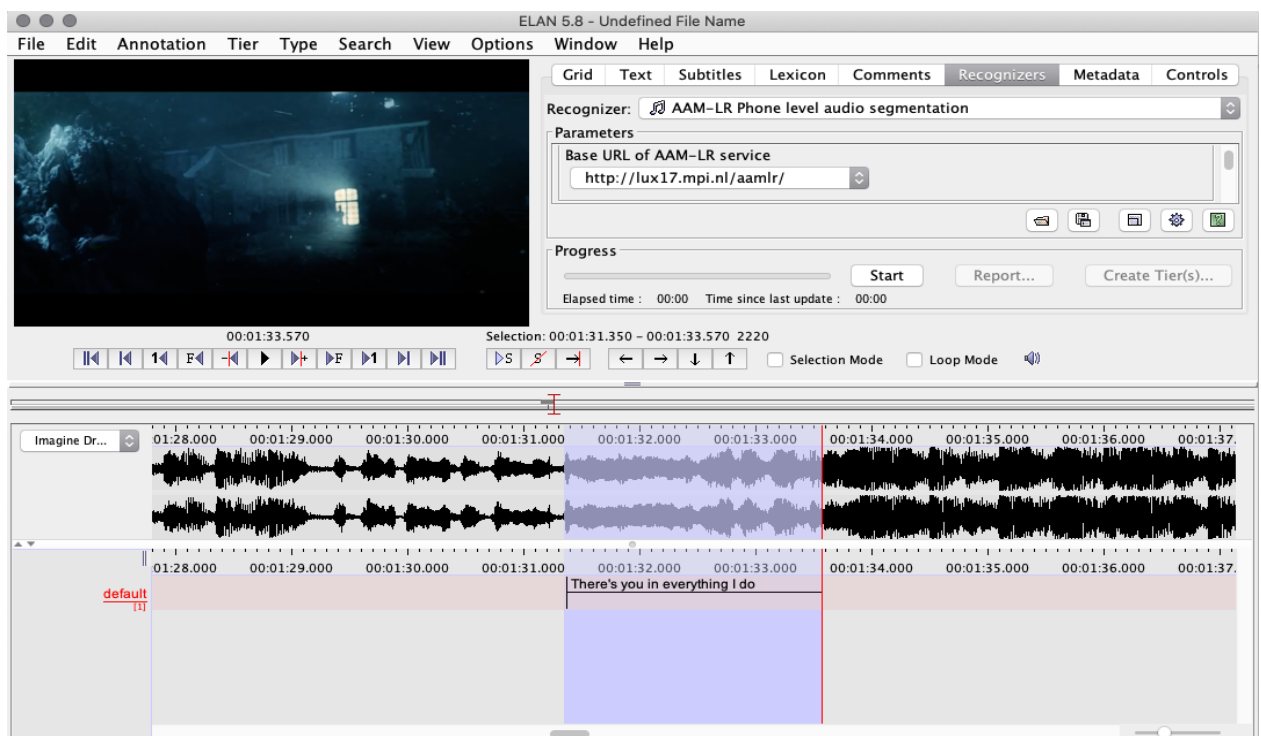


Figure 1. A screenshot of ELAN software analyzing “*I bet my life*” (2014) by *Imagine Dragons*.

Moreover, it does not interpret the data and does not analyze it. The software could be of a great help in analyzing videos as it helps to speed up the analytical processes and provides the possibility of splitting and overlapping video dimensions for successful analysis.

Film scholars introduced one more stylistic approach. They deal with trans-medial manifestations that are aimed to construe fiction's potential with the help of different stylistic frameworks. Thus, Forceville (2002) reassesses existing models for stylistic analysis from a multimodal perspective. He studied how means of non-verbal communication are presented in screen version (Schrader, 1990) of Ian McEwan's *The Comfort of Strangers* (1981) can transmit the same message as in a printed version.

In *Pictorial Metaphor in Advertising* (1996), Forceville puts at work a theoretical framework for pictorial metaphor analysis, which is applicable when two visual elements from different categories or semantic field are compared. He draws an example of an advertisement in London underground to illustrate this phenomenon. However, the pictorial metaphor grounds on visual representation and cognitive processes the representation evokes. At that, the verbal analysis lacks in this perspective.

The stylistician also suggests that multimodal constructions should be considered identical to mental structures that represent specific linguistic forms. This idea goes back to Lakoff & Johnson (2003 [1980]) and their theory of conceptual metaphor and gives rise to new developments in multimodal stylistics from a cognitive perspective.

Furher, Forceville & Urios-Aparisi (2009, p. 5) point out a significant role of cognitive metaphor that bounds to the multimodal one. As stated by Forceville (2009), there are two types of metaphors: monomodal and multimodal. The author insists that there are, at least, nine different modes: "(1) pictorial signs; (2) written signs; (3) spoken signs; (4) gestures; (5) sounds; (6) music; (7) smells; (8) tastes; (9) touch" (Forceville, 2009, p. 23).

One more mind-film theory is developed by George Currie (2014), who suggests that moving pictures are more like mental illusions (dreams, fantasies) than thoughts. Cognitive scholars attempt to figure out the process of engaging human minds into cinematic reality and seek to find how our experience of moving image and moving narrative coincides with our knowledge of seeing and perceiving the reality.

The core idea is to reveal addressees' feedback to film and find the most efficient way to perceive, construe, and interpret it. Currie advises that there is only one level of fabrication in films where the audience only forms a particular attitude to fictional events in the movie, but not to their relation to the fiction. This makes him reject the idea that the audience is an imagined spectator that participates in a scene or play the role of an observer: "What I shall object to is the idea that cinematic works encourage us to imagine ourselves to be observers of the fictional events, placed within the world of the fiction" (Currie, 2014).

All in all, the stylistician outlines two main factors of personal imagining:

1. audience imagines seeing made-up events in the movie;
2. audience perceives everything not from the film perspective, but the camera position.

To sum up, there is a high number of scholars who made a significant input in the multimodality studies. Based on their works, one can define at least three basic approaches to multimodal research:

1. systemic functional linguistics (SFL);
2. social semiotic (SS);
3. cognitive approach (CA).

SFL was developed by Halliday (1978) and further studied by Baldry and O'Halloran (2004). They focused on the way how the multimodal approach can link the verbal and visual aspects of an image. Jewitt (2009), Machin (2007), O'Halloran (2011), Kress & van Leeuwen (2006), O'Toole (2011 [1994]) studied the social semiotic approach. Their works proposed the opinion that the theory of semiotic resources performs four functions and underpins the grammar of relations

of the resources. Forceville & Urios-Aparisi (2009) and George Currie (2014) developed a cognitive aspect of multimodality, trying to explain the process of engaging human minds into cinematic reality. They also were aimed to find out how our experience of moving image and moving narrative coincides with our knowledge of seeing and perceiving the objective reality.

1.2. Multimodal analysis: Opportunities and challenges

One of the benefits of multimodality is its great essence, which suggests a new approach to the investigation of multimodal digital environments in terms of interaction and representation. However, multimodal perspective is still on the stage of its development in terms of the practice of language transcription, description as well as video data analysis.

There are two main benefits of multimodal analysis conducted with the help of digital technology (Smith, Tan, Podlasov, & Ohalloran, 2011):

1. ability to analyze with the help of different annotations within the same setting;
2. capacity to collect annotations for later investigations.

What is more, all these annotations can be generated either by human beings or by computer. The human-generated analysis includes verbal texts (usually written) and graphic overlays. So, these comments are made to a particular part of the text (visual or verbal), while computer-generated analyses can be automated or semi-automated and applied to videos, slot recognition, gesture, or speech recognition. By means of multimodal analysis software, one can use both types of annotations to conduct successful analysis and get valid results. At the same time, there are a lot of challenges for the stylisticians who deal with multimodal analysis.

However, revolutionary frameworks, analytical details and technological scope offered by a wide range of software resources evoke a lot of questions like *‘which aspects of a phenomenon should be examined?’* and *‘what*

approach/software to use for a particular aspect?`. Consequently, the matter of choice becomes a crucial factor of the study.

One more crucial factor is that usually, all the attempts to analyze any texts from a multimodal perspective become monomodal. Since researchers typically concentrate on a single mode, and, as a rule, it is a verbal one (e.g., Graesser, McNamara & Louwerse, 2004; Louwerse & Jeuniaux, 2009). This way blocks the other modes, and the analysis becomes simplified since they are viewed more like non-verbal elements, and that is the reason why they are not considered, for instance, in corpus linguistic analysis (Martin & Rose, 2008; Parodi, 2008; 2010).

Even though several markup languages like Standard Generalized Markup Language (SGML) and Extensive Markup Language (XML) (Bryan, 1988; CES, 2000) were developed in order to identify semiotic characteristics in multimodal texts, they cannot analyze present-day texts at the full scope. Moreover, computer-generated analysis lacks a well-grounded theory of multimodal language in the framework of visual analysis.

Conclusions to Chapter One

This Chapter presents a phenomenon of multimodality that is widely used in present-day society. Four milestones of multimodal study, developed by García, Flores, & Spotti (2016), are also outlined there. The milestones serve as the basis for collecting the multimodal data and its further interpretation.

Apart from that, there is a significant number of scholars studied multimodality from different perspectives. Based on their works, three approaches to multimodal studies can be outlined: systemic functional linguistics (SFL), social semiotic (SS), and cognitive approach (CA). However, these approaches are not universal since SFL works with verbal and visual modes only, SS concentrates on sonic, musical, pictorial, or linguistic model (but on stylistically marked elements only), and CA is more about the human mind and cinematic reality.

Based on the way the data is collected and interpreted, all the perspectives of the study can be divided into human analysis and computer analysis depending on the devices and approaches stylistician use. Besides, both ways of collecting and interpreting the information are not perfect if used separately. Machine alongside artificial intelligence lacks emotional intelligence and empathy, however, the human being needs more time to accumulate and process all the information, without omitting a single detail.

At that, apart from opportunities to analyze various texts from different modes of perception, several challenges are described in the Chapter as well. There are two main pitfalls of multimodality: a variety of methods the research and be conducted and interpreted and availability to analyze various modes in the text, but not their interaction and they produce. Choosing only one method of analysis limits a scholar in receiving a valid result of the research, however, using a lot of them is not only challenging but might lead to confusion. Concentration on all the modes that participate in a text is only a small part of multimodal research as not only the number of modes are crucial, but their interaction. In case the modes are investigated separately, the research becomes monomodal.

CHAPTER TWO. MUSICAL DISCOURSE: ANALYTICAL TOOLS

Musical discourse can be construed as a discourse of music or discourse about music (Blacking, 1982). Its structure is a momentous element since it contributes to the general impression of a particular piece of music, hence all elements overlap and create one indivisible unit with its similar and contrasting sections. The notion of structure presupposes that a human being acquires the ability for `musical` thought. Moreover, there is a high chance of musical forces which postulate that skillful listeners of tonal music "not only talk about music in terms used to describe physical motion, but experience musical motion as if it were shaped by quantifiable analogs of physical gravity, magnetism, and inertia" (Larson, 2012).

2.1. Structure of musical discourse

`Semiotics`, `social agents, relations and context`, as well as `text` are the main elements of musical discourse. They are connected, and they play a significant role in the discourse. However, each component can be viewed from a specific perspective since the absence of even one element can drastically change the whole perception of a music piece.

Songs can provide background for the structure of musical discourse. Some researchers do not believe that such discourse exists at all since all human responses towards the interaction of words and music (or text and melody) contribute to the idea of usage and evaluation of music. As a result, more contradictory issues, such as how words can create melody and vice versa, or how a particular idea is embodied in the interaction of words and melody.

Blacking (1982) assumes that if there is a specific pattern of musical discourse, it will be outlined by tracing the coherence of various musical systems and structures of their discourse in terms of reflection of concepts, intentions, and perceptions of its users. In such a case, similar features appear from such

ambiguity, one might be able to dwell upon the musical process with its universal structure of kinship and verbal language. Blacking created the following set of rules that could be applicable to identifying how speech and song interact and how they affect each other:

- Verbal mode only does not fully convey the main message of the song since words have symbolic meaning only, and they do not usually refer to the thing at the point.
- Words have an only formal sense as they contribute to the patterns of speech such as melody and tone.
- Music can shape the sequence of words even if it contradicts the syntactic structure of the sentence and violates grammatical rules of a particular language.
- The correspondence between mood, sense of words, and the structure of melody is not mandatory.
- One should abandon words since they distract further musical development and might lead to the discrepancy between all elements of a composition.
- The structure of the song is more likely to be influenced by prosodic features than by its verbal part.

Basically, Blacking (1982) argues that the song and the speech are equal. This equality presupposes that the speech (verbal text) is no longer the part of the song. At this point all the components of musical discourse and the song should be considered as equal. Based on this assumption, the song is, indeed, a multimodal unity.

One more important step in identifying the structure of a musical discourse is to step away from the terms that do not refer to the music but are applicable in different spheres. This, however, contradicts the present-day view on any study, since the idea of interdisciplinary research is of extensive use today.

