LISTENING STRATEGIES FOR MUSICAL BRIGHTNESS CONTOUR: CONTINUOUS RATINGS BY USING SLIDER CONTROL

Henri Kärkkäinen
Master's thesis
Department of Music, Art and Culture Studies
University of Jyväskylä
Spring 2023

UNIVERSITY OF JYVÄSKYLÄ

Faculty Humanities and Social Sciences	Department Department of Music, Art and Culture Studies
Author Henri Kärkkäinen	
Title Listening Strategies for Musical Brightness Contour: Continuous Ratings by Using Slider control	
Subject Musicology	Level Master's Thesis
Month and year March 2023	Number of pages 40

Abstract

Timbre and timbre perception have been notoriously difficult subjects to study and for years, timbre lacked a clear definition. Research in past decades have formulated definition for timbre and a timbre space model of timbre perception according to which different timbre dimensions corresponding to different acoustical features of sound, interplay and creates the perceptual timbre of sound.

Continuous rating studies have been used in music psychology studies as a method for collecting data for example about how subjects perceive different elements of music.

This study explored how listeners might perceive musical brightness contour. The data was gathered by employing continuous rating type of method. Brightness contour means that similarly as a melody forms a contour when its unfolding in time, changing brightness forms a contour when the music unfolds in time. The task of the participants of this study, who were recruited from the community of the University of Jyväskylä, was to continuously rate by using a slider control, how bright they perceived the music. The stimulus that was used were classical music and were considered as ecologically valid. The movements of the slider control were captured by motion capture technology and the data was later analysed mathematically by comparing the movement data to timbre related music feature data retrieved from the music files.

As a result, this explorative study found indications that the timbre space model works in the ecologically valid listening situations as well and that the brightness is perceived multidimensionally according to the interplay of different timbre dimensions. However, indication that timbre dimension corresponding to attack/onset of tones is dominant over other dimensions was found.

Keywords

timbre, continuous ratings, timbre space, brightness, ecologically valid

Depository University of Jyväskylä

Additional information

JYVÄSKYLÄN YLIOPISTO

Tiedekunta Humanistis-yhteiskuntatieteellinen tiedekunta	**Laitos** Musiikin, taiteen ja kulttuurin tutkimuksen laitos
Tekijä Henri Kärkkäinen	
Työn nimi Listening Strategies for Musical Brightness Contour: Continuous Ratings by Using Slider Control	
Oppiaine Musiikkitiede	**Työn laji** Maisterin tutkielma
Aika Maaliskuu 2023	**Sivumäärä** 40

Tiivistelmä

Sointiväri ja sen hahmottaminen ovat olleet vaikeita aiheita tutkia. Sointiväriltä puuttui ylipäätään selkeä määritelmä vuosien ajan. Viime vuosikymmenten tutkimus on muovaillut määritelmän sointivärille ja mallin sointivärin hahmottamiselle, jossa erilaiset sointivärin hahmotusavaruuden ulottuvuudet korreloivat tiettyjen äänen akustisten ominaisuuksien kanssa. Näiden ulottuvuuksien keskinäinen vuorovaikutus muodostaa havaintomme sointiväristä.
Jatkuvaa arviointia on käytetty menetelmänä musiikkipsykologian tutkimuksissa ja tästä on saatu dataa esimerkiksi siitä, miten tutkittavat hahmottavat erilaisia musiikin elementtejä. Tämä tutkimus selvitti, miten musiikin kuuntelijat voisivat hahmottaa musiikin kirkkauden muotoa. Tutkimuksen data kerättiin käyttämällä jatkuvan arvioinnin menetelmää.
Musiikin kirkkauden muoto vastaa ilmiötä, jossa melodialle syntyy muoto, kun se etenee sävelestä toiseen ajan mittaan. Samaan tapaan sointivärin kirkkauden vaihtelulle syntyy muoto, kun kappale etenee ajan myötä.
Tämän tutkimuksen osallistujat rekrytoitiin Jyväkylän yliopiston yhteisöstä ja heidän tehtävänään oli arvioida jatkuvasti, käyttäen liukusäädintä, kuinka kirkkaaksi he kokivat kuulemansa musiikin. Kuuntelutilanteessa käytetty ärsyke oli klassista orkesterimusiikkia, jota pidettiin ekologisesti validina. Liukusäätimen liikkeet tallennettiin liikkeenkaappaus teknologialla ja data analysoitiin matemaattisesti vertaamalla sitä sointiväriin liittyvään, musiikkitiedostoista irrotettuun dataan.
Tuloksena tässä aihetta kartoittavassa tutkimuksessa löydettiin merkkejä siitä, että aiemman tutkimuksen esittämä, sointivärin hahmottamisen malli toimii myös ekologisesti valideissa kuuntelutilanteissa ja että sointiväri ja siihen liittyvä kirkkaus hahmotetaan moniulotteisesti, sointivärin hahmotusavaruuden ulottuvuuksien keskinäisen vuorovaikutuksen mukaan. Ulottuvuus, joka liittyy äänen syttymishetkeen, näyttää tämän tutkimuksen perusteella olevan sointiväriavaruuden muita ulottuvuuksia tärkeämmässä roolissa, sointivärin kirkkauden hahmotuksessa.

Asiasanat sointiväri, jatkuva arviointi, sointiväriavaruus, kirkkaus, ekologinen validiteetti

Säilytyspaikka Jyväskylän yliopisto

Muita tietoja

TABLE OF CONTENTS

1 INTRODUCTION AND RESEARCH QUESTION

When conducting research in the field of music psychology, researchers are often interested in what happens when we listen to music and how we perceive different events and elements of music. Many of the elements of music and how they are perceived are well understood. Most importantly, most of the elements of music has a clear definition for what they are. However, this was not always the case with timbre. First attempts to describe timbre resulted to definitions describing more of what timbre is not rather than to explaining what it is. The last few decades have shown good progress with research and understanding of what timbre and timbre perception is. Early work in the seventies and eighties consisted of theoretical approaches and produced better definitions for timbre and theoretical timbre space and timbre dimension models of timbre perception. Also, acoustic correlates for the perceptual timbre dimensions and space have been found.

Recent approaches to research in the field of music psychology typically includes the paradigm of embodied cognition and experimental settings where the participants are continuously rating some element in the music they are listening to. This is the approach also used in this study which aims to survey what kind of listening strategies listeners apply when they are focusing to brightness contour in music. Listening strategy in this case means that this study is interested of the acoustic correlates of sound that the participants focus on for to perceive the brightness and its changes in music. Brightness is considered as one aspect and experience of a resulted perceptual timbre usually related to sound containing high frequencies and/or high amount of energy on them. Brightness contour is a concept developed in a pilot study in the university of Jyväskylä. As a melody forms a contour when its unfolding in time, similarly, changing brightness forms a contour when the music unfolds in time. For studying the brightness contour and for collecting quantitative data, an experimental setting was designed for the pilot study. This same setting was applied to this study and the data was collected at once to be used in the two studies.

Research has shown that body movement not only express different music related meanings through gestures, but body movement is also part of our sensorimotor perception of music. Composers of classical music have sometimes described sun rise, light, and brightness with the means of musical expression. Even though relatively little is known about timbral processing and about body movements that would express timbral contents or timbre contour of music, brightness contour might be possible to perceive and rate the perceived brightness with the aid of using body movement. During the experiment of the studies, participants were asked to continuously rate the brightness contour by using gestures, and motion capture technology were used to capture those movements. The data was mathematically analysed to see if correlation between features of music and the motion capture data existed.

Embodied cognition, gestures, and timbre research literature form the theoretical framework for this study. It is assumed that when participants rate the perceived brightness in music, the gestures they make reflect the activity of listening and carry meaning that can be interpreted by the researcher. As will be reviewed in detail later in the literature review section, expressive intentions or meaning can be recognised from gestures and body language which is based on embodied cognition processes.

As mentioned earlier, timbre is known to be difficult to study and as will be revied in detail in literature review section, timbre is a multidimensional perceptual attribute of sound. Studying brightness contour in music is a way to gain access to timbre perception which would be otherwise difficult to study. Music that was used in the experiment of the studies, as stimuli, were real classical music compositions by Schoenberg, Ravel, Strauss, and Rommiteli. This makes this study ecologically valid. In comparison, especially the earlier studies of timbre have used for example artificial and single instrumental tones as stimuli which participants then rated (Grey, 1977).

The research question in this study is: What types of listening strategies, does listeners apply, when focusing on musical brightness contour? The purpose is to find out by which acoustical feature of sound, participants perceive and rate the brightness in music. Spectral content meaning the sounds overtone series, meaning the series of tones that the sound is comprised of (Killick 2021), and events such as spectral fluctuation and changes in energy distribution across the series (Lartillot 2019) is considered most promising, so the hypothesis is that spectral content of the music will be the main acoustical feature that participants focus on during the experiments.

2 LITERATURE REVIEW

The perception and rating of musical brightness contour is a complicated phenomenon which touches topics of many areas of research inside the field of music psychology. Most important matters to understand are the definition of timbre, the multidimensional nature of timbre perception, and the role of embodied cognition in the timbre perception processes.

2.1 Definition of timbre

In the past, American Standards Associations used to define timbre as a property of sound by which a listener can tell the difference between sounds with similar pitch and loudness. This definition is however problematic, and has met criticism: First, it makes timbre an ill-defined "wastebasket term" for everything that sound is except for its pitch and loudness. Second, this definition implies that timbre is a property of a pitched instruments and sounds only. (Bregman, 1990, p. 92.), (Dolan & Rehding, 2021.) Probably many would agree that different drums of a drum set, and different cymbals of drum set can have very different and rich timbres.
Third, the definition of timbre by American Standards Associations also suggests that timbre is an automatic perceptual result of certain acoustic input. It seems however that timbre is to some degree created by our cognitive processes. According to Wallmark and Kendall, timbre exists at the confluence of physical properties of sound and perceptual result of cognitive processes. (Wallmark & Kendall, 2021). This creates an odd paradox since acoustical properties of sound (timbre in one sense) influences cognitive processes but cognitive processing influences timbre. (Bregman, 1990, pp. 488-489.)
Today, psychoacoustic studies have revealed that timbre is a multidimensional perceptual attribute of a complex tones, which means that timbre is better described and

defined as a collection of dimensions constructing a timbre space. Each timbre dimensions corresponds to a particular acoustic parameter. (Caclin et al., 2008.) Timbre is an emergent quality and there is no single physical measure of sound stimulus that can be correlated with perceptions of timbral identity (Fink, Wallmark & Latour 2018). Wallmark and Kendall have given timbre an operational definition which expresses the nature of the variable in terms of how it can be measured. "Timbre is that attribute of auditory sensation that listeners react to by providing a certain kind of behavioral and/or physiological response". (Wallmark & Kendall, 2021.)