Even though there is a drastic interest in popular music studies and music discourse analysis, it has not become a subject too complex linguistic research and analysis. There are three ways the discourse analysis can be found popular music:

in the research of song lyrics in the role of performed language, interpretation of music as discourse, or discourse as music (Bradby 2003). Therefore, there is no general definition of this phenomenon it is either 'music discourse' (Moore 2001) or 'musical discourse' (Tagg, 2003), 'song discourse' (Murphey, 1992), 'discourse of music' or 'discourse about music' (Blacking, 1982).

Fairclough (2003) differentiates three levels of abstraction in social analysis: events, practices, and structures. From his standpoint, discourse is used in two meanings: an abstract noun aimed to define a language or any other types of semiotics as elements of social life or count noun that underpins specific ways of representing facets of social life. He also claims that there is a correlation between text, context, and interpretation. So the "text" is the "product" of the "process of production," and the "resource" of the "process of interpretation" (Fairclough, 2001). All these principles shape the Critical Discourse Analysis (CDA).

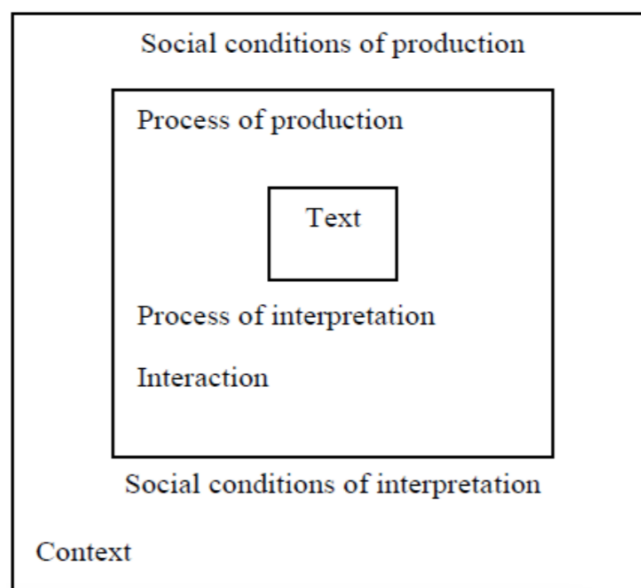


Figure 2. Fairclough (2001, p.21)

Apart from the fact that CDA allows us to understand how the text is composed, it emphasizes the choice of words. It also deals with the way text make meaning, how its components work the way they work and how the resources of the discourse are used. Transferring CDA into multimodal context evokes the interest in how discourses of other expressive domains (not only a verbal one) produce the social structure and constitute the world.

However, it provides not much information about the way the audience perceive the text or the idea the author wanted to convey

Blommaert (2005) steps away from CDA and switches toward ethnography and equates music to language, claiming that it is music that helps us in understanding the world around us. He presents five milestones for music discourse analysis that goes back to anthropology and sociolinguistics (from the above mentioned, word ‘language’ can be easily replaced by ‘music’):

- The role of the language in society depends on the importance of language to the people.
- To grasp the idea of how the language works, one should consider the linguistic environment since the context is of great importance here.
- Deviations of social meaning in the language is determined by its deviation in different contexts should be in focus.
- The language applies a specific pattern to its users, and they have always bound to them.
- Structure of the world system influence communicative events, what contributes to the globalization process in the modern world.

These principles facilitate new ideas in discourse analysis, and mainly the frames for utterances. Further suggestions in the ways people utter are presented in the structures such as *contextualization*, *uptake*, *intertextuality*, *indexicality*. The patterns are at work in the musical setting since they advocate the powerful effect of language in different settings (social, cultural, and political).

The concept of *contextualization* was developed by Gumperz (1982, 1992) to find ways people transmit the message and if they use direct or indirect speech acts. Blommaert (2005) utters that context does not exist intrinsically; it always requires dialogue (either a speaker and their interlocutor or audience). In most cases, this leads to an *uptake* - a part of a sequence of interactions (any action that is to be interpreted by the addressee). Furthermore, uptake has a temporal dimension, so it is relevant only after particular words or utterances.

While speaking people often use words of other people to communicate their ideas and thoughts, that is called *intertextuality*. Hence the term itself is borrowed from the literary theory, Blommaert (2005) finds its application in discourse analysis and points out that this phenomenon helps us to look a long way off the frames of a particular communicative event and define roots of certain words, their source, receiver and relation to reality.

Gumperz (2001) states that all our conversations depend on two types of verbal signs: symbolic signs that transmit the message with the help of grammatical and lexical rules and *indexical* signs that indicate a correlation between sign and context. These are usually personal pronouns, temporal, placement, spatial expressions.

One more approach is nominated by Scollon (2004). It is called a nexus analysis. The scholar elaborates on Wertsch (1991) and his core idea that a person usually employs ‘mediational means’ as culture and language clusters, and these means influence the action in a certain way. As social practice comes from real experience and real places, there are some complex aggregates (nexus) of different discourses with different relevance to certain social parties. Scollon (2004) defines nexus analysis as one of the approaches to discourse analysis that is attuned to music transmission. They also, like Blommaert’s, are linked to ethnography to analyze specific social practices (setting, drums in a local studio). Researchers seek the input of the language for its users, as well. What is more, they provide a set of recommendations for conducting nexus analysis:

- First of all, one should enter a zone of identification and get to know its main participants.
- After that, figure out how the participants came to be at this place in a particular moment and what makes them act in this way.
- The next step is to analyze concepts as mediation means.
- Further, it is significant to target interpersonal links and the structure of involvement of each participant.

- Then, it is essential to define the way social interest is reflected in the discourse.

2.2. Music videos as fragments of musical discourse

Since the beginning of the 1990s, scholars started to compare music videos with films (Vernallis, 2008) as these art types have a lot in common: camera movements and shots reflected in words (lyric) with accompaniment of a number of musical instruments or other sound effects, modes replacement (auditory mode can stand for visual one and vice versa).

It is a common practice today that film directors begin their career not with graduating a film-making school or writing scripts, but with creating music videos. This experience serves as individual training and an excellent start for the film-making industry. With the extensive usage of the Internet and clip thinking, music videos become a great way of influencing the audience and conveying different types of messages with the help of combining a range of perception media.

Since early 1980s, music videos have been created by means of mixing, remixing, and overlaying different modes (Vernallis, 2008). It greatly influenced the shaping of contemporary music video aesthetics. This extraordinary view facilitated a new approach to music, sound, and context analysis, and, of course, broadening the horizons of multimodal research. Gondry, a famous film director and ‘the most sought after video director in the world’, according to New York Times, started making a music video DVD compilation that was revolutionary for pop music of those times.

Music video imagery contains narrative elements. However, it has its well-organized structure, which is aimed to reflect the movement of the song, but it does not concentrate on the music itself. It happens so that music videos become a center of one or a couple of processes. As a reason, it is not easy to define the way that the audience will see; the best option to the viewer is to go with the flow of the video, since the duration might be unexpectedly prolonged or unusually cut off or

disturbed by images of singer/band performing. The video's main points might be scattered among several sections of the video, and the footage material is an excellent hint for an uncanny sense of return. Moreover, not only the footage but the other parts like sound (as well as the other codes perceived by different modes) can be a great linker between different parts of one music video.

There is a pattern of musical video described by Vernallis (2008), where a structure is shown as a system of cycles, loops, strings of motifs with inserted video section (nested patterns). There is a particular way the sound sizes in the foreground. Mainly, there is a cluster of sounds: electronic noise, keyboard tremolo, clarinet motif, record scratch, drum beep, animal calls, computer clicks. Other sounds often stem from a mix of that basic sounds. Visual material phrased with the sound links to some part of our memories and helps the audience to grasp the idea of the music video. At that, images in a music video can draw the audience's attention to an audio recording and offer various ways of its interpretation.

Furthermore, visual motifs can be attached to different musical materials such as a rhythmic motive, bassline, odd timbral effect, without disturbing or influencing each other. As a rule, they are multimodal, and they might present conflict, contradictory evidence, or support each other at the same time.

One more crucial aspect of the music video is color. Color timers are usually bound to each shot individually so that they can reflect movements in the music (Vernallis, 2004). The change of colors attracts to materiality, texture, and surface. Rapid change of color and visual gap motivate the viewer to look for the meaning where they can create an additional source close to the soundtrack.

Hence, music videos are mainly short and compressed, they lack dialogue, and they are to advertise a singer/band in some way, emphasize lyric and put musical material into the light, music video directors look for the slots that can reflect human emotions or feelings that can be projected over the music. Besides, such videos are not classical performances since they lack an effect on relationships. This is the reason why music video directors cooperate with

performers or dancers who possess hypermobility and can speak through their bodies (Kaplan, 2016).

Music videos incorporate various kinds of experience, and at that, music can disappear from the audience. The same is with imagery concerning music, as it often comes and goes so fast that one cannot experience it to the full extent. One more ambiguous part is the sense of time, as the audience cannot understand it as it leads to memory. Moreover, such videos often deal with flashbacks, flashforwards, and different shifts in time as the image might be projected based on the lyric, harmony, or arrangement of memories from the past. This effect is created with the help of the montage sequence and montage shift, which is of high importance in this sphere. The most substantial response from the audience is received by combining montage and narrative as in this way the scene appears to be the most memorable and touching.

There are different approaches to defining interaction in music videos, and one of them is the interplay of audio and visual modes (Gillet, Essid, & Richard, 2007). In this case, high-level structures of song and video are separately viewed as the one to measure the connection between artists' performance and sound and video. For example, most mainstream music videos present dances or performance in order not to reveal the message, but to attract attention to particular performers. However, there are videos with high-level features such as sophisticated shot structures that contribute to mood or visual metaphors creation. In music sequences, a significant number of 'silent,' as well as foregrounded events, can be presented with the help of music signals or chord changes.

Music pieces can be segmented in groups or sections united by different characteristics such as dynamics, tonality, timbre, chores, verses. The segments can be differentiated by outlining large blocks of matrix or by various methods that determine boundaries between sectors.

As for a video part, a high level of description is acquired using shot segmentation. Shots are semantically crucial as they usually correlate with the rhythm or section changes in music.

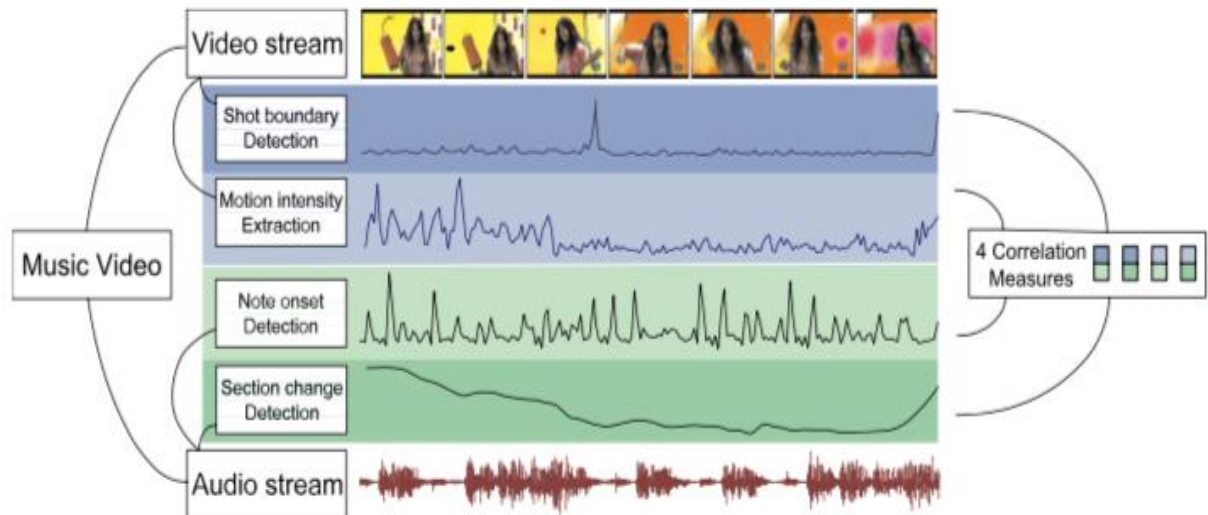


Figure 3. Overview of the audio-visual content structuring system (Gillet, Essid, & Richard, 2007).

The four-step segmentation procedure presented above reflects the peaks where event or section change is detected. Functions of these peaks are aimed to obtain the temporary location of unmarked events and segment boundaries to define correlations.

Several studies contribute to solo transcription of a particular musical instrument (Joder, Essid, & Richard, 2009) or monomodal studies that dwell upon audio only (Gillet & Richard, 2008). Still, there is a limited number of studies exploiting multimodality. It is worth mentioning, that many multimodal experiments focus on connections between movements and emotions, as well as relationships between musical accents in drumming and video features, were analyzed by S. Dahl (Dahl, 2000).

However, some scholars (Poh & Kittler, 2010) consider Joder, Essid, and Richard`s (2009) research to present a couple of early-fusion and late-fusion approaches to drum-solo with a feature-level fusion and succession of audio and video features that can archive a remarkable upgrade of monomodal techniques.

A slightly different approach is presented by McGuinness, Gillet, O`Connor, & Richard (2007), where multimodality in the video is compared to a detection process (Figure 5).