2.2 Timbre and physical properties of sound

While timbre is a multidimensional perceptual attribute of sound, it is often referred in everyday language to as tone quality or tone colour. Different adjectives are often used while trying to describe a tone colour or a quality of timbre. Adjectives like rough, bright, warm, or cold for example, are often used.
Differences in timbre tend to involve three acoustical properties of sound: attack/onset of tones, overtone series, and noise. Attack describes how abrupt is the onset of sound, meaning, does it begin with full volume or is there a momentary fade-in. Overtone series describes the number of different frequencies from which a pitch is comprised of. (Killick, 2021.) These three properties of sound are the acoustical corresponds to the perceived timbre dimensions. Timbre dimensions constitutes of two main dimensions, spectral dimension, and temporal dimension, which can be divided into more specific dimensions. (Grey, 1977), (Fink, Wallmark & Latour, 2018).
The description of factors behind differences of timbre (in acoustical sense) by Killick is however only good if one is studying single musical tones or sounds with short duration. When studying musical brightness contour with ecologically valid stimulus, timbre changes during time. There are number of acoustical parameters which affect to the changes in timbre, and Caclin et al. (2008) offers terms that support temporal aspect of timbre and how sound changes over time: attack time, spectral centroid, and spectral fine structure. For example, spectral centroid refers to overtones but in a larger time scale. It tells statistically about spectral distribution, to what frequency range are the overtones centred.

2.3 Timbre dimensions, timbre space, and timbre perception

One of the first attempts to study the timbre dimensions and the timbre space, and its acoustical correlates, includes a study by John M. Grey (1977), from which it's observable that understanding about timbre started shifting from the American Standards Associations view of timbre towards a multidimensional view of timbre. In the multidimensional scaling study by John M. Grey, multidimensional perceptual scaling experiment was used as a method of gathering data and to construct knowledge about timbre dimensions and timbre space. In the experiment, participants rated similarities between instrumental tones in multiple different scales, as a distance of similarity, between sounds. Based on the answers, different instrumental tones were placed on a multidimensional perceptual scale constituting of three dimensions. These three perceptual dimensions were found to correlate with acoustical features of sounds. One dimension was found to correlate with spectral energy distribution, one with attack and the onsets of the tones, and one with temporal patterns of the tones. (Grey, 1977.)

The concept of timbre space with its dimensions is a model of timbre perception. Our experience of timbre will position into the space according to the interplay between the (subjectively weighted) dimensions. The timbre space varies across subjects, as according to McAdams and Giordano (2009), weight of different factors varies according to which properties of sound a listener is focusing into. The timbre space models integrate individual and class differences as weighting factors on the different dimensions and the set of specificities. For example, some listeners might pay more attention to spectral properties and ignore temporal aspects, and vice versa. (McAdams & Giordano, 2009.)

The perceptual experience of timbre comes from the emergent, interactive properties of the vibration pattern. We identify the different sound sources by listening to the acoustical properties of sound which results from the sound production process. (Handel, 1995.) Subjective differences about the weights of timbre dimensions and on what the listeners focus on while listening about the sound production process, results to different experiences about the timbre.

Works from psychoacoustics field such as the studies about auditory stream segregation by Bregman (1990), and studies about the role of sound production (and sound production gestures) for timbre has had an important role in complementing the theory about multidimensional nature of timbre perception.

2.3.1 Timbre and auditory stream segregation

The auditory scene analysis is a concept mainly formed by Albert S. Bregman, which considers psychological processes that are involved when we hear sounds: how sounds are identified and segregated from each other, and how the grouping of different auditory streams works. In the book, auditory scene analysis (Bregman, 1990), the concept is formed, and the role of timbre is discussed.

Järveläinen (2010), has summarised how timbre can affect to the auditory scene analysis, based on the concepts introduced by Bregman. According to the concepts, hearing separates different sound sources in time and frequency domains. Similar sounds, heard through time, will be grouped into an auditory stream. (Järveläinen, 2010.) Timbre can influence the forming of an auditory stream, but the weight of timbre is not as high for auditory stream formation as of the weight of other factors such as melody, for example (fundamental frequencies/pitch), heard at the same time. Timbre can affect for example if the timbre of notes in a melody are very different. This makes it hard to follow the melody and two auditory streams are formed instead of one. Frequency components (the overtone series of sound) affects to the identification of the sound source. (Järveläinen, 2010.)
Uniform movement of the frequency components of sound can influence and strengthen the formation of an auditory stream. For example, the soloist (instrumental or singer) can use vibrato to separate him/herself from the sound of the rest of the orchestra. Uniform movement of the frequency components can also aid timbre perception since it is easier for hearing to group the partials to belong to the same group. Changes in the time domain can also aid timbre perception. The most important aspect is the simultaneous attack and decay of the harmonic components. (Järveläinen, 2010.)
Timbre can aid sound source identification via auditory stream segregation as is mentioned by McAdams and Giordano (2009). According to them, an important way by which timbre contributes to the organisation of musical structure is related to the fact that listeners tend to perceptually group together sound events that arise from the same sound source. Both spectral and temporal aspects of timbre affect to the auditory stream segregation. (McAdams & Giordano, 2009.)

2.4 Embodied cognition paradigm and musical gestures

Embodied cognition is a paradigm of psychology which consider that cognition is deeply dependent upon features of the physical body of an agent. Aspects of the body beyond the brain play a significant causal or physically constitutive role in cognitive processing. (Wilson & Foglia, 2015.)
Next three theoretical assumptions, by Cowart (2021), describe few of the important qualities of embodied cognition: 1) cognition is goal-directed and organism act with its environment, 2) the form of agent's body determines and constrains the type of cognition, 3) cognition is constructive. The sensorimotor apparatus of an organism

determines the way the organism will experience the world and makes it possible to interact with environment (Cowart, 2021).

The experiences and the type of body that an agent has, contributes to the contents of the cognition. Larry Shapiro (2007) mention Lakoff and Johnson's work on concepts. People often make a use of concepts and metaphors in their reasoning. Basic concepts stem directly from the type of body human beings possess and the way this type of body interacts with the environment. Among basic concepts are spatial ones like up, down, front, back etc. For example, basic concept of up can give metaphorical meaning to concept of happy: I'm feeling up, my spirits rose or you're high in spirits. (Shapiro, 2007.)

When it comes to music, metaphors of movement give meaning and understanding over different elements and events in music. We can hear melody move, rhythm move, and harmonic progression move. Music can be understood as a moving landscape where details in the landscape represent different events in music. Different metaphors of time also relate to the listening of music. (Johnson & Larson, 2003.)

In this regard, some metaphors of movement, or actual movements or rather gestures, might help experiencing and to perceive, and to give meaning for the musical brightness contour. This is viewed in more detail in next sections.

2.4.1 Musical gestures and meaning

Not only do people understand music as movement but they also move to it and make various gestures. A straightforward definition of gesture is that it is movement of part of the body. While studying movement made to music however, more specific definition is often needed. When speaking about the musical activity of musicians, the involved embodiment is called gestures rather than movements. The notion of gesture blurs the distinction between movement and meaning. Movement denotes physical displacement of an object in space while meaning denotes mental activation of an experience. (Jensenius, Wanderley, Godøy & Leman, 2010.)

Musical gesture is movement that expresses an idea or meaning and is part of our sensorimotor perception of music. Gestures can be made for different reasons: to produce sound and to control a musical instrument or to communicate with other musicians and audience. Gestures can also be movements that accompany or express the activity of listening, such as tapping along the beat or dancing. (Leman & Godøy, 2010.)

Musicians body language and sound-producing gestures (gestures used for sound production and control of an instrument) can communicate a certain emotional or expressive state as indicated by various studies.

Dahl et al. (2010) give an example of a Johansson (1973) study, where musicians were instructed to play a same piece of music with three different expressive intentions: deadpan, projected, and exaggerated. The performances were then rated by the observers in scale of expressiveness. They were about equally successful in identifying the expressive intent regardless of whether they were allowed only to listen, only watch, or both watch and listen.

Dahl & Friberg (2007) has conducted a study where participants watched and rated silent video clips of musicians performing the emotional intentions happy, sad, angry, and fearful. Except for fear, other emotional intentions were successfully identified. Different type of movements was used by the musicians for different emotional intensions. For example, happy playing was characterised by fast, jerky, and large movements whereas sad playing was characterised by slow, smooth, and small movements. Anger and happiness were sometimes confused by participants, and it indicates that those two expressions have some features in common.

As aforementioned studies suggest, meaning and intention can be communicated through gestures and perceived from them. Not only are gestures important for communicating meaning, but they have significant role in embodied cognition and perception of music.

2.5 The role of sound production process for timbre

In the theory of sound color (Slawson, 1980), a primary basis of how certain sounds are produced, is defined by formulating a source/filter model of sound production. According to the model, sound is produced when an object is excited by mechanical energy and the object in some manner changes the excitation. The mechanical excitation is called source and the object filter. Both the source and the filter have quite distinct and independent characteristics: The source can affect the intensity and timing of sounds. The pitch of sound and spectral characteristics results from when the filter is excited by the source. The source and the filter are considered independent and sound color (timbre in a sense) is associated with the filter. (Slawson, 1980, pp. 22-24.) The theory considers sound color as a multidimensional attribute of sound and argues that those dimensions include both physical and psychological dimensions (perceptual timbre dimensions and the acoustical correlates for them).

2.5.1 Sound-producing gestures and timbre

An important part of timbre perception is listening to the actions that was required to produce the sound. When considering the source-filter model by Slawson, it seems that the acoustic properties underlying timbre comes from the sound producing processes that excite the source and the physical material of the instrument (filter). The changes in the source-filter combination correlates to the changes in the perceptual timbre.

Sound-producing gestures are movements that are used while playing an instrument to produce sound and are different for different instruments. Godøy (2010) have formed three categories to describe sound-producing gestures distinct in terms of biomechanics and in motor control: iterative, impulsive, and sustained.
Iterative sound-producing gestures means rapid repetition of small movements which fuse into a single gesture. These gestures can be associated with the playing of drums, where the stick bounces back from the surface of drum. Iterative movement could also be rapid shaking movement of a string instrument tremolo.
Impulsive sound-producing gestures means discontinuous movements such as hitting, kicking, rapid stroking, or bowing. (Godøy, 2010.) Impulsive sound-producing gestures are used with piano and plucked string instruments for example.
Sustained sound-production gestures mean continuous movement such as in continuous bowing or blowing (Godøy, 2010).
The amount of effort that was used for gestures is an important part of expression. Godøy (2010) gives an example of a scene in Charlie Chaplin movie, where a barber makes shaving movements which mimic the sound events in music (Brahms Hungarian dance number 5). Gestures with similar sensations of effort may be generalized into schemata. These schemata may be projected onto new sounds.
The amount of effort that was used for a gesture is an important factor in music perception and perception of meaning as indicated by Dahl & Friberg (2007). Lhommet & Marsella (2015) have also found similar results, the amount of effort that is used for gestures communicate different emotional intentions: Happy expression was characterised by fast, jerky, and large movements whereas sad expression was characterised by slow, smooth, and small movements.

2.6 Timbre perception and emotions

In addition to multidimensional nature of timbre perception where sound production process and sound-producing gestures has important role for perception, there is an interesting way of cognition to couple emotions to movement. According to Ferrer (2021), individuals perceive, differentiate and experience emotions with various timbres. Wallmark, Deblieck & Iacoboni (2018) have found out with fMRI scans that there is a motor component to timbre processing.
Wallmark (2014) in his doctoral thesis states following: timbre is directly connected to emotions and that timbre is not understood only by feeling but with the feeling body. Timbre perception is fundamentally motor mimetic, listening to musical timbre demands that we wake up a bodily attitude consistent with the corporeal articulation of

each particular timbral quality. The key to understanding timbre's connection to emotion is the bodily motion it requires, both to produce and to perceive. (Wallmark, 2014.) The sound producing gestures, bodily motion, that is required to produce a sound, seems to have a dominant role in timbre perception. Fink, Wallmark, and Latour demonstrates this in the context of playing electric guitar, where vast possibilities to alter the timbre with different amplifiers, pedals (effects) and so on exists. They point out that it is ultimately the player who makes the real difference to the timbre: no matter the gear, it is always possible to identify a "sound" of certain player for example. (Fink, Wallmark & Latour, 2018.)

In another words, it appears that when we listen to music, one of the ways we perceive timbre is as movements that was required to make the sound. As mentioned earlier, people can perceive different emotional intentions from recordings. There seems to be support for the idea according to which different music related meanings, and timbre, can be understood and perceived by embodied cognition processes.

2.7 Gestures and musical brightness contour in classical music

When we think of classical music, musicians can use different type of moves to express different meanings. They will likely make a use of different playing techniques and styles. When reflecting to Dahl & Friberg (2007), if happiness is expressed by fast, jerky, and large movements, that would mean that in the movements has a lot of energy, compared to sadness which was expressed by slow, smooth, and small movements. According to Friberg (2006), different energy dimensions in so called kinematics-energy space, which is used to describe expressivity of music, are associated with musical parameters such as dynamics and articulation. High energy in performance would mean staccato articulation for example.

If we think of bowed instruments for example, the bowing force, speed, (both influenced by the amount of effort used to make the sound-producing gesture) and position all contributes to timbre of the instrument (Halmrast, Guettler, Bader & Godøy, 2010). Players actions can influence to attack segment of sound envelope, which has particular importance in timbre. If attack is rapid, the sound will be perceived as "percussive" and if it is slow, the sound will be perceived as "bowed" like. Although the attack segments are at the beginning of the sound, these attack features are remembered and influence the rest of our perception of sound event. (Halmrast et al., 2010.)

2.8 Summary of the literature review

Previous study has shown that body and environment beyond the brain has a causal and constitutional role in human cognition. Our sensorimotor system has an important role in cognition especially when it comes to music. We use wide range of different gestures to perceive music and its elements and structures, and to play musical instruments. Since music is a form of communicating different meanings, gestures are used for the communication of those meanings. Composers of music, and musicians alike, describe different meanings with music and those meanings can be recognized and interpreted to some extent from the musicians' gestures.

When we express different meanings, we use various gestures that differ in energy or in effort, and size and intensity. Different emotions for example were characterised by fast, jerky, and large movements (happiness) and slow, smooth, and small movements (sadness).

Timbre is a multidimensional attribute of a complex tones and is defined as a collection of dimensions constructing a timbre space. Each timbre dimensions corresponds to a particular acoustic parameter which are: attack time, spectral centroid, and spectral fine structure. Timbre helps us to recognise different sound sources and it seems that there is a motor component to the timbre processing. When we listen to different sounds, we try to imagine the cause of the sound and the movements that was required to make that sound.

However, according to Handel (1995), no predominant cue (certain acoustic input) uniquely determines sound source identification and timbre. Any single cue will provide some level of identification performance, and combinations of cues usually will produce better performance than one. Moreover, the effectiveness of any cue will vary across contexts. The cues that determine timbre quality are interdependent because all are determined by the method of sound production and the physical construction of the instrument. For this reason, acoustic properties cut across the traditional segmentations of the sound (e.g., attack, steady state, decay) and can lead to the same perception using just part of the sound. The cues for timbre also depend on context: the duration, intensity, and frequency of the notes, the set of comparison sounds, the task, and the experience of listeners all determine the outcomes. (Handel, 1995)

No known acoustic invariants can be said to underlie timbre. Timbre is not correlated with a simple fixed acoustical property. It is inherently multidimensional and emergent, coming out of the interdependent but somewhat uncorrelated parts of the signal. (Handel, 1995.)

3 CURRENT STUDY

The purpose of this study is to find out, how listeners approach the task of listening and rating of the musical brightness contour. The literature referred previously implicates, that ideas and meanings can be communicated trough gestures. In that regard, it is reasonable to assume that it is possible to use a method where participants rate and express, by using gestures, how they perceive the musical brightness contour. As body and its sensorimotor system contributes to cognition, it also contributes to the perception of timbre and the musical brightness contour. Based on the literature, we perceive timbre by multiple acoustical phenomena which correlate with our subjective timbre dimensions. Those dimensions interact based on our experiences of music and on our body and sensory-motor system. Our experiences of a timbre vary since the perceived timbre will position into the timbre space differently between subjects. As timbre is a multidimensional perceptual attribute, it is necessary, while conducting research, to view all the dimensions and the acoustical correlates of those dimensions. The research question in this study is: What types of listening strategies listeners apply when focusing on musical brightness contour? The purpose is to find out by which acoustical feature participants perceive and rate the brightness in music.
Hypothesis is that participants will be found to listening to the changes in spectrum of the music since attributes of spectrum has been found to correlate with perceived brightness. (Bregman, 1990, pp. 96-97).

4 METHODOLOGY

To investigate how musical brightness contour is perceived by listeners, an experimental setting was designed for the pilot study, which included continuous rating tasks for the participants. This study shared the same setting and the data that was collected. However, this study has its own research question and its own analysis of the data. The participants listened to music and gave the ratings by continuously moving a slider control. Motion capture technology was used to capture the movements of the slider and to collect the data. The experiment had also a second part where the participants moved freely and expressed their brightness ratings using their whole body. However, that data is not analysed in this study as the focus is on the slider control data.

Quantitative approach was chosen for this study as it makes possible to gather objective and accurate, numerical data of the attributes of sound, and of the movements of the participants. The experimental setting that was used allowed objective and non-restrictive way to observe the participants moves while allowing participants to fully engage and concentrate on the tasks. Quantitative approach also enables testing the hypothesis and analysis of the data by using statistical tests of probability (Williamon et. al., 2021, pp. 37-38). By using these statistical methods, reliable conclusions can be made whether association between variables existed.

The data was analysed with software called MATLAB, which enabled the statistical tests and the retrieval of the relevant musical information form the music used as a stimulus.

4.1 Participants

Total of five participants, two male and three female, was recruited for the experiment, from community of the university of Jyväskylä. The recruitment of the participants was done by the researchers of the pilot study.
Simple and understandable instructions for the participants was created for the experiment, since as pointed out by Williamon et al., potential threats to reliability include participant error which may be made because questions or instructions have been misunderstood (Williamon et. al., 2021, p. 42).
After the tasks, participants were asked to give an interview where they could share their experiences of doing the tasks. This was for the purposes of the pilot study and the interviews are not analysed as part of this study.
During the tasks, participants wore a motion capture suit and hat with total of thirty-two reflective markers attached on the head, torso, hands, and legs of the participants. Four of the markers were attached on the hat, the rest on the suit. The slider control-device that the participants operated, was placed on a table and the participants sat by the table while doing the task. The table had two reflective markers on it and the slider control had one marker on it.
As the sample of participants is small, the results of this study cannot be generalized, but merely used in designing future studies.

4.2 Experimental strategy

The experiment consisted of two parts. In the first part, participants were asked to continuously rate the musical brightness contour by moving a slider control. The slider control was bidirectional (in a unidimensional scale) device placed on a table. The slider control could be moved to forward to indicate brighter timbre, or backwards to indicate less bright timbre. By default, the slider control was positioned in the middle of its travel at the beginning of experiment. This "zero point" represented not either bright or dark timbre. In the second part of the experiment, participants were also asked to continuously rate the musical brightness contour but this time they were asked to use their whole body to make gestures and to reflect the perceived brightness. The stimulus consisted of four, ecologically valid, classical compositions by Arnold Schoenberg, Maurice Ravel, Richard Strauss, and Fausto Rommiteli. The compositions were: Gürrelieder, Daphnis et Cloé-part3 sunrise, Also Sprach Zarathustra, and Flowing Down too Slow. Ecological validity concerns the situation in which the phenomenon of interest is being studied while trying to retain its "real-life" context (Williamon et. al., 2021, p. 42). The brightness perception was wanted to study with music that could be listened in everyday contexts like home, concerts etc. Ecologically valid stimulus also triggers “natural” cognitive processes compared to artificial listening task stimulus.

4.3 Apparatus

Motion capture technology were used to record participants and slider controls movements for later analysis in MATLAB. Motion capture technology is a valid meter for measuring participants continuous ratings of musical brightness contour. (1) The technology does not constrain participants movement or doesn’t set any other restrictions for them (There is of course limits set by the space in which participants move). (2) The motion capture technology measures exactly what must be measured, and those measurements can be used in analysis to answer the research question of this study. The motion capture is also reliable since it measures participants movements the same way any time, regardless of who is using the technology or who the participants are. When sample rate is high enough and when the cameras are calibrated prior to recording, reliability is high.
Total of thirty-five markers were tracked by fifteen infrared cameras. Thirty-two out of thirty-five markers were placed on the head, torso, hands, and legs of the participants. Remaining three markers were tracked in the first part of the experiment where one marker was placed on the slider control and two on the table on which the slider control was placed. (The data of the slider control marker was used on the analysis). The tracked movements were recorded by Qualisys-software with sample rate of 120Hz. Fifteen infrared cameras ensured that all the reflective markers that participants wore were “seen” most of the time by at least one camera. The sample rate of 120Hz ensured that all the even tiniest movements are captured. Music was played through four Genelec-loudspeakers and two subwoofers for surround-sound and good low frequency reproduction.