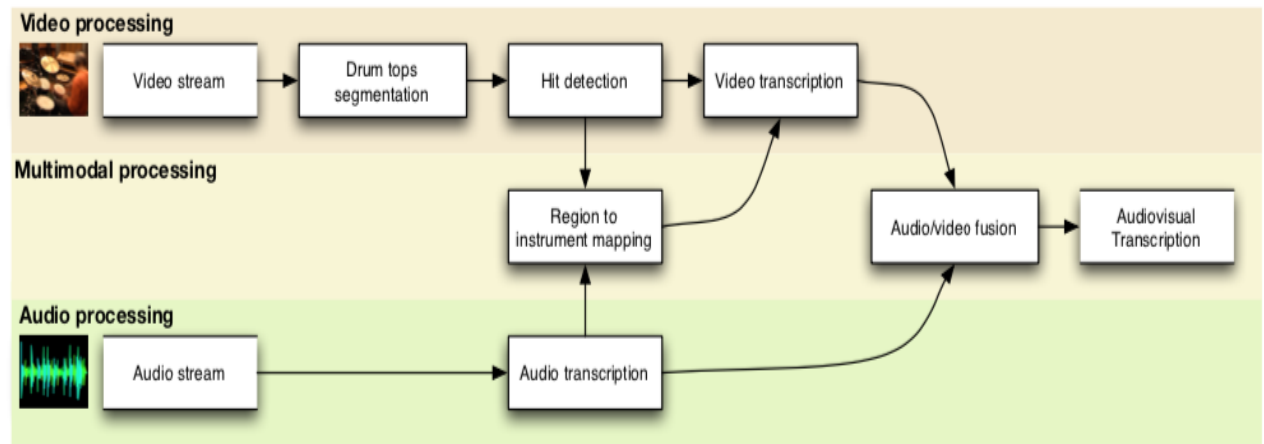


Figure 4. Overview of the audio-visual content structural system (McGuinness, Gillet, O'Connor & Richard, 2007).

The first step in their scheme is an analysis of video sequence to detect the position of each drum element in the scene, and, to be more precise, the part of the instrument used in a detection process. The drum tops are presented in a circular shape, and they are detected with the help of geometric criteria. Further, on each selected drum tops, drum strokes are defined with the help of a simple motion intensity features, coupled with a foreground object of segmentation. Transcription is made of identification of drum instruments that correlate with the detected drum top. Once a video transcription is applied, it can be combined with audio (or any other) transcription received from different cameras. Also, the same transcription system can be used to assigning each detected region to a particular drum instrument.

This multimodal system leaves behind the monomodal system as well as the system based on the traditional fusion methods (Gillet & Richard, 2006). To sum up, using high-level information received within one mode aimed to process the other one can be more successful than the use of direct feature-level or decision-level fusion.

Music videos are in the focus of not only multimodal linguistic and stylistic studies, but also for various media and cultural investigations. McQuail (2000) outlines music videos as one of the most comprehensive and successful postmodern television services since it is a dominant part of contemporary art that resulted in a new aesthetics of cultural revolution. Schwartz and Ratner (2007)

suggested not only the ways the audience perceives this kind of art but also how the music industry is shaped to produce an effect upon the audience. This insider's approach facilitates understanding of how music videos work from both sides: as an instrument and as an entertainment.

Verbal and nonverbal elements of the music video are in the focus of content analysis (Wallis, 2005). This analysis is closely related to cultural patterns and McKee and Pardun's qualitative study (2003), where a college-student provides their interpretation of music videos based on views of present-day society, influenced by the culture of popular music texts. One more communicational approach was suggested by Gow (1992), who categorized the videos according to the combination of context forms.

Moreover, there are relational approaches (Moore 2003) that work from another perspective, aiming to blur the lines of musical genre (Mittell, 2000), context, and textual specifications. Vernallis (2004) pays attention to multi-layered discourses and critical analysis of appearances and stereotypes. Martinec (2000) puts rhythmic hierarchies, language level, phrases, and foregrounding music, and action into light in his review.

Application of transcriptional strategies is closely associated with multimodal analysis due to structural specifics of a lyric video. Free editing software Adobe Premiere allows study of a juxtaposition of sound, images, movements, words at the frame, and shot level throughout video continuity.

One more approach to music video analysis is suggested by Professor Lori Burns from University of Rochester (UR Institute for Popular Music, 2016). She suggest that the interpretive framework for the music videos goes out of three theoretical perspective: genre, CDA, and narrative theory. Each of the theories is concerned the ways of doing things and. Genre theories explore cultural norms and values, create shared realities and shape understandings of the world. CDA aim is to bear a discursive determinants that dive texts and doing so examines how texts do persuasive work. Narrative theories are concerned with how stories are told, who tells them and who is the speaker. Basically, this framework facilitates

systematic thinking about how the individual domains of words, music and images, that go together in mutually reinforcing ways to be culturally productive and constitutive of the social realm. Professor uses CDA in a multimodal video and musical context to explain how images can ‘say’ things people are not able to communicate in words.

2.3. Analysis of music and lyrics in musical discourse

In the past few years, a lot of open source applications were released to research in the musical field, and MSAF (Music Structure Analysis Framework) is not an exception. This is an open-source framework written in Python that enables us to analyze the structure of music with the help of combining computation features, algorithm implementations, annotated datasets, evaluation metrics and music information retrieval (MIR) (Bruderer, 2008).

It is worthwhile pointing out that different music structure algorithms accept various features for uncovering loudness, rhythm, or melody. There is a set of features that can be revealed not only by means of MSAF but with libROSA (one more python package that helps to analyze music) as well (Mcfee et al., 2015). Here are some features MSAF can elaborate on using libROSA according to Mcfee, et al.: Pitch Class Profiles (PCPs, representing harmony), Mel-Frequency Cepstral Coefficients (MFCCs, representing timbre), Tonal Centroids (or Tonnetz, representing harmony), and Constant-Q Transform (CQT, representing harmony, timbre and loudness). What is more, each feature highly correlates with additional parameters as hop size, sampling, and bit rate.

Particular algorithms detecting the boundaries and structure of music were presented. They are classified based on the task they are aimed at. Mainly, MSAF includes eight algorithms:

| Algorithm | Boundary | Grouping |
|--|----------|----------|
| 2D-Fourier Magnitude Coeffs (Nieto & Bello, 2014) | No | Yes |
| Checkerboard Kernel (Foote, 2000) | Yes | No |
| Constrained Cluster (Levy & Sandler, 2008) | Yes | Yes |
| Convex NMF (Nieto & Jehan, 2013) | Yes | Yes |
| Laplacian Segmentation (Mcfee, Bertin-Mahieux, Ellis, & Lanckriet, 2012) | Yes | Yes |
| Ordinal LDA (Mcfee & Ellis, 2014) | Yes | No |
| Shift Invariant PLCA (Weiss & Bello, 2011) | Yes | Yes |
| Structural Features (Serra, Muller, Grosche, & Arcos, 2014) | Yes | No |

Table 1. Approaches included in MSAF and used in experiments (Mcfee, Nieto, Farbood, & Bello, 2017).

Different types of evaluation metrics are applied to detect and evaluate boundaries and grouping. For instance, for boundary detection, usually Hit Rate is used (limits are considered to be `hits` under the condition that they fall within a specific timeframe from the reference). In such a way, it is easier to determine the

number of correct boundaries (Precision) and the number of reference boundaries (Recall) within a certain time (F). As a rule, F is typically 3 or 0.5 seconds (Nieto. O. et al., 2014).

With the help of different experiments, it has been proven that the stronger weight on Precision is, the better is perception of music (Nieto & Jehan, 2013). One more classic metric is the Median Deviation (Torrent et al., 2010), where the median deviation from each reference is to the estimated boundary and vice versa.

Another interesting metric is the Pairwise Frame clustering (Levy & Sandler, 2008) since it contrasts the parameter based on the cluster (label) the estimation and reference belongs to. The Recall here is viewed as the ratio between sets (usually two) over the amount of the same texts and the Precision - number of similar estimation. The F-measure connects these two parameters.

All three metrics are at work at MSAF, depending on the set of methods the researchers use to achieve the aim of the study. However, there is a significant oversight in the metrics due to the subjectivity of the results. That is why the application of different parameters and the usage of more than one set of boundaries will help to receive valid results on an experiment.

Therefore, using a single approach or only one metrics might mislead the results of the whole study, so it is highly vital to accumulate as many references as possible for effectively accessing the boundary algorithms.

It is highly essential to elaborate on our music perception to analyze a piece of music from a multimodal perspective. Figure 3 presents the human signal analysis framework. The first three stages stand for listening and the last three – for learning. Two stages in the middle, where listening and learning overlap, stand for musical cognition.

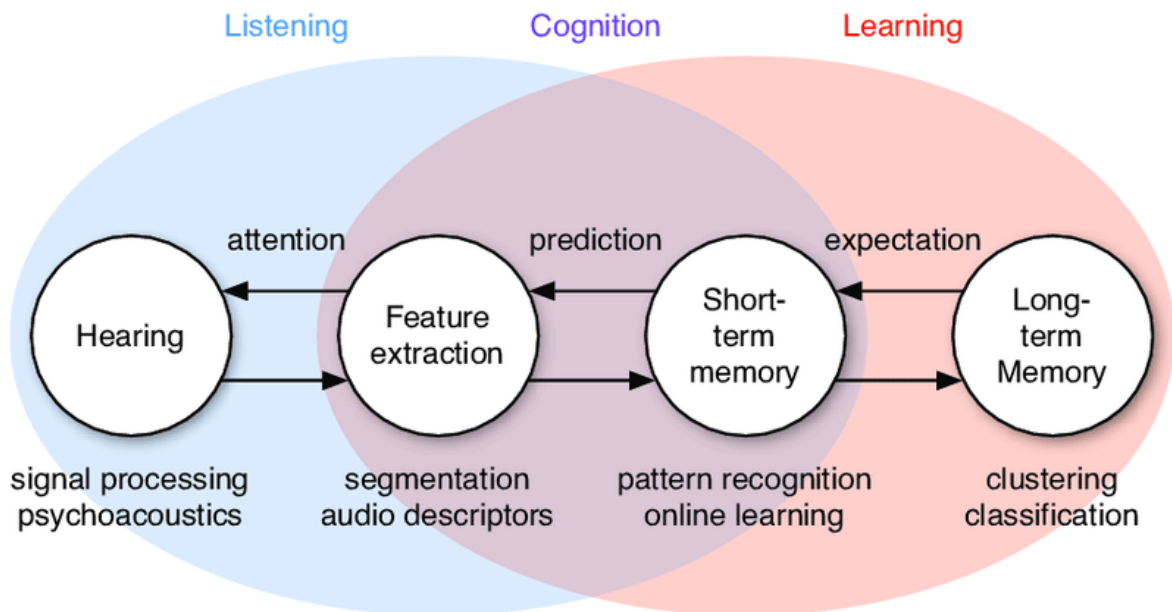


Figure 5. Music signal analysis framework (Jehan, 2005).

Hearing is the first stage that serves as a filter for further music analysis. Our ear perceives less than 10% of the incoming signal, and that is even though it is also psychologically limited, meaning that we hear the things that we want to hear or things that are foregrounded. The signal we receive can be converted in an auditory spectrogram, where the things that are represented by the time-frequency correspond to the particular parts of the audio we hear. All the data collected is construed by scholars in psychoacoustics.

Feature extraction is the next stage, and it deals with converting signals into symbolic representation (something alike “DNA” sequence, but applied to music). This is a point where source separation starts. Here perceptual characteristics or description signals might be singled out. So, this stage compresses the musical content by taking its segments apart.

The third stage is a short-time memory domain, where the information is analyzed to detect foregrounded patterns and redundancy. In such a way, associations and logical chains appear, and they lead to the algorithm of long-term memory (Jehan, 2005).

2.4 Analysis of visual elements in musical discourse

Since we live in the era of technology, the way we perceive visual information influences the way we construe it. That is why visualization tools are of great importance (Spence, 2007). Visualizations often are viewed as a core of complex information analysis as its tools influence the way we think about the data. The work of these instruments is apt to the cognitive process, which was proved by human-computer interaction researchers (HCI). The nature of visual information analysis is a popular topic for scholarly research.

Stylisticians outlined the frameworks explaining the way individuals perceive the visual information and how they make use of it. These tools have some standard features of forming addressees participation in the visualization process as an interactive chain of elements, where each of them has its level of abstraction and focus.

Card et al. (1999) come up with a knowledge crystallization cycle, a level model of activity where the aim is to grasp understanding from data that is relative to some action/task. The elements of this pattern contain different activities from searching for data to elaborating on researched findings.

Spence (1999) improves this model by exploring the ‘foraging for data’ element in terms of visual operation. Thus, he connects visual navigation to cognitive activities, claiming how the audience can collect and interpret data. Other researchers focus on the tools design and the way its users manipulate the picture and on visualization transformation parameters (Card et al., 1999; Chi & Riedl, (n.d.); Jankun-Kelly & Gertz, 2007).