4.4 Procedure

The recruited participants were asked to come by the motion capture laboratory one at the time. First, a participant was asked to put on the motion capture suit. Once the participant had the suit on and felt comfortable in it and that the size of the suit was right, researchers carefully placed the reflective markers on the suit by following the instructions provided by manufacturer of the motion capture system. After the preparation of the suit, the participant was asked to sit by the table with the slider control on it, and the instructions for the task was provided.

After the participant confirmed that he/she understood the instructions, the recording was started. After the tasks, the participant was asked to give the interview. This procedure was repeated until the data was gathered with all five participants.

5 DATA ANALYSIS

The numerical data of motion capture and music files was analysed by using a software called MATLAB. Timbre related features of the music were retrieved from the audio files by using Mir-toolbox. Mir-toolbox, developed by the university of Jyväskylä, is a library to MATLAB and it contains functions for various analytical needs. Mir-toolbox was used for example to calculate spectral features such as spectral flux from the music files. MATLAB-functions outside Mir-toolbox were used to plot the movement and music feature data and to calculate correlations between the two.
The three timbre dimensions constructing the timbre space, according to Grey (1977), were dimensions of spectral energy distribution, attack and the onsets of the tones, and temporal patterns of the tones. (Grey, 1977.)
As musical brightness contour unfolds over a long period of time, suitable functions from Mir-toolbox was used to calculate musical features, describing the acoustic correlates of the three timbre dimensions over the whole duration of a music piece.

5.1 Movement data

As the participants moved a slider control to continuously rate and express their perception of brightness in music, the movements of the slider control were captured using motion capture technology. The recordings were pre-processed for data analysis in Qualisys software, meaning that the reflective markers were labelled and then the data was exported from the software into a matrix containing columns from each variable (marker that were tracked) and rows for each observation. Data of a single variable such as the slider control can be extracted from the matrix and is on a vector form. The values of the vector represent observations of the slider position at given time. The slider movements were captured with 120Hz resolution, meaning that 120 observations were recorder each second.

The following figure, containing plots of the movement data, illustrates each five participants brightness ratings in all four listening tasks:

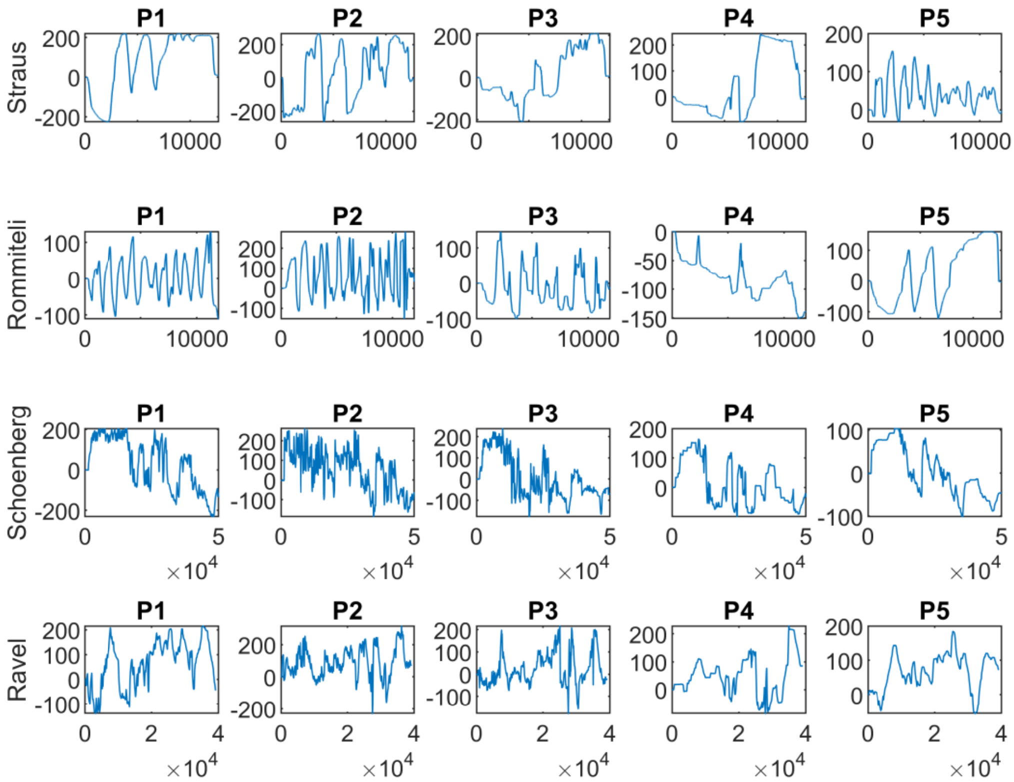

Figure 1 slider data (brightness ratings)

Note that the movement of slider starts at the center of the y-axis since the slider control were a bi-directional device positioned at the center by default. Rows of plots represents the listening tasks: Straus, Rommiteli, Schoenberg, and Ravel.

5.2 Music feature data

Mir-toolbox by the University of Jyväskylä, is a collection of functions that can be used inside MATLAB, to extract music features typical for music information retrieval (Lartillot, 2019, p. 7). This collection of functions is added to the MATLAB's library, and the functions are used similarly as functions original to the MATLAB.
The music features that are extracted, varies from research to another. In this work, the focus is on timbre related music features. Acoustic properties of sound which

correlates with perceived timbre are spectral properties of sound and attack/onset properties of sound (Grey, 1977). Three functions from the mir-toolbox that were selected for analysis are, mirflux, mirbrightness, and mirenvelope.

5.2.1 Mirflux-function

Spectral flux can be calculated by using mirflux-function, meaning the distance of spectrum between each successive frames (Lartillot, 2019, p. 58). The values returned by the function describes as distance how much the spectral content of a frame differ from previous frame. It is important to note that the function does not describe to which direction the perceptual or acoustical change occurred, for example did the perceptual timbre changed from brighter to darker timbre or did the overtones shift to higher frequencies.

The following figure illustrates the spectral flux data from all the music files used in the listening tasks. 120 Hz sample rate were used to make the data comparable with motion capture data. The spectral flux data describes the temporal patterns of the tones.

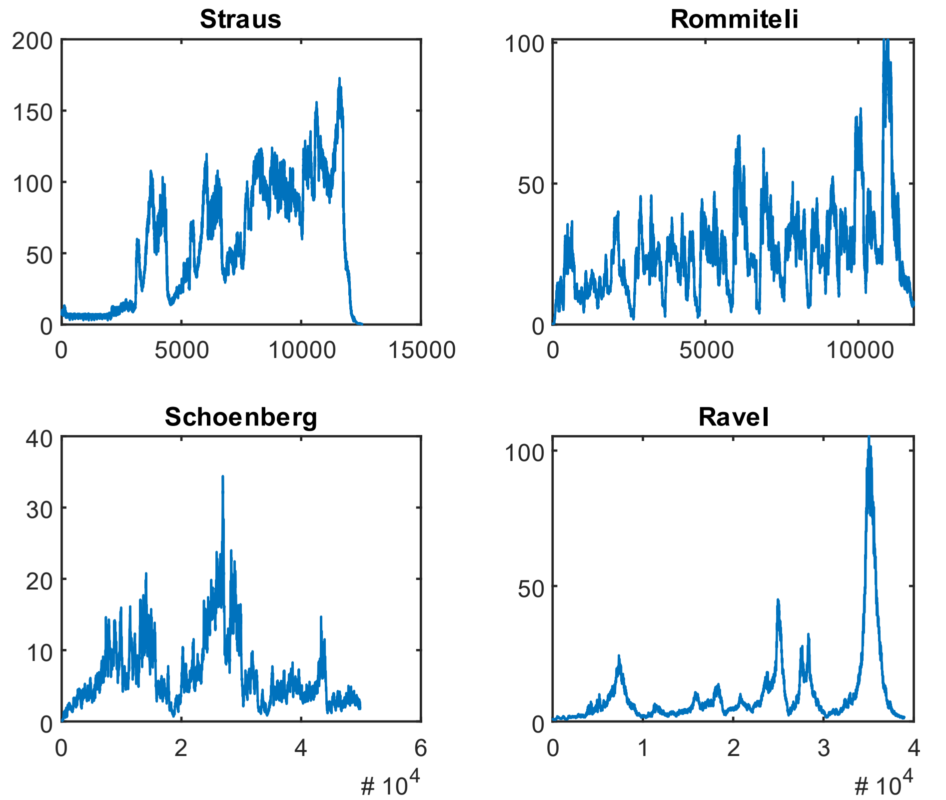

Figure 2 spectral flux data

5.2.2 Mirenvelope-function

Amplitude envelope of a music file can be calculated by using the mirenvelope function. Attack/onset-portion of sound from every single instrument or event in music would be impossible to calculate but results of all those events contributes to the global amplitude of the sound. With mirenvelope function, amplitude envelope is calculated from an audio waveform, and plotting the result shows the global outer shape of the signal (Lartillot, 2019, p. 30). A value zero represents complete silence in the audio and values getting higher represents the sound getting louder.
The following figure illustrates the amplitude envelope data from all the music files. 120 Hz sample rate were used in the calculation. The mirenvelope data describes the global result (a sum) of amplitude to which attack/onset parts of the individual sounds (tones) contributes.