Jankun-Kelly and Gertz offer a model for analyzing the audience interplay with the visualization system, which is called a model of visual exploration (Jankun-Kelly & Gertz, 2007). The main idea was to record interaction in the visualization of the system and bound the results to the controlled parameters. The milestone of this pattern is that the fundamental operation in the visual investigation is in the manipulation of visual parameters. However, the models are

useful in the temporal setting of visual analysis, while higher-level semantic of users` interaction cannot be interpreted with the help of these frames.

The way people work with information is also viewed from a task-centric perspective. Shneiderman identifies a two-step procedure and advocates various assignments that information visualization tools have to support to keep the problem-solving process: overview, zoom, filter, details-on-demand, relate, history, and extract (Shneiderman, 1996). These activities serve as a rule for the design of information visualization tools (Craft & Cairns, 2005).

Amar and Stasko identify higher-level analytic activities that the audience of the visual system would usually do. These are complex decision-making, learning a domain, identifying the nature of trends, and predicting the future (Amar & Stasko, 2005).

In visual analysis using cave automatic virtual environments, there is a five-step pattern of behavior from problem interpretation to negotiation of discoveries (Park, Kapoor & Leigh, 2000). Mark et al. (Mark, Carpenter, & Kobsa) (Figure 2).

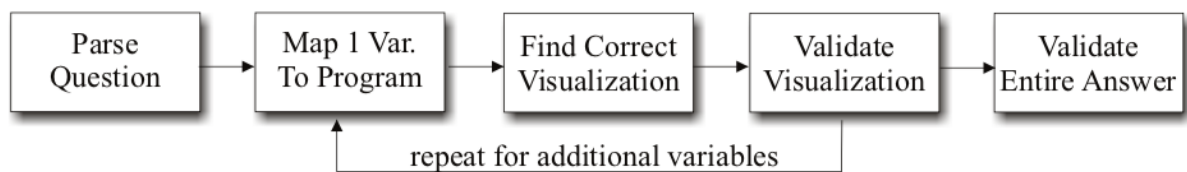


Figure 6. Mark, Carpenter, & Kobsa: outline a five-stage model of collaborative information visualization tool use.

Temporal succession of steps in the model roots from the study of pairs solving free data discovery as well as concentrated question tasks in distributed and co-located settings. The models have some similar features, but they are far from being identical.

Conclusions to Chapter Two

The Chapter underlines the importance of music elements and their structure. The structure plays a crucial role in musical discourse as it not only organizes the piece of music but also contributes to its meaning. Broken structure of composition might lead to misinterpretation of the main idea of the song. This idea is well-supported by Blacking (1982) and his pattern of musical discourse. What is more, the researcher studied the interaction of speech (lyrics) and songs.

Scollon (2004) and Blommaert (2005) have a rather ethnographical standpoint. However, the idea remains the same as that music is equal to the language and helps us understand the world better. Their research and principle gave start to new ideas in musical discourse analysis. Further theories in the ways people utter are presented by Gumperz (2001) and Scollon (2004) in structures such as contextualization, uptake, intertextuality, indexicality, and nexus analysis.

As videos have been created employing mixing, remixing, and overlaying different modes, they became a part of musical discourse. There are different approaches to defining interaction in music videos, and one of them is the interplay of audio and visual modes (Gillet, Essid, & Richard, 2007; Poh & Kittler, 2010).

Music videos have their structure that contributes to the general idea of the music video. An arrangement of a music video is described by Vernallis (2008), where a structure is shown as a system of cycles, loops, strings of motifs with inserted video section (nested patterns), color. At that, this message will be broken in case of omitting one of the modes used in the video. That is why transcriptional strategies are of great importance in the music video with multimodal analysis due to structural specifics of this type of art.

Analysis of music and lyrics in musical discourse is at work in present-day society due to a wide range of software and applications. MSAF (Music Structure Analysis Framework), MIR (music information retrieval), libROSA enable music structure analysis with the help of combining different computation features, algorithms, datasets, and metrics.

Analysis of visual elements in musical discourse is of great importance in the Chapter due to the high significance of video tools in the music video. The tools are aimed at addressees participation in the visualization process. Card et al. (1999) create a pattern where the audience is intended to understand the data based on some action. The model was improved by Spence (1999) as he connects this visual activity to a mental one. Other researchers focus on the tools design and the way its users manipulate the picture and on visualization transformation parameters (Card et al., 1999; Chi & Riedl, (n.d.); Jankun-Kelly & Gertz, 2007).

CHAPTER THREE. MULTIMODAL ANALYSIS OF 'IMAGINE DRAGONS' MUSIC VIDEOS

A multimodal analysis of *Imagine Dragons* Pop Rock Band's Discourse is conducted with the help of software ELAN. The results are based on the data collected and the emotions and feelings the videos evoke. Besides, each pattern of analysis outlines a particular topic that the video is famous for. At that, the correlation between the modes of perception, topic of the video and response from the audience is presented in the Chapter.

3.1 “Believer” (2017): structure and interaction of modes

Believer (2017) by *Imagine Dragons* is considered to be one of the most popular songs in the world (McIntyre, 2018) and one of the biggest songs in 2017 (McDermott, 2017). The main reason why the song, as well as the music video, is so popular is a great interplay of all the modes that participate in meaning making and while perceiving this piece of art.

Mainly, the music video is a great reflection of the lyrics not only in its direct meaning, but in its metaphoric interpretation as well. One more important thing to mention is that the song is believed to be autobiographical. As is known, real stories told by real people evoke understanding and sympathy even among the huge audience all over the world.

As in *Believer* (2017), the music video reflects lyrics, let us consider the structure of the video in detail to prove this assumption. From the very first seconds, one can see Dan Reynolds, the band's frontman, who is boxing.

The general atmosphere is quite gloomy due to the visual effects (the red light at the beginning signalized about danger and fight) (Figure 6), alongside with visual effect, the audience hears a loud vivid beat that serves as an accompaniment to the punch.



Figure 7. Believer at 00:01 (Imagine Dragons, 2017)

Furthermore, every punch alongside the video will go with the same musical bid. After the punch, something fragile breaks up and no one exactly knows what. This is also a milestone moment as it contributes to the further development of the video as the next is zoomed eyes of a child, who appears later on.

The next scene is created with the help of striking visuals and violet colour, where a human body is created from small digital pieces and this point on, it is not exactly clear where the action takes place: either dreamlike or outer space location. All these sounds and visuals are like small prehistory to the verses.

To sum up, before the song starts we have three key figures in the video: a frontman, a child and an unknown human body created from digital pieces.

Verses (verbal mode) are very sequential as narrative in them starts in the same manner:

Verse 1: *First things first*

Verse 2: *Second things second*

Verse 3: *Third things third*

Verse 4: *Last things last*

Moreover, these verses always go together with the same melody that reinforces the hook for the audience.

Right after the first stanza, one can see a man (Dolph Lundgren, a legendary actor who was famous for boxing scenes), and a DNA structure outline right after

that shot (mainly, three samples of the same outline). This sequence brings some light into the story and makes the audience think that Dan Reynolds, boxer and a child are somehow connected. One more proof is the shot with the pyramid and a drop, that splits into three smaller drops. On the background of the pyramid, one can see three different colours that shape one stripe.

A close-up appearance of a digital boxing glove (Figure 7) followed by a close-up of Dolph Lundgren (who is lying on the ground and stands up with the music beats like a digital creation) gives the idea that a fight is on the way.

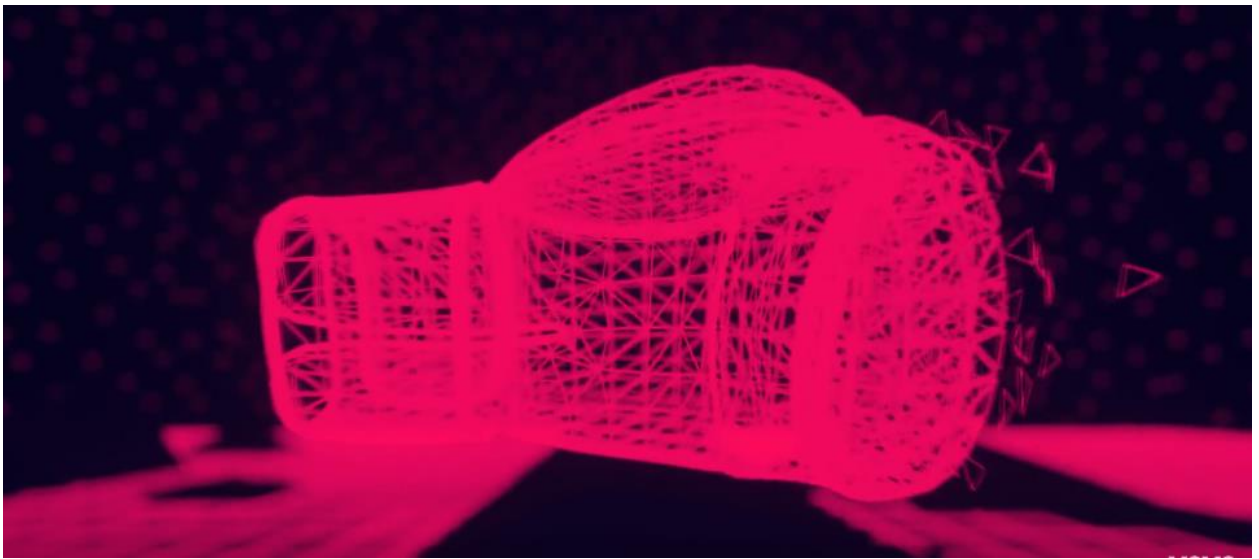


Figure 8. Believer at 00:05 (Imagine Dragons, 2017)

Words “*Write down my poems for the few*” are pronounced while a child is sitting and writing something, what makes the audience think that a child is like a flashback of the frontman, this is a reflection of his personality in childhood. Two next lines “*Taking my message from the veins // Speaking my lesson from the brain*” the camera zooms the boxer’s face and alongside with the word lesson and strong beats it leads the audience to one of the most powerful episodes of the video. Before that, there is a dramatic pause that is followed by annoying woosh, which contributes to the creation of a tense and anticipating atmosphere.

The effect of each chorus is intensified by this woosh and lull that extend the two last beats of the pre-chorus and the first two beats of the chorus. The tension peak is at the end of a pre-chorus stanza Seeing the beauty through the and resolves

in the *PAIN*. The mentioned metaphor is based on contrast, it does not just activate the physical fight, but also the mental one, where Dan Reynolds seeks for answers.

Interestingly, there is antithesis in lines about pain “*(Pain) you made me a believer, believer // You break me down, you build me up, believer, believer*” which is also reflected visually: the camera does not emphasize the fact that somebody is punching somebody, but the fact that somebody is being punched. To evoke an emotional response among the audience, camera emphasises the first punch is such a way that it is not clear if the boxer punches Dan Reynolds or the audience. At that, a great tactile image is created.

What is for a fight, at the beginning of it (the first minute of the video), the frontman looks very eager to participate in this confrontation as his mood is quite elevated and he seems to tell ‘*come on*’ without any words, but only with the help of his body and lips movements.

Moreover, a few seconds later a close-up of his widely-opened eyes is presented (the moment when he opens the eyes goes along with the musical bid and it contributes to the tense creation). One more milestone episode in this sequence is a close-up of the mouth shouting *PAIN*, that creates an atmosphere that the everything absorbed the pain not only from the outside (bruises and wounds) but it also affects the body from the inside.

The pre-last stanza of chorus highlights the core idea of the song as well as the music video “*My life, my love, my drive, it came from PAIN*”. At that point, the boxer and the child closed their eyes, underlying the visual and emotional aspect of pain. The words “*Send a prayer to the ones up above....Spirit up above*” are visualized with raising the hands towards the sky. This gesture is connected with the Bible and prayers who lift their hands.

What is more, this gesture (as well as the mentioned words) is closely connected with an orientational (conceptual) metaphor (Lakoff & Johnson, 2003 [1980]) MORE IS UP, CONTROL IS UP, RATIONAL IS UP. In general, up is always connected do something good and positive.

Similarly, “*All the hate that you've heard has turned your spirit to a dove*” is a metaphor based on contrast. As a *dove* is a symbol of *peace*, not *hate*, *pain* or *struggling*. However, both the lyrics and the video, construct the images of *pain* and *hate* not as destroying the person, but as something people have to learn to live with to become stronger. Good things in life come from that strength and overcoming adversities.

Moving on, the tension is increased with the help of parallel constructions in the lyrics *I was choking in the crowd // Building my rain up in the cloud // Falling like ashes to the ground // Hoping my feelings, they would drown // But they never did, ever lived, ebbing and flowing // Inhibited, limited* as well as visual and audio effects. Due to quick change of sequences presented in various colors and combined with a dynamic rhythm of the song. The audience is waiting for a new round in this fight, which, by the way, starts from a very rapid close-up change of music video main participants` coming along with a drum beat and ending with almost complete illumination of the scene.

The last stanza starts from last things last and it makes the audience think that the conflict is resolved as Dan Reynolds is sitting in a chair and singing, what looks like he is concluding the story. However, later on, the scenes of the fight are shown with even more cruel moments. Even though, the rhythm and the musical does not differ from the one that was presented initially.