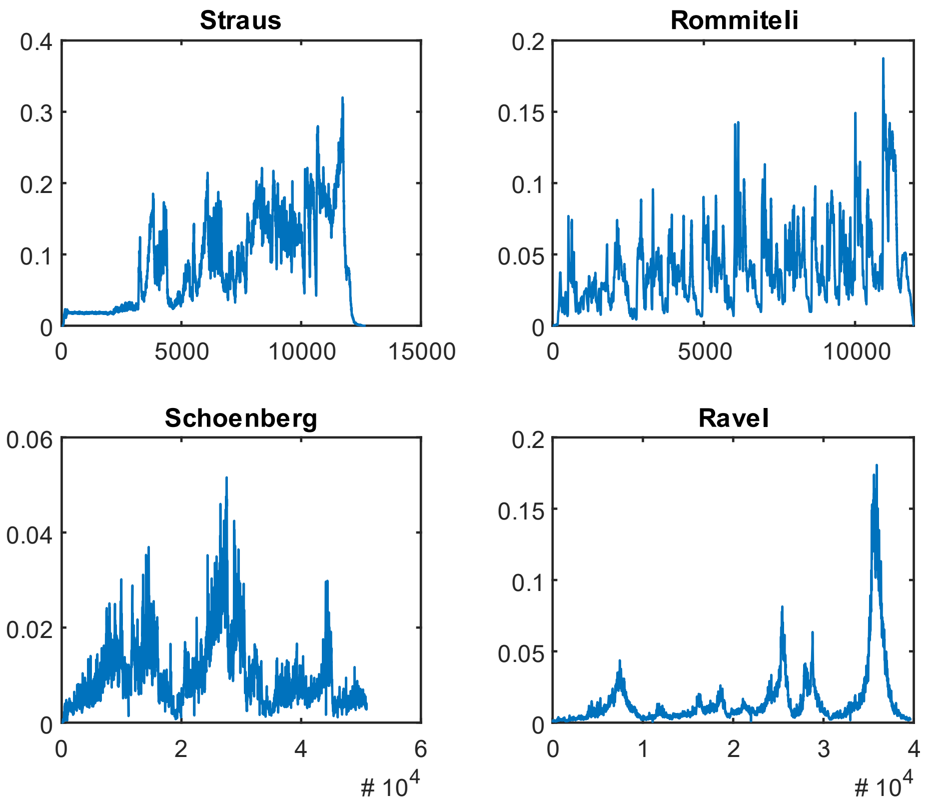

Figure 3 amplitude envelope data

5.2.3 Mirbrightness-function

Spectral energy distribution of a music file can be calculated by using the mirbrightness-function. The function calculates the amount of energy in the harmonic partials above a certain cut-off frequency which is set by default to 1500Hz. The values returned by the function range from 0 and 1. (Lartillot, 2019, p. 133.)
A number 0.6, for example, would mean that 60% of the amount of energy in the harmonic partials is above the cut-off frequency. The following figure illustrates the spectral energy distribution data from all the music files. The mirbrightness data describes the spectral energy distribution of tones. The default cut-off frequency was used.

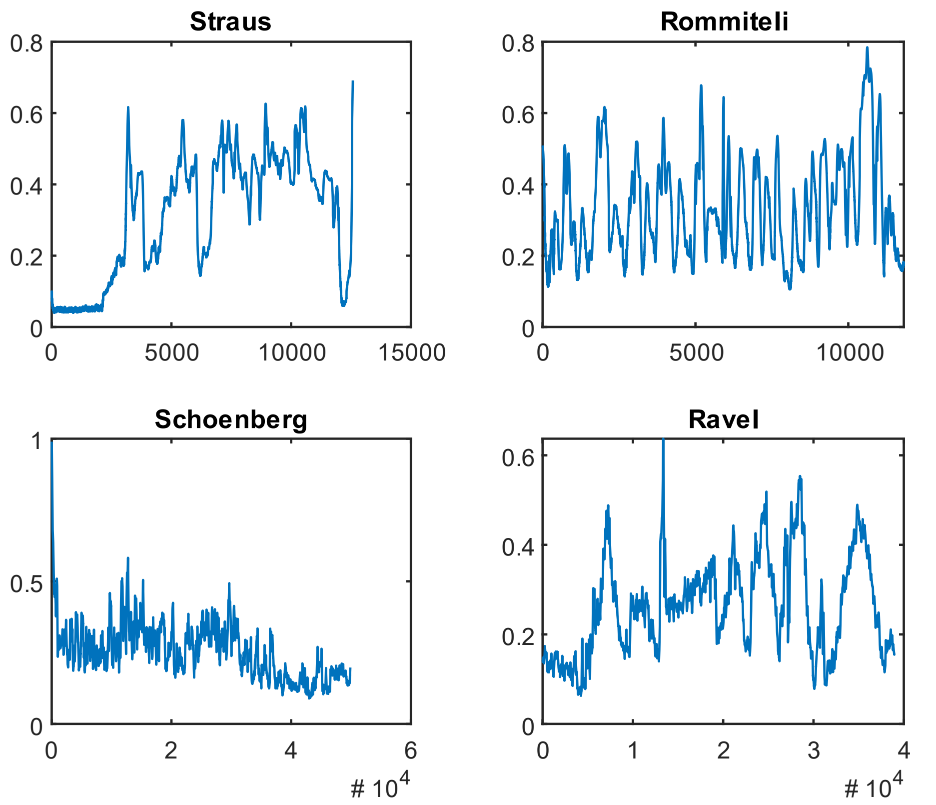

Figure 4 spectral energy distribution data

5.3 Investigation of relationship between music and movement by using scatter plots

Scatter plot is a figure containing values describing two variables. It is observed prior to further analysis to determine if relationship between the variables exists and what is the type of the relationship.
If the measurement points in the plot aligns according to a line representing an increasing or decreasing function, then the relationship between the variables is linear. Should the measurement points align according to parabola or logarithmic curve, then the relationship between the variables is non-linear. If the measurement points distribute randomly across the plot, then there is no relationship between the variables. The importance of viewing the scatter plots is high since the type of relationship, either linear or non-linear, must be considered when choosing the correlation coefficient to be used in analysis. (Nummenmaa, Holopainen & Pulkkinen, 2019, p. 211.)
The variables, movement data and the music feature data, were compared using scatter plots. The movement data of all participants in all four listening tasks were compared against the three different types of music feature data, spectral flux, amplitude envelope, and spectral energy distribution data, to see what types of listening strategies the participants had.

5.3.1 Movement data and spectral flux data

To compare movement data against the spectral flux data, the differences between each frame of the movement data was calculated. This was due to fact that the mirflux-function returns values describing the change between each frame. Comparing spectral flux data straight to the movement data would mean comparing amount of change in spectral content against to a single slider position at any given moment. To be able to see if slider movement and spectral flux data correlates, one must compare the amount of change in the spectral content against the amount of change in the slider position. The following figure illustrates the scatter plots between each participants movement data against the spectral flux data. The rows of plots represent the listening task, and the columns represents each participant.

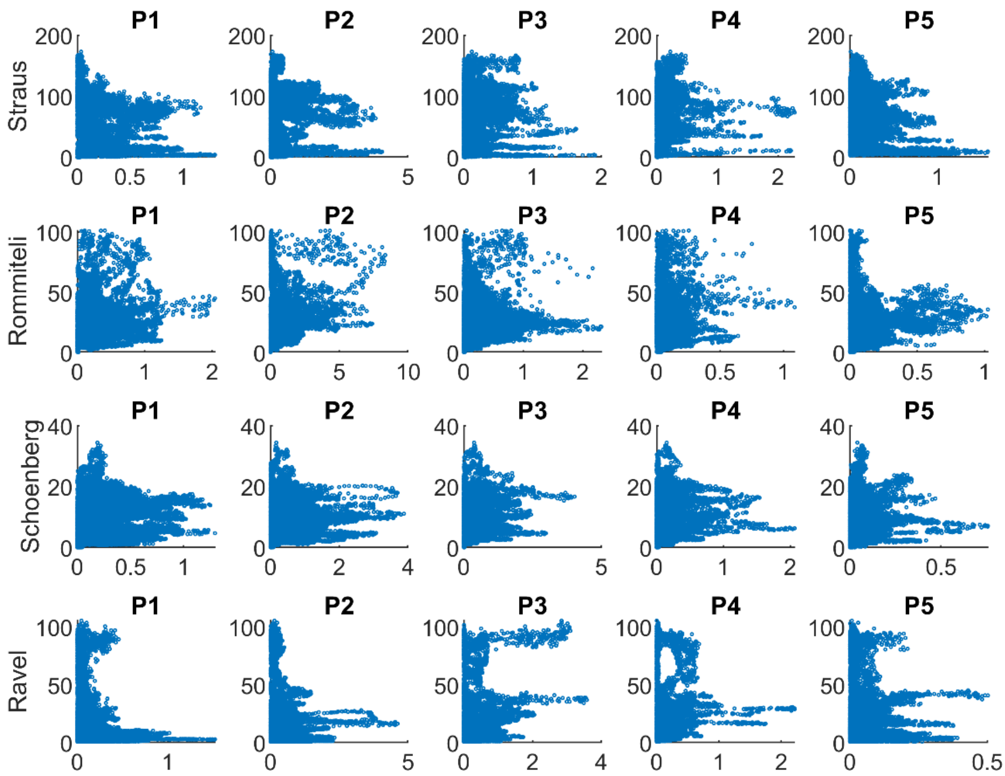

Figure 5 movement and spectral flux data

As can be observed from the figure, it seems that in some cases, for example with participant one and participant two in Ravel listening task, there seems to be a weak non-linear relationship between movement and music. With most of the cases however, there is no relationship between the temporal patterns of tones and the brightness ratings of participants.

5.3.2 Movement data and the amplitude envelope data

Since amplitude envelope data has a direction, for example when numbers are getting higher values the music is getting louder, it is possible to compare the amplitude envelope data straight against the movement data. The following figure illustrates the scatter plots between each participants movement data against the amplitude envelope data. The rows of plots represent the listening task, and the columns represents each participant.

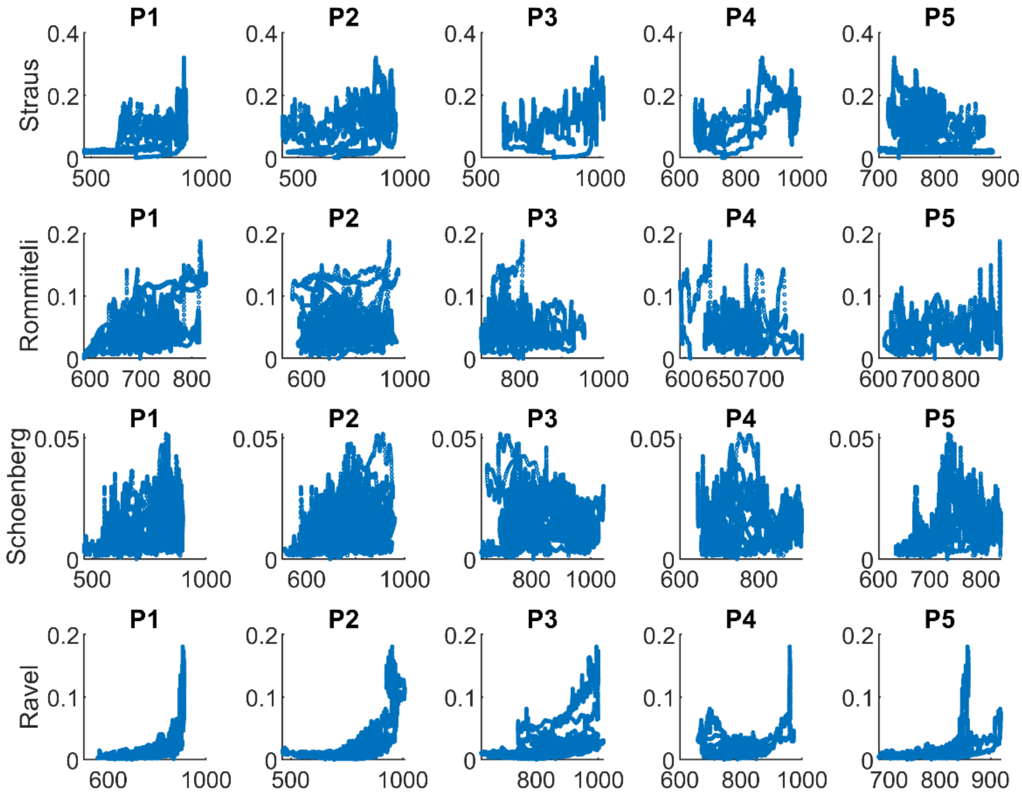

Figure 6 movement and amplitude envelope data

As can be observed from the figure, clear non-linear relationship between each participants movements and music in a listening task with Ravel's music exists. Weaker non-linear relationship between participants one to four movements and music in a listening task with Straus's music exists. Also, somewhat linear relationship exists with participant one's movement and music in a listening task with Rommiteli's music. Following two figures illustrates one of the cases with non-linear relationship with added least square line and the case with linear relationship with added least square line. Least square line can be used to model the direction of the relationship and correlation (Nummenmaa, Holopainen & Pulkkinen, 2019, p. 237).