The controversy is mostly in visual effects as the black screen appeared and, along with it, the music turned off for a second and it leads the audience to the most unexpected element in the video, where the band's frontman says that he wants to stop this fight, but the boxer replies that they cannot do it. Interestingly, that scene leads to more dramatic one, where it is really hard to understand who will win that fight.

Despite the close-ups with Dan Reynolds' blood and tears, he destroyed his adversary, who was, the most likely, better or worst version of him. The destruction was also shown visually with the help of the digital body that was built

up at the beginning of the video. At that, the words “*You break me down, you build me uptake on a new meaning*”.

At the end of the music video, the addressees see what the boy was drawing during the whole fight. The picture is the same as the tattoo on Dan Reynolds’ chest. This fact unites the child with the frontman and famous boxer, making them the same person, who through struggles and pain become better.

3.2 “Radioactive” (2012): Apocalyptic motives and allegory

The band’s very first single *Radioactive* became a sleeper hit, occupying the third position in the *US Billboard Hot 100* chart and being the third best-selling song in that country in 2013 (Edwards, 2015).

At first sight, the song is about the first ex-prisoner’s impressions of the new world and their readjusting experience. At that, the world is depicted through the prism of ash, dust and dimmed colour. Alongside with musical effect, this visual atmosphere brings some pre-apocalyptic agurity that will spot some light on the course of further events.

The music video begins in a bit unusual way as the audience see a stranger coming on, carrying something in their hands. The colours and the lights are gloomy and dimmed, which also contributes to creating a general mood. What is for the weather, it is autumn outside, that represents the fall of everything. More than that, the wind is blowing (one can not only see how the trees and leaves are moving, but also hear the howl of the wind) and the noise of raven is warning that something is going to happen. Judging by the first five seconds, this something will be serious and mysterious.

The next shot is taken from a rather different perspective. All the audience can see is a back of the stranger. The camera works in a way that consciously or unconsciously the one who watches the video, follows the stranger and goes into the heart of the forest. The music escalates atmosphere so that this scene works as a hook for attracting attention. All these tools create an effect that the person wants the audience to follow them and to hear their story.

From the next scene, it becomes clear who the strange is: a blue-eyed girl, who carries something like a box. Taking into account the setting and the atmosphere it all looks like that is a Pandora box that will be opened soon. The focus of the camera is switching between the girl, the box and the dark cellar that looks like a prison. The girl keeps moving and the music beats are more frequent, intensifying the motion scenes. While the camera is zooming the girl's face, the alarming sound is heard. It is like an alarming siren that informs about the danger.

The siren, as well as the scene with the girl moving to the cabin, is interrupted by the drum beat that goes together with a completely new scene and new people. A loud drumbeat is "*waking up, to ash and dust*" with an element of irony as waking up is usually *clarity*, but instead, the audience can observe *obscurity* and *darkness* of the post-apocalyptic world. Moreover, *ash* and *dust* are visually presented in the video further, namely, they appear on Dan Reynolds' guitar. At the moment when ash and dust are mentioned, the man is throwing around the money, that might not mean anything, only shallowness that comes and goes with the speed of the wind.

What is for new settings, it seems like a girl entered the cabin she was moving to. There is a lot of hustle and bustle inside as the men are gambling, but it is not clear what exactly they are playing.

Both visual and auditory images enhance the lyrics in the line "*I'm breathing in the chemicals*", where the ex-prisoner, instead of fresh air, inhales chemicals of a toxic city of the new age, reflected in the video. What is more, in the following line one can hear the prisoner's gasp which is not only struggling for air but struggling for a new life after prison. After a loud gasp, there is a dramatic pause that underlines the drops of the water. It makes the audience think that the prisoner is somewhere in the basement.

In the scene when prisoners are looking through the bars (Figure 9) as they are afraid of what is on the outside. It can be construed by non-verbal signs: wide-open eyes, lips moving in silent "*On, no*". The episode is connected to the feeling that prisoners are going to feel when they get out of prison and start living in society. This is also a representation of how hard life in real world could be. The

moment becomes ironical when Dan Reynolds looks up and hears the voices from above.



Figure 9. Radioactive at 00:52 (Imagine Dragons, 2012)

In the scene when prisoners are looking through the bars (Figure 9) as they are afraid of what is on the outside. It can be construed by non-verbal signs: wide-open eyes, lips moving in silent “*On, no*”. The episode is connected to the feeling that prisoners are going to feel when they get out of prison and start living in society. This is also a representation of how hard life in real world could be. The moment becomes ironical when Dan Reynolds looks up and hears the voices from above.

Apart from those voices and noises, a bell rings, symbolising that a fight has begun. The prisoner`s posture is changing and he stoops his head, showing that he is disappointed hearing the battle above.

The world turns upside down once again and the scenes from above are at focus. The men are sitting at the table with a key hanging on his neck. Usually, a key is a symbol of opening something. To open something can only someone who has the power or knows where the key is.

The line “*This is it, the apocalypse*” sounds right after a blackboard with a score (Figure 10) is shown. This goes for a pre-history of a puppet fight that will be shown later on. As is seen, the ‘*champ*’ always wins and no one has broken this system yet. As usual, here is a dramatic fight before an epic fight.



Figure 10. *Radioactive* at 01:07 (*Imagine Dragons*, 2012)

A purple monster stands for champion and it is supported by the company of the men standing around. The chorus starts from the words “*I’m waking up, I feel it in my bones*” and a loud drum beat, that activates a violet monster and it provides a great “*welcome to the new age*” to a toy in a usual prisoners` clothes, a stripper orange shirt.

The repetition of “*Welcome to the new age, to the new age*” draws addressees’ attention to what is going on in the video, where two toys are fighting and the crowd is observing it. In turn, lexical unit ‘*welcome*’ is a signal of verbal irony and, to some extent, defeated expectancy effect as the *new age* is meant to be full of hope and new opportunities, but not a boxing ring.

The puppet of the prisoner is knockdown by the monster and it is spanking down not only with a loud drum bid. Interestingly, the moment of the bid is shown as Dan Reynolds’ beats the drum, it produces the sound and a toy is falling. So here is a logical sequence of events presented both visually and auditory.

Right after knockdown, a hand crosses a new mark on the blackboard, reflecting one more haunted. The man with the key gives monster`s nonverbal hints and encouragement (the man is running his finger across his throat), what

brings to new victories of a champion. As a rule, all the fallen are sent to the prison with the help of a special hatch in the floor, controlled by a man with the key.

The purple monster stands for life in the prisoner`s eyes. As is seen from the previous scene, it is aggressive and difficult to beat. Monster`s red eyes connect back to a toxic chemical atmosphere mentioned before, representing the new age.

On pre-last repetition of the chorus, the girl reveals what was in her hands. It was not a box, it was a little cage with a pink teddy-bear, whom she lets fight. During the chorus everything remains the same: the monster seems to win this fight as well. The men encourage monster with throwing money to the ring, which makes the situation even worse (for a pink bear).

However, as “*All systems go, the sun hasn't died*” the bear fights the violet monster with. As a symbol of victory, a girl steals the keys from the man and lets the prisoners go away from prison. The man (now without the key), gets into prison and the squeaking sound of the puppets, as well as their image, follows him.

The song and music video represents a fighting ring and those surroundings that are shown in the video is an allegory to a pressing issue ex-prisoners have gone back to a normal life. The message is conveyed with the help of the interplay of visual and aural modes. Moreover, conveying a message by having a speaker the victim of the issue makes the audience sympathize the prisoners.

At that, the repetition of the word ‘*radioactive*’, as well as the title of the song, bring clarity as radioactive elements are usually unstable and it is not easy to determine what is going to happen with them or what they will cause. This is also with the prisoners shown in the video: no one expected them to get out of prison in such a way.

3.3. Visual symbolism in “Shots” (2015)

From the very beginning of the music video, the audience can see the lines “*Imagine Dragons Shots inspired by the paintings of Tim Cantor*”, which reveals the concept of the music video. It is a collaboration of *Imagine Dragons* and Tim

Cantor, a surrealist American artist, who designed their album *Smoke + Mirrors*. As will be seen later on, the video breathes life into paintings and makes them a part of reality. The paintings are of great importance in construing the message of the music video as they help to reveal the main concepts of the song, alongside with painter`s interpretation of their meaning. More than that, after a successful collaboration between the band and the artist, a song gallery was created, where not only the pictures but also Tim Cantor`s interpretation of *Imagine Dragons* songs was provided ("STORY").



Figure 11. Shots at 00:05 (Imagine Dragons, 2015)

The second shot of the music video presents two ladies who are standing in front of the closed door and it one might think that they are about to open it (Figure 11). If to take a closer look at the scene, it is seen that there is another door or arch before the door the girls are standing and there are open curtains just before the closed door. One more important detail here is the iron heart, hanging right above the door. The heart consists of two parts, what is rather symbolical and it draws the audience back to the topic of love. The camera zooms the door and it creates tension on the addresses`.

Once the door is open, one can see the band's frontman and he starts singing "*I am sorry for everything, oh everything I've done*". In terms of visuals, the effect of *mise en abyme* or 'picture in picture' is created. Moreover, this effect goes throughout the whole music video and the audience will be encountering with it over and over again. This technique is of great importance in the video as it helps in building us a story world and helps to connect this world with reality. It also creates an effect that the ladies invite the audience to something personal and hidden.

The next shot is with Daniel Wayne Sermon, band's guitarist. The camera works in the way the musician has moved away from the screen to that extent that one might see a lady who has something like a magic snow globe in her hands and Daniel Wayne Sermon is inside that globe. This combination is rather contradictory as, despite the fact the musician is playing the guitar, he is inside the globe, so space is limited as well as space where the lady is located.

The first lines of the song "*Am I out of touch? // Am I out of my place?*" are rhetorical questions and they are aimed to attract listeners' or viewers' attention. What is more, the lines have their visual manifestation as the picture with the lady. She is getting away in the becomes outside the room where band's frontman is standing and singing the first lines of the song.

This visual representation is of great importance here as it seems to be an answer to the questions asked before: as the picture is behind the window and Dan Reynolds is standing in the house, he seems to be out of the place. Interestingly, space is limited here as well as he is inside the house and the window is closed, so there is no way for him to get in touch with the lady.

The next line "*When I keep saying that I'm looking for a space*" is also well-supported by visual effects. A close-up of space appears on the screen. However, there are some metal gears, cogwheels and screws, that look like details from an old clock.

A great antithesis in lyrics supported by the video is presented during the next few seconds. "*Oh, I'm wishing you're here*" the band's drummer appears in

the space and he is playing the drums on the clock face, but as the lines “*But I'm wishing you're gone*” are communicated, the scene moves away and becomes just a picture or a window view from another room or even from a space shuttle. This steam-punk representation of the space is a perfect fit for a surrealistic world created in the video.

“*Oh, I'm going to mess this up*” and zoom of the picture, where snakes are moving towards man's throat, like predicting the consequences of what can happen if something goes wrong. This combination supports uncertainty expressed in lines before (“*Oh, I'm wishing you're here // But I'm wishing you're gone*”).

However, they stop moving at some point and tension is decreasing. With the line “*Over and over and over again*” the camera works the way the audience sees picture moving back from the screen. This effect turns into one more *mise en abyme* as the picture becomes a part of the scene that was presented at the beginning: a long hall or alley with Dan Reynolds.

The difference between two scenes with this hall is that closer to the beginning of the video, the singer was standing, but the camera worked the way he was getting closer to the screen, and in the current scene he is getting farther and space is extending.

Word ‘*shot*’ is of great importance in the music video. The reason for it is not only the title but the next scene as well. This word goes always with a vivid bit of drum (what creates an effect of a shot). What is for the music video, each time the word ‘*shot*’ is pronounced, the picture changes. It can be pointed out on 1:09 - 1:19 of the music video, where a repetition of ‘*shot*’ reflects in a change of scene.

After a close-up of a bird appears (Figure 12), everything changes. This is twist attract attention to the changes in visual effects and increases the tension in the episode.

Even though the hall with Dan Reynolds is shown for the third time, the camera works the way frontman moves right into it, asking rhetorical questions as it was at the very beginning of the music video. This movement underlines the

phrase “*I’m looking for a way to escape*”. It also makes clear as there is no way to escape as it seems that frontman and the road move is different directions.



Figure 12. *Shots* at 01:27 (*Imagine Dragons*, 2015)

The second chorus is performed in a new setting - field with houses, one of which is burning. The burning house is a visual representation of ‘*everything I’ve done*’. One more visual representation of ‘*shots*’ can be found there as well when a lady with six doves is shown (at 02:02 - 02:13) as the word *shot* is repeated for six times.

The line “*At the roadside we used to know*” is reflected in the picture called “*Hopeless Opus*” (at 2:39), where the man is looking at his past and everything that was happening before. Past is created with the help of *mise en abyme* (the band playing in the red room became a part of a drum and a drum becomes a small circle at the man from the picture looks at). The mask might symbolize either past or present. Most likely, it refers to the experience acquired before. As the man is not in the mask and it is located a few inches from his face “*And there’s always time to change your mind*”, as there is space for a change between past (the mask) and present (the man).