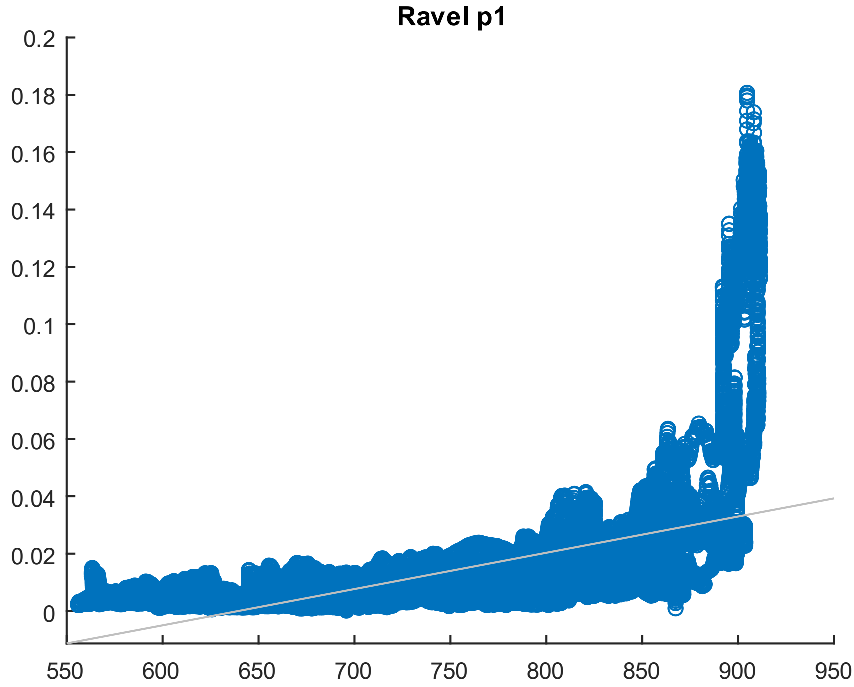

Figure 7 non-linear relationship between music and p1 movement

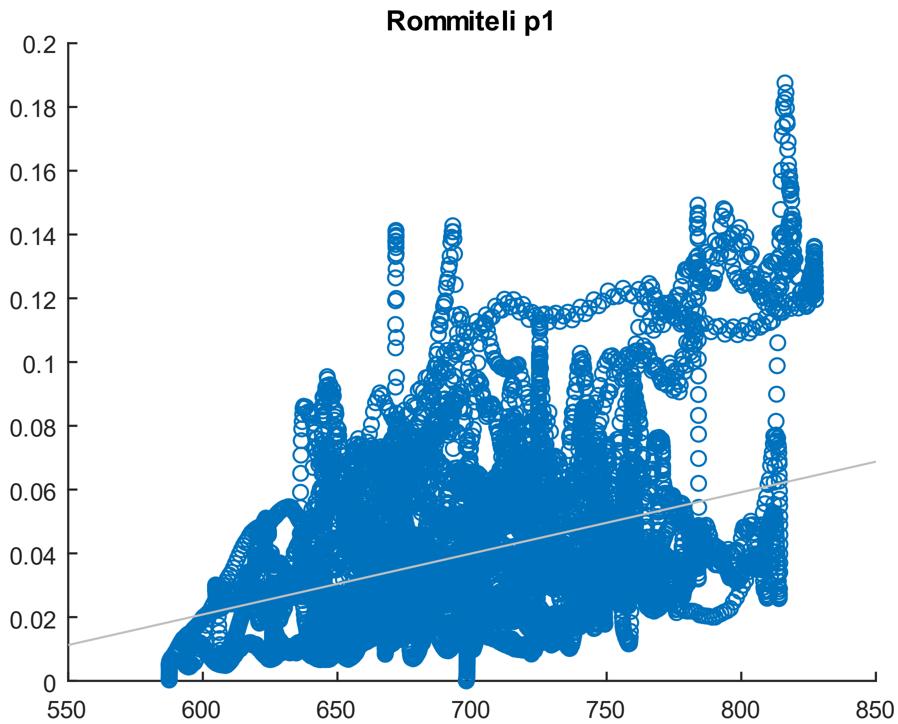

Figure 8 linear relationship between music and p1 movement

5.3.3 Movement data and the spectral energy distribution data

Since spectral energy distribution data has a direction, for example when numbers are getting higher values the energy in high partials are getting higher and thus music is getting perceptually brighter, it is possible to compare spectral energy distribution data straight against the movement data. The following figure illustrates the scatter plots. Each participants movement data against the spectral energy distribution data. The rows of plots represent the listening task, and the columns represents each participant.

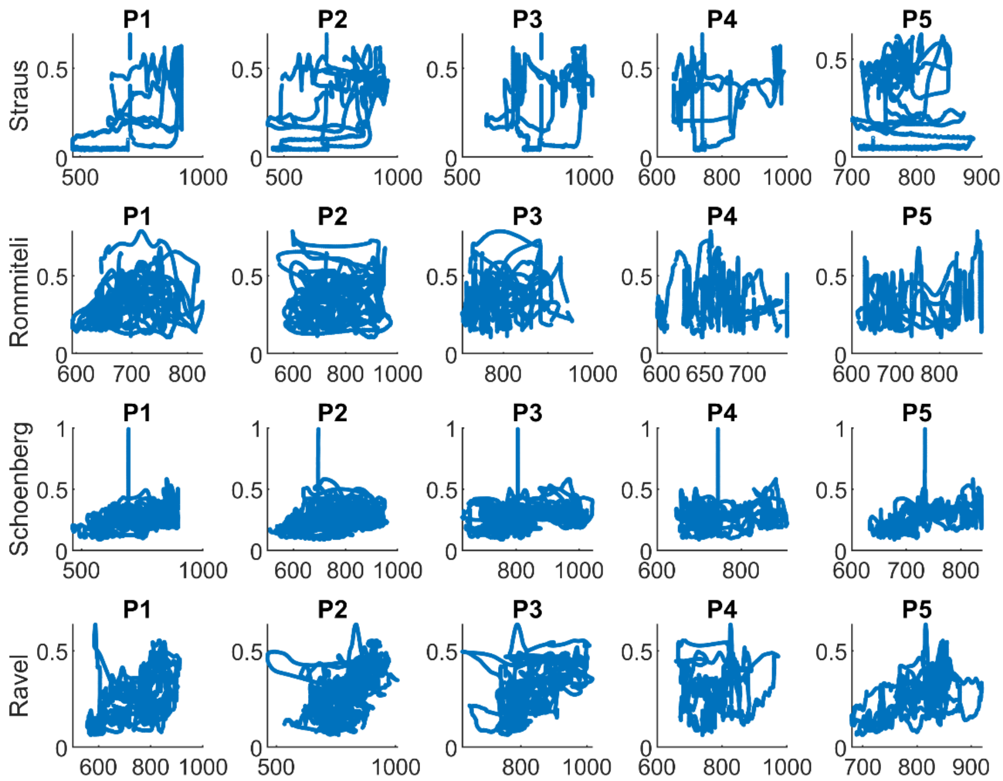

Figure 9 movement and spectral energy distribution data

By observing the scatter plots, it is possible to find somewhat linear but noisy relationship between all the participants movements and music in the Schoenberg listening task. Participants one to three, and five's movement and music in the Ravel listening task shows linear but rather noisy relationship, and similarly, participant one's movement with music in the Straus listening task.

The following figures illustrates the some of the cases with the linear relationship, with added least square line.

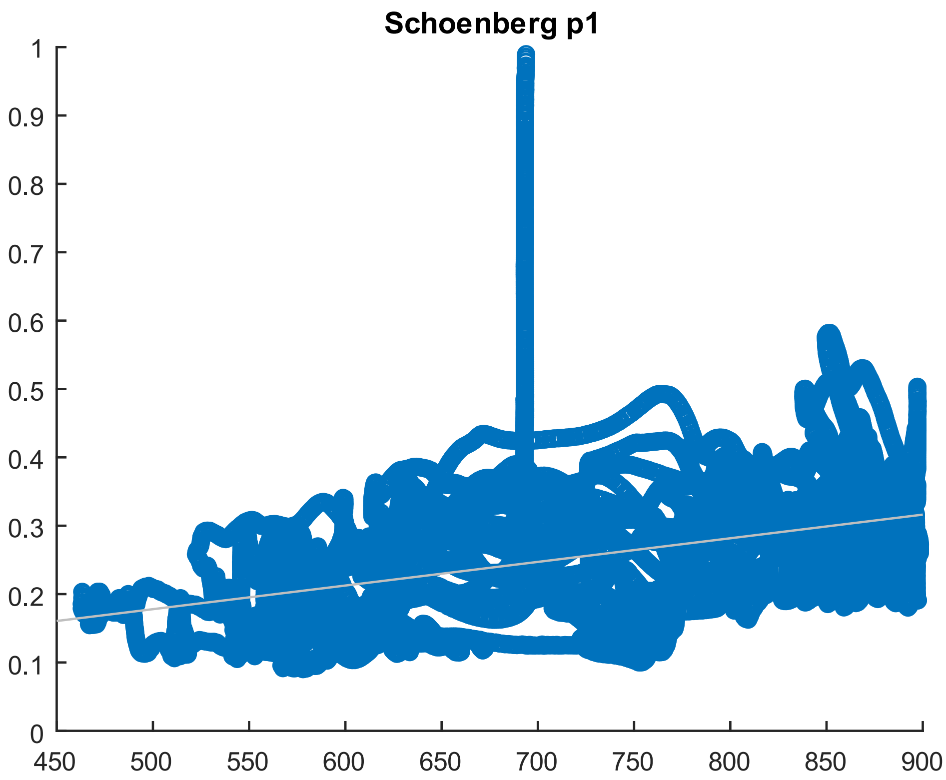

Figure 10 linear relationship between music and p1 movement

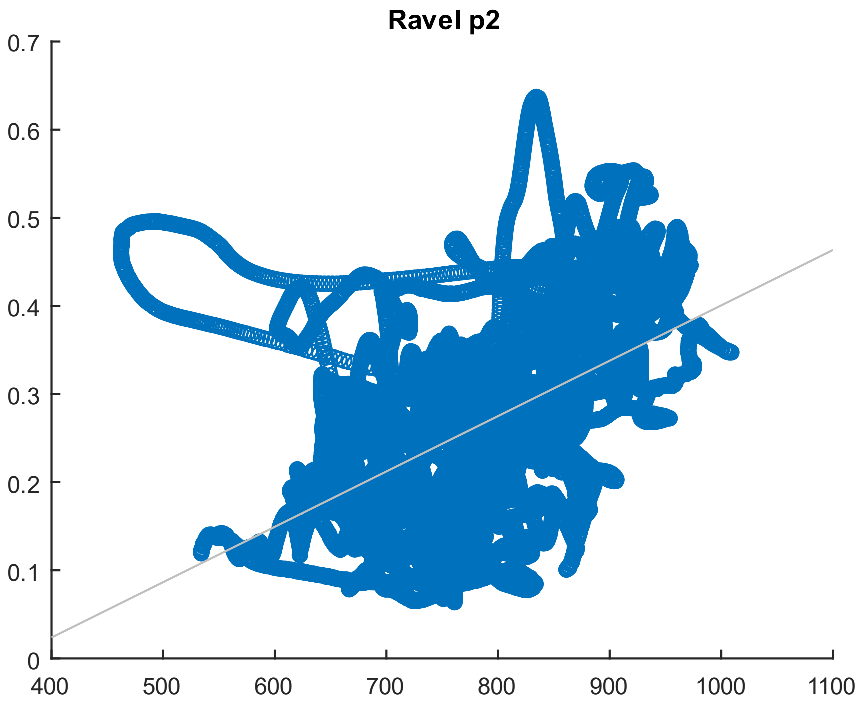

Figure 11 linear relationship between music and p2 movement

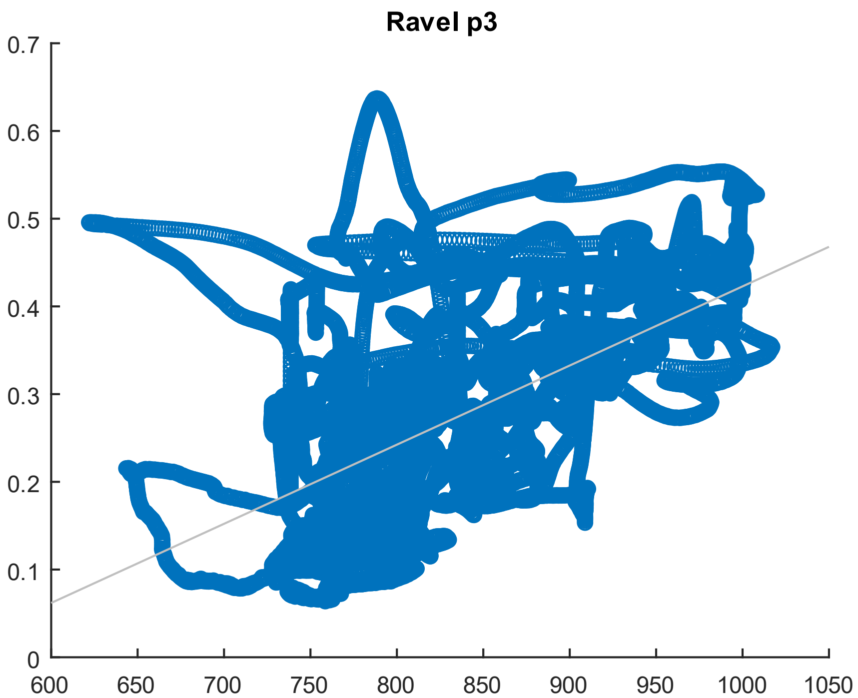

Figure 12 linear relationship between music and p3 movement

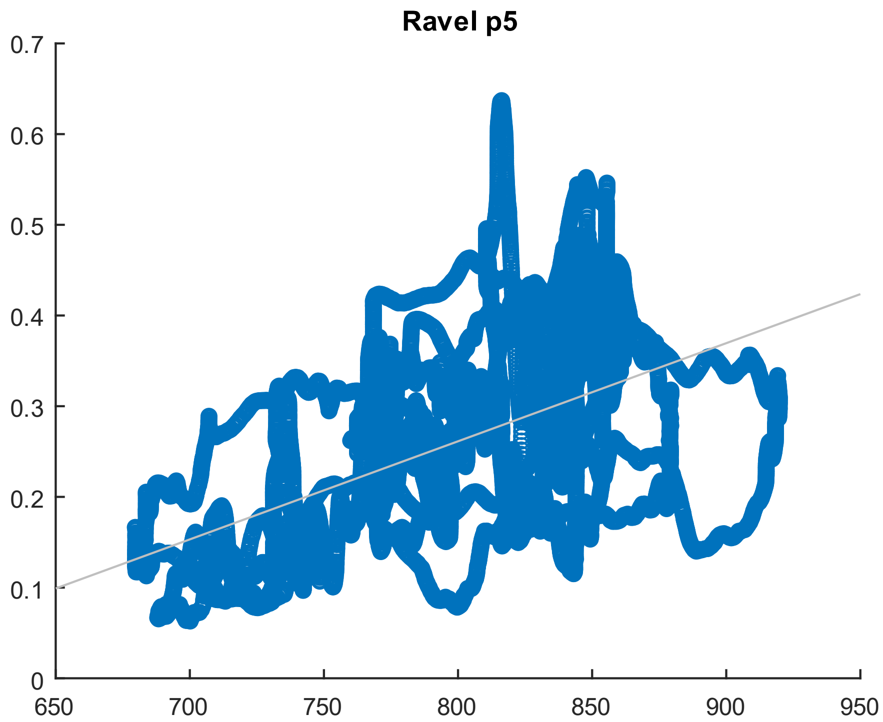

Figure 13 linear relationship music and p5 movement

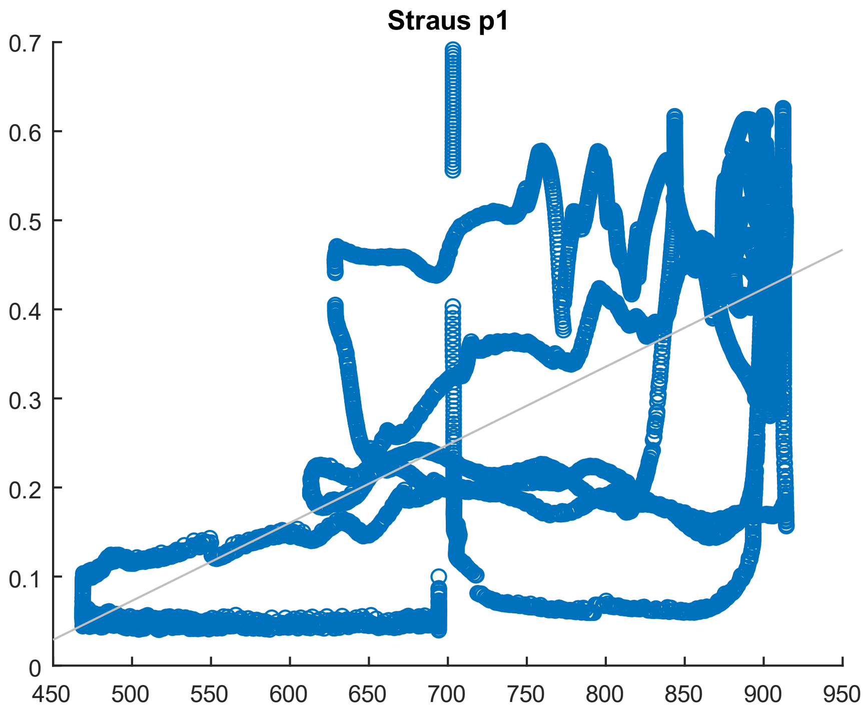

Figure 14 linear relationship between music and p1 movement

6 RESULTS

After investigation of scatter plots to determine the type of relationship that existed between participants brightness ratings and music feature, correlations were calculated to see: 1) how strong the relationship between the variables was, and 2) to determine the listening strategies used by participants to percept and rate the musical brightness contour. The following tables contains the correlation coefficient values between each participants movements and music in the four listening tasks. The R-value for possible listening strategy is highlighted and was chosen based on either high correlation coefficient value (over .50) and/or if the value is clearly higher compared to the other two. Pearson's correlation coefficient was used with the cases of linear relationship between the variables and Spearman's with non-linear.