Finally, the last chorus and the *shot* there is visualized with the help of a close-up of the shot from the main gun (there were cases of *shot* visualization

before, but they were not that vivid). After that, the piano solo was introduced and the tempo of the song became faster, which reflected in a faster change of shots and pictures in the footage sequence moving from the screen. Moreover, the shot transforms into a bullet, what is rather symbolic. It dramatizes the line '*I shot everything I loved*' as people are able to kill without affords or special equipment, but with a glance only.

The video has a rather predictable denouement. It video ends with the scene where the girls are standing in front of the open doors and they are, most likely, about to close it. This scene frames the video and, to some extent, it foregrounds the idea that everything was happening only in the head and soul of a lyric hero. There are numerous proofs of it, not only the framing but also limited space in the music video and the man with the birds (that usually symbolize ideas and thoughts), that flew away from his head.

3.4. Pictorial and conceptual metaphors in “I bet my life” (2014)

The name of the song "*I bet my life*" is rather encouraging and promising as it brings certainty that something will happen. From the first seconds, one can see a car on the road. This is an excellent hook as the way is not only the beginning of something momentous, but it also employs some uncertainty and unpredictability about the things that might happen on the road. Alongside with first musical bid, the sound of the car moving on the stony road is heard. More than that, the camera works the way the vehicle moves to the audience, creating an effect that the car is about to reach the destination. However, not everything is as simple and clear as this episode.

The next shot represents a group of people that look like close friends or even family as they are laughing, and most likely, they enjoy spending time together. The rising action happens when Dan Reynolds plays skipping stones, and the audience's attention goes along with the stone, thrown on the surface of the river. The camera follows the stone and brings the action to the opposite bank of

this river. At the moment when a stone is moving, the music sound in the video changes as well.

To be more specific, the sound waves are likely to reflect the waves in the river as it fluctuates from loud to soft. The frequency of flotations is equal so that it creates an effect of flow and ebb. The work of these two effects intensifies the action and draws attention to the setting of the music video. More than that, the river looks like a road in this case so that the audience perceive the shift of settings not only mentally, but visually as well.

On the opposite bank, one can see two boys that are about to fight. Aural effects underline the tension between them. Some noises, very similar to screams, appear in the background, and the boys start their fight. The noises remain fluctuating from soft to loud depending on the action in the video: the sound is soft when the scene is known to the audience and loud at the moments when something new happens.

The setting changes once the song starts, and from the very first line, "*I know I took the path that you would never want for me*" a path is visualized with the help of river, surrounded by mountains, what represents the lyric to the full extent. The second part of this line reveals the conflict between boys (perhaps brothers) as someone made a choice that is opposite the desire of his opponent. To announce that a journey has begun one of the steps into the water, and it becomes clear that he is about to take the wrong path. In terms of music, it becomes more dynamic and rhythmic. In this particular scene the lyrics shapes the tonality and dynamic of the song. They interact and become indivisible, contributing to the plot of music video.

Moreover, the scene with fighting boys was interrupted two times by the scene with the road. This sequence contributed to the idea that someone is going to make a choice. Considering a verbal aspect of the song, it is going to be a wrong path.

However, with the line "*Well, I'm just a slave unto the night*" it becomes clear that the protagonist is forced to choose this way either because of the

circumstances or because of someone else's will. This line also reflects in a plot of a music video. The boy in a blue shirt knocks down a boy in a white shirt to that extent that the punched boy plunges into the water.

The scene with the water is presented in a very unusual way with the help of the camera. The footage is shown both from above the water as well as from underwater. It creates the effect that both worlds (under- and above water) are equal there.

The words “*Now remember when I told you that's the last you'll see of me*” go alongside with the scene when the boy in blue swims away from his opponent. There is a close-up of the opponent's face, and, due to his facial expressions, it becomes clear that he is surprised by such a twist.

Moreover, and he does not agree with the choice of the boy in white. The very first line is repeated once again in such a way that it frames the stanzas, and signals that a chorus (culmination in this case) is on the road. This line does frame not only the stanza but also frames the scenes that were shot above the water.



Figure 13. *I bet my life* at 00:49 (*Imagine Dragons*, 2014)

More than that, the last scene above the water (Figure 13) underlines the contrast between two worlds, two locations, and at least two worldviews. This

division is created with the of the bridge (even though usually the bridge is a metaphor of connection, unity, and community).

After finding a hole in a bridge, the protagonist is falling down the water channel. Falling usually symbioses something groundless and unreasonable as RATIONAL IS UP; EMOTIONAL IS DOWN (Lakoff & Johnson, 2017).

Along the line *"I've been around the world and never in my wildest dreams"* the protagonist finds himself in an underwater house. Despite all the proofs (the bridge, falling through the channel, water, and fishes that can be observed through the window), there is no water inside the house. The atmosphere in the house is quite gloomy as the light is deemed blue due to the fact the house is underwater. The boy looks surprised and lost, judging by his non-verbal behavior. The tone and atmosphere of the scene are emphasized with the noises heard on the background (the same screaming sounds that were at the beginning of the music video when the boys were fighting).

"Would I come running home to you" proves the hypothesis that boys might be brothers or they might have close kinship relationships.



Figure 14. *I bet my life* at 01:32 (*Imagine Dragons*, 2014)

Besides the path, the boy has chosen and the thing that he became alone in an underwater world, *"There's you in everything I do."* The line from the song is supported by the work of the camera (Figure 14) as it looks the way someone

observes the boy in white. It creates an effect that the boy in a blue shirt is following his contestant even after the fight or differences they had.

The boy finds a bed, and he seems to fall asleep before the chorus starts. As SLEEP IS A TAILOR (Lakoff & Johnson, 2017), the settings change once again, and the sequence of settings supports this conceptual metaphor as the boy appears to be a tailor of his own life. It is manifested not only in choosing his path but also in changing the setting he is afraid of. The sleep is working as a mental and physical portal from undesirable to desirable, and the protagonist appears in a boat above the water world.

As the boy hoists up the sails, the chorus begins. Interestingly, there is a dramatic pause in terms of audio accompaniment and moving pictures. The boy is sailing smoothly, and the music becomes very slow and relaxing. The boy in the white shirt looks encouraged and ready to continue his journey as HAPPY IS UP. The camera is immersed in the water so that the audience can see that the actions take place in the world above and that the boy seems to get back home. This assumption is supported by the line *"Don't tell me that I'm wrong // I've walked that road before"*.

It is momentous to mention that the last stanza looks like a confession and repentance. The protagonist is returning home, and this way seems to be the new path he chose. However, based on the lyrics, he seems to regret the things he did and the path he had decided before, *"And left you on your own ...//...That it's left for yesterday"*.

The most powerful idea of the song and music video is expressed in the line, *"Please forgive me for all I've done."* The way how it sounds and how it is reflected in footage creates a strong effect upon the audience. First of all, this line has a strong semantic load as not everyone can understand their mistakes, and, more than that, not everyone asks for forgiveness.

Second of all, the line is powerful in terms of its aural representation as to the words *"for all I've done"* are pronounced with big pauses. This way of pronouncing draws the audiences' attention not only to the words themselves but

also to the manner they are communicated. It also helps to understand that these words are significant, and that is why not easy to pronounce.

Third of all, the visual representation is quite dramatic as the protagonist is almost at the edge of the waterfall. It is not clear what will happen next and whether it will be a happy end or not.

The slow manner of singing the last line of the pre-chorus created a dramatic pause before the upshot. The boy soared into the sky, what hints on the happy end of the video. However, the protagonist starts shattering the boat, and it overturns, throwing back the boy from the sky down in the city right to the crowd of people that are holding him. This scene is like a step away from loneliness.

Once the music ends, the picture overlaps, and right from the crowd, the boy appears to be back in the water. Someone is trying to drag him out of there right after it becomes clear that the protagonist is lost again. The boy in a blue shirt rescues his opponent. As there is no sound in the video, the attention is paid to the action only. Later on, the sounds of the waterfall appear, reminding about the things that have happened before.

The whole music video contains a widely-used conceptual metaphor LIFE IS A JOURNEY (Lakoff & Johnson, 2017). The metaphor goes through the footage, visualizing the key points that are applied by this construction. As is seen, the journey here is presented verbally, visually, and aurally. These modes intertwine and contribute to the general idea of the song, pointing out that life consists of choices we make, the paths we take, and people around. Here the name of the song “*I bet my life*” has a continuation “*for you*” making clear that not only individual matters but also the people who surround them.

3.5. Hidden movies in “On top of the world” (2013)

The *Imagine Dragons`* song “*On top of the world*” is closely connected to Stanley Kubrick`s two films, *The Shining* and *2001: A Space Odyssey*. The music

video appears to be an interpretation of the band's vision of what would happen if musical group *The Beatles* was the one who made a fake moon landing.

Before footage appears, there is a black screen and the sound of small wheels moving on the ground. Another music is of spinning pedals, but one can see neither where that are pedals nor who turns them as there is no motion at the picture, and there is no picture at all. The spinning sound intermingles with the sound that is very similar to the sound people produce while clapping their hands. Finally, the light turns on, and the audience can see a boy riding a small bicycle.

From the very first shot, the connection of Stanley Kubrick's movie and the music video becomes vivid as a boy in a tin-foil hat is riding a small bicycle the same as Danny in the film. The hat is symbolical not only for the baby-world but also for conspiracy theories and pseudo-science. From this moment on, it becomes clear that the events in the music video will be fabricated. If to take a closer look at the bicycle, the superscription 'SNOGARD' can be observed.

Once the boy reaches his destination, he dismounts the bike. After that, he stands in front of a showcase, and superscription on his shirt is also 'SNOGARD.' From first sight, this word looks like a strange name of some brand, but if to read this word back to front, it transforms into the word 'dragons.' That can be a reference to the band's name. Once the boy stands in front of a showcase, one can see the words 'Stan and Brick.' These words could be a shortening from Stanley Kubrick.

The next shot presents what is hidden beyond the showcase. It is an old TV named 'Monolith.' This name refers to *2001: A Space Odyssey* as there were individual machines, built by some unseen space inhabitants. Interestingly, those creatures were famous for intergalactic travel, and this might be a device of great help for them.

Similarly, space scenes from the screen were breathtaking, and they were also like a portal that opened a new dimension for the people of Earth. To be more specific, the scenes from the Apollo moon launch were presented on the screen.

This fact connects the view to *The Shining* as there is an excellent discussion on the Internet that Kubrick fabricated the landing of Apollo.

The song starts with the lines "*If you love somebody // Better tell them why they're here 'cause// They just may run away from you.*" At the beginning of this episode, there is a carpet on the background. The rug is the same as in *The Overlook Hotel* in *2001: A Space Odyssey*. Moreover, the lines are visualized with the help of showing the man (Dan Reynolds) and his family. He kisses wife and daughter alongside with the word love and runs away, like following the instructions from the song.

There is one more reference to the *Space Odyssey* here: the house number is 2001, which is a part of the name of the movie. The other three band member leave their houses as well. Interestingly, the fourth member (the one on the bike) has a gallon 'SNOGARD.'

The chorus starts with the words "*'Cause I'm on top of the world.*" More than that, it is underlined by a loud musical bid. Alongside with the bid, the biker starts his vehicle and gets on the way. On his way, a traffic sign is zooming. The sign says, "Route 237. Scenic Overlook" (*On top of the world* 0:53 *Imagine Dragons*, 2013) appears.

There is a reference to *2001: A Space Odyssey* as the room Danny saw ghosts was numbered 237. So it brings the idea that something mystical might happen. This number has a connection to Apollo 11, as this is the distance to the moon and back (in miles). What about the next sign "Scenic Overlook," it goes back to *The Overlook Hotel* from the same movie.

The second repetition of the first chorus line is illustrated with the help of TV. Here are some shots with Apollo preparing for the start. This episode shows the direct meaning of the collocation 'on the top of the world.'

Meanwhile, the *Imagine Dragons* are going somewhere (Figure 15). Once the words "*Been holding it in for a while*" are pronounced, they hold on for a

while. This effect evokes interest among the audience, and the question where are they going remains open.



Figure 15. *On top of the world* at 01:02 (*Imagine Dragons*, 2013)

The most powerful scene from a multimodal point of view is at 1:34-1:39 of the music video. While chorus "*Cause I'm on top of the world*" a man from NASA turns a button and gives a command to fly up. At that, the shuttle takes off. While the team is going into open space, the whole world is watching them doing that. The sounds from the space shuttle are reproduced in the song. It creates an effect that the audience is also following the heroes on their first trip to the Moon.

All this perfect picture is interrupted by a noise of communication jamming. After that sound, the image on the TV screen disappeared, and the splash screen appeared instead. All the people around TV screens and radios paralyzed, waiting for further development and action. As is seen from Figure 16, the reason for that technical interruption was a boom microphone that has fallen to the head of one astronaut. This epic collapse in the music video was presented with a loud biting noise. Even the music disappears in that very moment to underline footage. One can construe a situation irony from the sequence of events presented in this episode.



Figure 16. *On top of the world* at 02:09 (*Imagine Dragons*, 2013)

Luckily, the situation resolved, and everything went as it should be. The crowd was happy to see the American flag on the moon, however, judging by the boy's facial expressions, he knew that something went wrong, he was neither surprised nor satisfied. Meanwhile, the lyric goes on the line, "*And I know it's hard when you're falling down*" gains scene as, judging by the previous scene, the hero knows what is falling (at least direct meaning of this collocation).

One more interesting with the man who is sitting behind the scenes. Although the situation resolved in the right way, he was not satisfied at all. The man appears to be Stanley Kubrick, which is somewhat symbolic, taking into account all the references to his movies.

Conclusions to Chapter Three

Believer (2017) by *Imagine Dragons* is famous for the music video, hence it evokes a strong emotional response of the audience. The main reason why the music video, as well as the song, is so popular is a great interplay of all the modes that participate in perception of this piece of art. The interaction of auditory, visual, aural, and tactile modes makes a powerful effect on the audience.

A vast interplay of color and scenes, as well as the presence of three men that, as is seen, are the part of one person, keep the audience interested till the very end of the song.

Moreover, the music video is an excellent reflection of the lyrics not only in its direct meaning but in its metaphoric interpretation as well. Besides, the structure of the video is easy to perceive, which makes the audience follow the core ideas. Music effects underline moving pictures, creating a perfect fit for the audience.

Band's hit "*Radioactive*" is famous for the topic it raises and the way the problem is presented in the video. To be precise, the life of ex-prisoner is described, and the narrator is the prisoner himself. Gloomy atmosphere and visual effects create an impression of a post-apocalyptic world.

The video is full of verbal irony supported by visual effects, song, and the outline sounds (the blow of the wind, crows' screams, bits). The footage is rather allegorical as the toys were presented there, and in some scenes, they were on focus. The message is conveyed with the help of the interplay of visual and aural modes.

The video "*Shots*" is based on the visual representation of the lyric. The paintings are of great importance in construing the message of the music video as they help to reveal the main concepts of the song, alongside with painter's interpretation of their meaning. However, there is a slight mismatch between the imaginative world created in the video and the lyrics.

The song has a strong semantic load, and the footage is surrealist, which creates an atmosphere of a fairy-tale. In terms of sound, every shot in the song was not only highlighted visually but also underlined by the musical bid, very similar to the gun's shot.

Pictorial and conceptual metaphors are at work in video "*I bet my life*". Interestingly, they were presented as intertwine of sound, picture, and lyrics. The camera in the video was switching between up and down, so these were two

worlds presented in the video, opposite opinions communicated in the song, loudness, and totality in terms of music.

The main idea of the video "*On top of the world*" is communicated not only through several modes of perception, but also with the help of movies called *The Shining* and *2001: A Space Odyssey*. It is essential to mention that the films do not put the interaction of the modes aside, but highlight the most significant moments in the video.

The parallels between the video, movies and the Beatles in one piece of art does not only evoke the interaction of perceptions. It also refers to the mental processes and activates a cognitive mode in order to draw parallels between events and understand the message of the video.

GENERAL CONCLUSIONS

Multimodal studies are in the focus of present-day linguistics. The paper determines multimodal features of *Imagine Dragons`* pop rock band`s discourse. The notion of multimodality for the research of musical discourse is defined based on a number of studies conducted in this field. At that, four milestones of multimodality shape the study (analysis should focus on several modes of perception, interaction of modes within any context, each mode meets certain communication needs and contributes to the meaning equally) (García, Flores, & Spotti, 2016).

Opportunities and challenges of multimodal analysis are discovered with the help of revealing strong and weak points of different methods of analysis. As is seen, a researcher should use human-generated as well as computer-generated data and interpretation to conduct a valid study (CES, 2000). Based on that, structure and elements of musical discourse is to be revealed in order to conduct a grounded research. The components of musical discourse are: the melody (an auditory mode.), the lyrics (a verbal mode), and a music video (visual mode). What is about structure, it can vary. However, the main point is that all the elements should be equal.

Music videos are outlined as fragments of musical discourse at work of Critical Discourse Analysis (CDA). Professor Lori Burns (UR Institute for Popular Music, 2016) suggests that the interpretive framework for the music videos goes out of three theoretical perspective: genre, CDA, and narrative theory. Basically, this framework facilitates systematic thinking about how the individual domains of words, music and images, that go together in mutually reinforcing ways to be culturally productive and constitutive of the social realm.

Theoretical background contributed to successful analysis of music, lyrics, and video in *Imagine Dragons`* musical discourse. The videos were analyzed with the help of a software called ELAN. As this is a semi-automatic software, the results are based on the data collected and the emotions and feelings the videos

evoke. With the help of data collected in ELAN and observations, the video were analyzed in terms of the interaction of visual, aural and verbal mode. At that, each video revealed its message with the help of these dimension combined in a single message that the audience should receive.

The effective structure of music structure was elaborated with the help of modes interaction and reveled “*Believer*” (2017). The atmosphere in the video is gloomy due to the visual effects (the red light at the beginning signalized about danger and fight), alongside with visual effect, the audience hears a loud vivid beat that serves as an accompaniment to the punch (aural mode) and the word ‘*PAIN*’ (verbal mode).

Apocalyptic motives and allegory in the music video “*Radioactive*” (2012) are outlined with the help of the interplay of visual and aural modes. The song and music video represents a fighting ring and those surroundings that are shown in the video is an allegory to a pressing issue ex-prisoners have gone back to a normal life. Moreover, conveying a message by having a speaker the victim of the issue makes the audience sympathize the prisoners.

Visual symbolism in “*Shots*” (2015) is, to some extent, determined by collaboration of *Imagine Dragons* and Tim Cantor. However, some episodes are not only visualized, but supported with audio and lyrics. The last chorus and the *shot* (in the lyrics) goes together with the help of a close-up of the shot from the main gun. After that, the piano solo was introduced and the tempo of the song became faster, which reflected in a faster change of shots and pictures in the footage sequence moving from the screen. Moreover, the shot transforms into a bullet, what is rather symbolic. It dramatizes the line ‘*I shot everything I loved*’ as people are able to kill without affords or special equipment, but with a glance only.

Pictorial and conceptual metaphors play a crucial rove in music video “*I bet my life*” (2014). The whole music video contains a widely-used conceptual metaphor LIFE IS A JOURNEY (Lakoff & Johnson, 2017). The metaphor goes through the footage, visualizing the key points that are applied by this construction. As is seen, the journey here is presented verbally, visually, and aurally. These

modes intertwine and contribute to the general idea of the song, pointing out that life consists of choices we make, the paths we take, and people around.

The music video "*On top of the world*" (2013) contains allusions to Stanley Kubrick`s two films, *The Shining* and *2001: A Space Odyssey*. It is seen not only from visual dimension, but also from lyrics and song. From the very first shot, the connection of Stanley Kubrick`s movie and the music video becomes vivid as a boy in a tin-foil hat is riding a small bicycle the same as Danny in the film . The hat is symbolical not only for the baby-world but also for conspiracy theories and pseudo-science. From this moment on, it becomes clear that the events in the music video will be fabricated (*The Shining, 2014*).

While chorus "*Cause I'm on top of the world*" a man from NASA turns a button and gives a command to fly up. At that, the shuttle takes off. While the team is going into open space, the whole world is watching them doing that. The sounds from the space shuttle are reproduced in the song. It creates an effect that the audience is also following the heroes on their first trip to the Moon (*2001: A Space Odyssey, 2013*).

RÉSUMÉ

Темою даної магістерської роботи є «Лінгвостилістика дискурсу поп-рок групи *“Imagine Dragons”* у мультимодальному вимірі».

Мультимодальність у лінгвостилістиці почала активно досліджуватися впродовж останнього століття. Наразі це поняття асоціюють не лише з історією інтерпретації тексту, а й з різними формами медійного дискурсу. Саме тому у роботі висвітлено поняття мультимодальності як музичного дискурсу, представлені його основні підвалини, а також переваги і недоліки мультимодального аналізу музичних відео.

Крім того, для успішного аналізу лінгвостилістичного дискурсу групи *“Imagine Dragons”* описана структура і елементи музичного дискурсу для ефективного аналізу музичних відео, музики й тексту.

Робота базується на лінгвістичних засадах мультимодального і музичного дискурсу, які проілюстровані на прикладі відомої поп-рок групи *“Imagine Dragons”*.

Мета роботи — розкрити мультимодальний аспект у лінгвостилістиці дискурсу групи *“Imagine Dragons”*.

Магістерська робота складається з вступу, основної частини, яка включає в себе три розділи, висновків та списку використаної літератури.

У першому розділі дипломної роботи розкриває концепт мультимодальності розкритий в контексті музичного дискурсу. У розділі також представлені переваги і недоліки мультимодального аналізу.

Другий розділ пропонує огляд структури музичного дискурсу, визначає музичні відео як частину музичного дискурсу, характеризує музику, слова (пісні) і музичні відео у мультимодальному вимірі.

У третьому розділі представлений аналіз дискурсу поп-рок групи *“Imagine Dragons”* на базі програми ELAN. А точніше, лінгвостилістичний аналіз мультимодальному вимірі на прикладі музичних відео *“Believer”*

(2017), “*Radioactive*” (2012), “*Shots*” (2015), “*I bet my life*” (2014), “*On top of the world*” (2013).

Загалом магістерська дипломна робота розкриває музичний дискурс у мультимодальному вимірі, а також ілюструє його на лінгвостилістиці дискурсу поп-рок групи “*Imagine Dragons*”.

Ключові слова: Imagine Dragons, multimodal, mode.

REFERENCES

- Amar, R., & Stasko, J. (2005). Knowledge Precepts for Design and Evaluation of Information Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4), 432-442. doi:10.1109/tvcg.2005.63
- Baldry, A. & Thibault, P. (2006), *Multimodal Transcription and Text Analysis*. London: Equinox.
- Baldry, A. P. (2004). 'Phase and Transition Type and Instance: Patterns in Media Texts as Seen Through a Multimodal Concordancer', in K. L. O'Halloran (ed) *Multimodal Discourse Analysis*. London & NY:Continuum, pp. 83-108.
- Blacking, J. (1982). The Structure of Musical Discourse: The Problem of the Song Text. *Yearbook for Traditional Music*, 14, 15-23. doi:10.2307/768068
- Blacking, J. (1982). "The structure of musical discourse: The problem of the song text." *Yearbook for traditional music*, 14, 15-23.
- Blommaert, J. (2005). *Discourse*. Cambridge: Cambridge University Press.
- Bordwell, D. and Thompson, K. (2001). *Film Art. An Introduction*. New York: McGraw-Hill.
- Bradby, B. (2003). "Discourse analysis." In D. Horn, D. Laing, P. Oliver & P. Wicke (Eds.), *Continuum encyclopedia of popular music of the world* (67-70). London, New York: Continuum.
- Bruderer, M. J. (2008). *Perception and Modeling of Segment Boundaries in Popular Music*. PhD thesis, Universiteitsdrukkerij Technische Universiteit Eindhoven
- Bryan, M. (1988). *SGML: An Author's guide to the standard generalized markup language*. New York: Addison-Wesley.

- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.
- CES (Corpus Encoding Standart. (2000). Corpus encoding standatd. Version 1.5. <http://www.cs.vassar.edu/CES>
- Chi, E. H., & Riedl, J. (n.d.). An operator interaction framework for visualization systems. *Proceedings IEEE Symposium on Information Visualization (Cat. No.98TB100258)*. doi:10.1109/infvis.1998.729560
- Craft, B., & Cairns, P. (n.d.). Beyond Guidelines: What Can We Learn from the Visual Information Seeking Mantra? *Ninth International Conference on Information Visualisation (IV05)*. doi:10.1109/iv.2005.28
- Currie, G. (1999) 'Cognitivism', in T. Miller and R. Stam (eds), *A Companion to Film Theory*, Oxford: Basil Blackwell: 105-22.
- Currie, G. (2014). *Image and mind: Film, philosophy and cognitive science*. Cambridge: Cambridge University Press.
- Dahl, S. (2000). The Playing of an Accent ? Preliminary Observations from Temporal and Kinematic Analysis of Percussionists*. *Journal of New Music Research*, 29(3), 225–233. doi: 10.1076/jnmr.29.3.225.3090
- Edwards, G. (2015, March 15). Billboard Cover: Imagine Dragons on Being 'Atypical' Rock Stars, and Singer Dan Reynolds on His Depression Struggles and Conflicts With His Mormon Faith. *Billboard*. doi:<https://www.billboard.com/articles/6472705/imagine-dragons-cover-smoke-and-mirrors-touring-grammys>
- Fairclough, N. (2001). *Language and power* (2nd ed.). Harlow: Pearson Education.
- Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. London: Routledge.

- Foote, J. (n.d.). Automatic audio segmentation using a measure of audio novelty. *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)*. doi: 10.1109/icme.2000.869637
- Foote, J. “Automatic Audio Segmentation Using a Measure of Audio Novelty.” *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)*, doi:10.1109/icme.2000.869637.
- Forceville, C. (1996), *Pictorial Metaphor in Advertising*. London and New York: Routledge.
- Forceville, C. (2002), ‘The conspiracy in *The Comfort of Strangers*: Narration in the novel and the film’. *Language and Literature*, 11, (2), 119–35.
- Forceville, C. (2009). Non-verbal and Multimodal Metaphor in a Cognitivist Framework: Agendas for Research. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal Metaphor* (pp.19-42). Berlin: Mouton de Gruyter.
- García, O., Flores, N., & Spotti, M. (2016). *The Oxford handbook of language and society*. New York: Oxford University Press.
- Gibbons, A. (2010), “‘I contain multitudes’”: Narrative multimodality and the book that bleeds’, in R. Page (ed.), *New Perspectives on Narrative and Multimodality*. New York and London: Routledge, pp. 99–114.
- Gillet, O., & Richard, G. (2008). Transcription and Separation of Drum Signals From Polyphonic Music. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(3), 529–540. doi: 10.1109/tasl.2007.914120
- Gillet, O., & Richard, G. (n.d.). Comparing Audio and Video Segmentations for Music Videos Indexing. *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*. doi: 10.1109/icassp.2006.1661202

- Gillet, O., Essid, S., & Richard, G. (2007). On the Correlation of Automatic Audio and Visual Segmentations of Music Videos. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(3), 347–355. doi: 10.1109/tcsvt.2007.890831
- Gow, J. (1992). Music video as communication: Popular formulas and emerging genres. *Journal of Popular Culture*, 26(2), 41-70.
- Graesser, A., McNamara, D., & Louwrese, M. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*, 36, 193–202.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, J. (1992). Contextualization revisited. In *The Contextualization of language*, ed. P. Auer and A. DiLuzio, 39–53. Amsterdam: John Benjamins.
- Gumperz, J. (2001). Contextualization and ideology in intercultural communication. In *The Culture in communication*, ed. Di Luzio, Günthner and Orletti, 35–53. Amsterdam: John Benjamins.
- Halliday, M.A.K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning* (London: Arnold).
- Imagine Dragons. (2012, October 29). *Imagine Dragons - Radioactive* [Video File]. Retrieved from <https://www.youtube.com/watch?v=ktvTqknDobU>
- Imagine Dragons. (2013, November 13). *Imagine Dragons - On top of the world* [Video File]. Retrieved from <https://www.youtube.com/watch?v=4ht80uzIhNs>
- Imagine Dragons. (2014, December 12). *Imagine Dragons - I bet my life* [Video File]. Retrieved from <https://www.youtube.com/watch?v=4ht80uzIhNs>
- Imagine Dragons. (2015, February 15). *Imagine Dragons - Shots* [Video File]. Retrieved from <https://www.youtube.com/watch?v=qQrgto184Tk>

- Imagine Dragons. (2017, March 7). *Imagine Dragons - Believer* [Video File]. Retrieved from <https://www.youtube.com/watch?v=7wtfhZwyrcc>
- Jankun-Kelly, T., Ma, K., & Gertz, M. (2007). A Model and Framework for Visualization Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 13(2), 357-369. doi:10.1109/tvcg.2007.28
- Jehan, T. (2005). Institute of Technology , Massachusetts.
- Jewitt, C. (Ed.). (2009). *Handbook of Multimodal Analysis*. London: Routledge.
- Joder, C., Essid, S., & Richard, G. (2009). Temporal Integration for Audio Classification With Application to Musical Instrument Classification. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(1), 174–186. doi: 10.1109/tasl.2008.2007613
- Kaplan, E. A. (2016). Rocking Around the Clock. doi: 10.4324/9781315623993
- Kress, G. & van Leeuwen, T. (2001), *Multimodal Discourse – The Modes and Media of Contemporary Communication*. London: Arnold.
- Lakoff, G. & Johnson, M. (2003 [1980]), *Metaphors We Live By*. 2nd edition. Chicago: University of Chicago Press.
- Larson, S. (2012). *Musical forces: Motion, metaphor, and meaning in music*. Bloomington: Indiana University Press.
- Levy, M., & Sandler, M. (2008). Structural Segmentation of Musical Audio by Constrained Clustering. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), 318–326. doi: 10.1109/tasl.2007.910781
- Louwerse, M. M., & Jeuniaux, P. (2009). A computational psycho- linguistic algorithm to measure cohesion in discourse. In J. Renkema (Ed.), *Discourse, of course* (pp. 213–226). Amsterdam: John Benjamins.
- Machin, D. (2007). *Introduction to Multimodal Analysis*. London & New York: Hodder Arnold.

- Mark, G., Carpenter, K., & Kobsa, A. (n.d.). A model of synchronous collaborative information visualization. *Proceedings on Seventh International Conference on Information Visualization, 2003. IV 2003*. doi:10.1109/iv.2003.1218013
- Martin, J., & Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox.
- Martinec, R. (2000b). Rhythm in Multimodal Texts. *Leonardo*, 33 (4), 289-297.
- McDermott, M. (2017, December 22). Ranked: The 50 biggest songs of 2017. *USA Today*. Retrieved from <https://www.usatoday.com/story/life/entertainthis/2017/12/22/ranked-50-biggest-songs-2017/934344001/>
- McEwan, I. (1981), *The Comfort of Strangers*. London: Jonathan Cape.
- Mcfee, B., & Ellis, D. P. (2014). Learning to segment songs with ordinal linear discriminant analysis. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5197–5201. doi: 10.1109/icassp.2014.6854594
- Mcfee, B., Bertin-Mahieux, T., Ellis, D. P., & Lanckriet, G. R. (2012). The million song dataset challenge. *Proceedings of the 21st International Conference Companion on World Wide Web - WWW 12 Companion*, 405–410. doi: 10.1145/2187980.2188222
- Mcfee, B., Nieto, O., Farbood, M. M., & Bello, J. P. (2017). Evaluating Hierarchical Structure in Music Annotations. *Frontiers in Psychology*, 8. doi: 10.3389/fpsyg.2017.01337
- Mcfee, B., Raffel, C., Liang, D., Ellis, D., Mcvicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and Music Signal Analysis in Python. *Proceedings of the 14th Python in Science Conference*. doi: 10.25080/majora-7b98e3ed-003

- McGuinness, K., Gillet, O., O'Connor, N., & Richard, G. (2007) Visual analysis for drum sequence transcription. In *European Signal Processing Conference (Eusipco)*, Poznan, Pologne.
- McIntyre, D. (2008), 'Integrating multimodal analysis and the stylistics of drama: A multimodal perspective on Ian McKellen's *Richard III*'. *Language and Literature*, 17, (4), 309–34.
- McIntyre, H. (2018, April 25). These Were The 10 Most Popular Songs In The World In 2017. *Forbes*. Retrieved from <https://www.forbes.com/sites/hughmcintyre/2018/04/25/these-were-the-10-most-popular-songs-in-the-world-in-2017/#8d24ab41be69>
- McKee, K. and Pardun, C. (2003). Reading the Video: A Qualitative Study of Religious Images in Music Videos. *Journal of Broadcasting & Electronic Media*, 43(1), Winter, 110-122.
- McQuail, D. (2000). *McQuail's Mass Communication Theory* (4th Edn.). Sage: London.
- Mittell, J. (2000). A Cultural Approach to Television Genre Theory. *Cinema Journal*, 40(3), (Spring 2001) 3-24.
- Montoro, R. (2010), 'A multimodal approach to mind style: Semiotic metaphor vs. multimodal conceptual metaphor', in R. Page (ed.), *New Perspectives on Narrative and Multimodality*. New York and London: Routledge, pp. 31–49.
- Moore, A. (2001). "Categorical conventions in music discourse: Style and genre." *Music & letters*, 82(3), 209-223.
- Moore, A. (Ed.). (2003). *Analyzing Popular Music*. Cambridge University Press: Cambridge.

- Murphey, T. (1992). "The discourse of pop songs." *TESOL Quarterly*, 26(4), 770-774.
- Nieto, O., & Bello, J. P. (2014). Music segment similarity using 2D-Fourier Magnitude Coefficients. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi: 10.1109/icassp.2014.6853679
- Nieto, O., & Jehan, T. (2013). Convex non-negative matrix factorization for automatic music structure identification. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 236–240. doi: 10.1109/icassp.2013.6637644
- Nieto, O., Morwaread M., Farbood T. J., & Bello J. P. (2014). Perceptual Analysis of the F-measure for Evaluating Section Boundaries in Music. In *Proc. of the 15th International Society for Music Information Retrieval Conference*. 265–270.
- Nørgaard, N. (2010), 'Multimodality: Extending the stylistic tool kit', in D. McIntyre and B. Busse (eds), *Language and Style*. Houndmills, Basingstoke: Palgrave.
- O'Halloran, K. L. (2011). Multimodal Discourse Analysis. In K. Hyland & B. Paltridge (Eds.), *Companion to Discourse Analysis* (pp. 120-137). London: Continuum.
- O'Toole, M. (2011 [1994]). *The Language of Displayed Art* (2nd ed.). London & New York: Routledge.
- Park, K. S., Kapoor, A., & Leigh, J. (2000). Lessons learned from employing multiple perspectives in a collaborative virtual environment for visualizing scientific data. *Proceedings of the Third International Conference on Collaborative Virtual Environments - CVE 00*. doi:10.1145/351006.351015
- Parodi, G. (ed.) (2008). *Géneros académicos y géneros profesionales. Accesos discursivos para saber y hacer*. Valparaíso: EUV.

- Parodi, G. (ed.) (2010). *Academic and professional discourse genres in Spanish*. Amsterdam: John Benjamins (in press).
- Poh, N., & Kittler, J. (2010). Multimodal Information Fusion. *Multimodal Signal Processing*, 153–169. doi: 10.1016/b978-0-12-374825-6.00017-4
- Schrader, P. (1990), *The Comfort of Strangers*. Skoura Pictures: USA.
- Schwartz, L. and Ratner, B. (2007). *Making Music Videos: Everything You Need to Know from the Best in the Business*. Watson-Guptill: New York, N.Y.
- Scollon, S. W. (2004). Nexus Analysis. doi:10.4324/9780203694343
- Serra, J., Muller, M., Grosche, P., & Arcos, J. L. (2014). Unsupervised Music Structure Annotation by Time Series Structure Features and Segment Similarity. *IEEE Transactions on Multimedia*, 16(5), 1229–1240. doi: 10.1109/tmm.2014.2310701
- Shneiderman, B. (2003). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *The Craft of Information Visualization*, 364-371. doi:10.1016/b978-155860915-0/50046-9
- Simpson, P. & Montgomery, M. (1995), ‘Language, literature and film. The stylistics of Bernard MacLaverly’s *Cal*’, in P. Verdonk and J. J. Weber (eds), *Twentieth-Century Fiction: From Text to Context*. London: Routledge, pp. 138–64.
- Smith, B. A., Tan, S., Podlasov, A., & Ohalloran, K. L. (2011). Analysing multimodality in an interactive digital environment: software as a meta-semiotic tool. *Social Semiotics*, 21(3), 359–380. doi: 10.1080/10350330.2011.564386
- Spence, R. (1999). A framework for navigation. *International Journal of Human-Computer Studies*, 51(5), 919-945. doi:10.1006/ijhc.1999.0265

- Spence, R. (1999). A framework for navigation. *International Journal of Human-Computer Studies*, 51(5), 919-945. doi:10.1006/ijhc.1999.0265
- STORY. (n.d.). Retrieved from <http://timcantor.com/imaginedragons.html>
- Tagg, F. (2003). *Analyzing popular music: Theory, method and practice*. Cambridge, New York: Cambridge University Press.
- Torrent, A., Peracaula, M., Llado, X., Freixenet, J., Sanchez-Sutil, J., Marti, J., & Paredes, J. (2010). Detecting Faint Compact Sources Using Local Features and a Boosting Approach. *2010 20th International Conference on Pattern Recognition*. doi: 10.1109/icpr.2010.1117
- Twilhaar, J. N., & Bogaerde, B. van den. (2016). *Concise lexicon for sign linguistics*. Amsterdam: John Benjamins Publishing Company.
- UR Institute for Popular Music (2016, December, 8). "Music Video Analysis: Genre, Discourse, and Narrative in Music, Words, and Images" - Lori Burns [Video File]. Retrieved from <https://www.youtube.com/watch?v=RVga-QgY5HQ>
- Vernallis, C. (2004). *Experiencing music video: aesthetics and cultural context*. New York: Columbia University Press.
- Vernallis, C. (2004). *Experiencing Music Video: Aesthetics and cultural context*. New York: Columbia University Press.
- Vernallis, C. (2008). Music video, songs, sound: experience, technique and emotion in Eternal Sunshine of the Spotless Mind. *Screen*, 49(3), 277–297. doi: 10.1093/screen/hjn036 (n.d.). Retrieved from
- Wallis, C. (2011). Performing Gender: A Content Analysis of Gender Display in Music Videos. *Sex Roles*, 64(3-4), 160-172.

Weiss, R. J., & Bello, J. P. (2011). Unsupervised Discovery of Temporal Structure in Music. *IEEE Journal of Selected Topics in Signal Processing*, 5(6), 1240–1251. doi: 10.1109/jstsp.2011.2145356

Wertsch, J. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

YouTube Movies (2013, December 12). *2001: A Space Odyssey* [Video File]. Retrieved from <https://www.youtube.com/watch?v=pprmw5C1AAE>

YouTube Movies (2014, March 11). *The Shining* [Video File]. Retrieved from <https://www.youtube.com/watch?v=kxZbmSJZdfI>