Listening task: Straus

Participant	R-value spectral flux	R-value amplitude envelope	R-value spectral energy distribution
P1	-.25 *p = .00.*	.75 *p = .00.*	.61 *p = .00.*
P2	.043 *p = .00.*	.54 *p = .00.*	.49 *p = .00.*
P3	.09 *p = .00.*	.45 *p = .00.*	.40 *p = .00.*
P4	.11 *p = .00.*	.69 *p = .00.*	.48 *p = .00.*
P5	-.20 *p = .00.*	-.10 *p = .00.*	-.11 *p = .00.*

Listening task: Rommiteli			
Participant	R-value spectral flux	R-value amplitude envelope	R-value spectral energy distribution
P1	.13 *p = .00.*	.37 *p = .00.*	.23 *p = .00.*
P2	.26 *p = .00.*	.09 *p = .00.*	.10 *p = .00.*
P3	.08 *p = .00.*	.02 *p = .023*	.20 *p = .00.*
P4	.15 *p = .00.*	-.28 *p = .00.*	-.02 *p = .08.*
P5	.01 *p = .00.*	.37 *p = .00.*	.18 *p = .00.*

Listening task: Schoenberg			
Participant	R-value spectral flux	R-value amplitude envelope	R-value spectral energy distribution
P1	.18 *p = .00.*	.42 *p = .00.*	.45 *p = .00.*
P2	.24 *p = .00.*	.34 *p = .00.*	.30 *p = .00.*
P3	.31 *p = .00.*	.27 *p = .00.*	.25 *p = .00.*
P4	.22 *p = .00.*	.07 *p = .00.*	.09 *p = .00.*
P5	0.17 *p = .00.*	.40 *p = .00.*	.39 *p = .00.*

Listening task: Ravel			
Participant	R-value spectral flux	R-value amplitude envelope	R-value spectral energy distribution
P1	-.13 *p = .00.*	.72 *p = .00.*	.47 *p = .00.*
P2	.04 *p = .00.*	.64 *p = .00.*	.53 *p = .00.*
P3	.33 *p = .00.*	.54 *p = .00.*	.59 *p = .00.*
P4	.23 *p = .00.*	.47 *p = .00.*	.28 *p = .00.*
P5	-.01 *p = .008*	.65 *p = .00.*	.53 *p = .00.*

6.1 General observations

As can be observed from the tables, there was not a single listening strategy that the participants or a participant relied on during the listening tasks. For example, participant one apparently used amplitude envelope as a cue in the listening task to Straus's music but in the listening task to Schoenberg's music, the participant used spectral energy distribution as a cue. Similarly, participant three used either amplitude envelope or spectral energy distribution as a cue for brightness perception and rating based on the context of the listening task.

In some cases, it was not possible to determine the participants listening strategy for the musical brightness contour based on the data and the type of data analysis that was carried out. For example, there is no relationship between participant five's movement and any of the three music features in the listening task to Straus's music. Similar cases can be observed from the tables for other participant as well in the other listening tasks.

It seems that amplitude envelope out of the three music features were most used as a cue for precepting and rating the musical brightness contour. In the listening task to

Straus's music, four out of five participants used that as a cue in their listening strategy. In the listening task to Ravel's music, all the five participants used amplitude envelope as a cue.

Even though one correlation coefficient value were highlighted out of three options in the tables, it seems for example that while amplitude envelope was the main cue for the brightness perception for a participant, spectral energy distribution had also significant role. It is also interesting to note in some cases that if for example all the participant five's correlation values were to add up in the listening task to Schoenberg's music, the result is close to one (.9624). This could mean that almost all the movements can be explained by these three variables.

Based on this, it might be more appropriate to talk about main cue or dominant cue of brightness perception rather than a cue. It seems that it is possible in some cases that all the three variables which were viewed here affected to the brightness perception simultaneously even though some having more impact over the others. When this is reflected on the literature, it is reasonable that all the variables co-affected since its known that timbre perception is a multidimensional perceptual attribute where the perception of a timbre is positioned to the timbre space by the interplay of the different dimensions. The literature and the data of this study indicates that it is subjective, which the role of the different dimensions is. While talking about listening strategies of the listeners, it might be appropriate to consider about individual timbre spaces of the listeners meanwhile.

6.2 Hypothesis and the research question

The hypothesis in this study was that spectral content of the music would be the main cue or dominant cue for the brightness perception. It seems however that this is not the case. If all the data from the four listening task is viewed, the amplitude envelope played the most important role in the brightness perception of the participants.

The research question in this study was: What types of listening strategies listeners apply when focusing on musical brightness contour? The purpose was to find out by which acoustical feature of sound, participants perceive and rate the brightness in music.

As mentioned, the observations from the data suggest that there are situations where a single acoustical feature of sound can't explain alone how a listener would percept and rate brightness in music. To provide a satisfactory answer to the research question of this study, the question itself needs to be reformed slightly. While talking about listening strategies of the participants, it would be appropriate to refer and to consider about individual timbre spaces of the participants. This way the research question

would be: What types of timbre spaces listeners have and what is the role of the dimensions of the timbre spaces for the listening strategy while listener is focusing into musical brightness contour?
Based on the data-analysis the answer is that the timbre dimensions differ between the subjects and the role of the different dimensions vary between the context of listening. It seems that the timbre space is not a fixed perceptual structure that a listener has and applies as it is to the timbre perception, but rather flexible structure which changes possibly over time and at least between the listening contexts. Amplitude envelope as a feature of sound seems to have a more important role in the perception of musical brightness contour than the others. This corresponds to perceptual attack/onset dimension of the timbre space model and this dimension seems to be dominant over the other dimensions. Reflecting on the literature, this is reasonable given the importance of attack/onset part of the sound for sound source identification and for timbre of individual instrumental tones for example.

7 DISCUSS

The purpose of this study was to to survey what kind of listening strategies listeners have when they are focusing to brightness contour in music. Listening strategy for this study meant that this study was interested of the acoustic correlates of sound that the participants focus on for to perceive the brightness and its changes in music.
While this research was able to answer the research question about the listening strategies of the participants, other important discoveries were made as well during the process: 1) It seems that that the timbre space model of timbre perception, suggested by the literature, works well in the ecologically valid setting also. Previous studies have studied the topic using single instrumental tones as a stimulus and participants so for did not listen to actual music pieces during the experiments. Ecologically valid stimulus could stimulate all the timbre dimensions as the single instrumental tones did in the previous studies.
2) While the main cue for timbre perception to which the listener might or might not focus on during the listening changes between the contexts, also the timbre dimensions and space are not constant. Even for a single listener.

7.1 Suggestions for future studies

As the timbre space with its dimensions seems to be a flexible perceptual structure which changes at least between the listening tasks, it would be important to study whether the timbre space is formed at the beginning of the listening task and if it stays fixed, or if the timbre space is continuously chancing structure. This type of study would be possible to perform using the same dataset as this study but analyzing it differently. Rather that calculating a correlation value from the duration of the whole music piece, the analysis work could view the interplay of the timbre dimensions frame by frame. Deciding the frame length could prove to be a difficult task. If the study would use shorter pieces of music as the stimulus, then viewing all the measurement points might be possible.

Another future study could be simply repeating this study and collecting more data and evidence for the current findings. As the dataset of this study is small with only five participants and four listening tasks, the limitation of this study is that the results displayed here can’t be generalized yet but merely used as the guideline for planning future studies.

Another limitation of this study in addition to a small dataset, is that investigating how timbre could be perceived through movements and emotions evoked by them, was completely out of scope of this study. Literature by Wallmark (2014), and Wallmark and his colleagues (2018) suggested that while people might perceive timbre as movements that was required to produce certain sounds and timbre, these movements can be connected to emotions as well. Since some emotional states have typical movement patterns or manners to them, a future study could use an experiment to see if participants movements while rating timbre reflects the emotional content of the music and the kind of gestures that was required to produce the sound (and the music) in the first place. According to Handel (1995), with some instruments, increase in effort in sound producing process leads to non-linear increasement of energy in the partials of the soundwave created by the source-filter combination. (Handel, 1995.) This would make analyzing the timbre perception of participants interesting: normally, powerful sound producing gestures, used for expressing emotional intention of being angry for example, should lead to increased energy in the harmonic partials. If the increasement is non-linear, this could have an interesting effect on the interplay of perceptual timbre dimensions. If the spectral energy distribution dimension is getting less information, would another dimension have more dominant role in the process of creating the perceptual experience of a timbre? This idea of one dimension offering less information could be used in the context of this study also: the music that was used as stimulus could be investigated to see the changes in music, meaning where and when amplitude is low or where and when the spectral fluctuation or energy in the harmonic components is low. Then it could be analyzed if this has impact on and if it can be seen from the perceptual process of a participant.

7.2 Final conclusions

Even though the results of this study cannot be generalized, they suggest that the multidimensional model of timbre perception, formulated during the previous research in the field of music psychology, is valid and can be applied in the context of ecologically valid settings also. Prior to this study, it was assumed by the author that a listener could choose or rather could have an established listening strategy and way to deal with a task where he/she would concentrate on the timbre of music, and this could be observed from the data gathered. This turned out to not be the case since it was observed that a listener could have different approach for the different contexts of listening. However, amplitude of the music might be some sort of elementary cue in timbre perception which might at least make a listener to react to the changes in music. While this study answered to its research question, the important aspect with a question, is it possible to listener to choose to focus on a specific attribute of sound for timbre rating and perception, is still left open and unanswered.

REFERENCES

Bregman, S. (1990). Auditory Scene Analysis. The Perceptual Organization of Sound. The MIT Press.

Caclin A., McAdams S., Smith B. & Giard M. (2008). Interactive processing of timbre dimensions: an exploration with event-related potentials. *Journal of cognitive neuroscience 20:1 pp. 49-64.*

Cowart, M. (22.3.2021). Embodied cognition.
https://iep.utm.edu/embodcog/

Dahl, S., Bevilacqua, F., Bresin, R., Clayton, M., Leante, L., Poggi, I. & Rasamimanana, N. (2010). Gestures in performance. In R. Godøy & M. Leman (Eds.) *Musical gestures: sound, movement and meaning* (pp. 36-68). Routledge

Dahl, S. & Friberg, A. (2003). Expressiveness of musician's body movements in performances on marimba. *Conference paper January 2003*

Dahl, S. & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception June 2007 DOI: 10.1525/mp.2007.24.5.433*

Dohlan, E & Rehding, A. (2021) Timbre. Alternative Histories and Possible Futures for the Study of Music. In E. Dolan & A. Rehding (Eds.) *The Oxford Handbook of Timbre*. (pp. 3-20). New York, NY: Oxford University Press.

Ferrer, R. (24.3.2021) Embodied cognition applied to timbre and musical appreciation: theoretical foundation.
http://britishpostgraduatemusicology.org/bpm10/ferrer_rafael-embodied_cognition_applied_to_timbre_and_musical_appreciation_theoretical_foundation.pdf

Fink, R., Wallmark Z. & Latour M. (2018) Introduction. In R. Fink, M. Latour & Z. Wallmark (Eds.) *The Relentless Pursuit of Tone: Timbre in Popular Music.* (pp. 1-18). Oxford University Press, Incorporated

Friberg, A. (2006). pDM: An expressive sequencer with real-time control of the KTH music performance rules. *Computer Music Journal March 2006 DOI: 10.1162/comj.2006.30.1.37*

Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of. Acoustical Society of America., Vol. 61, No. 5, May 1977*. (pp. 1270-1277)

Godøy, R. (2010). Gestural affordances of musical sound. In R. Godøy & M. Leman (Eds.) *Musical gestures: sound, movement and meaning* (pp. 103-125). Routledge

Halmrast, T., Guettler, K, Bader, K. & Godøy, R. (2010). Gesture and timbre. In R. Godøy & M. Leman (Eds.) *Musical gestures: sound, movement and meaning* (pp. 183-211). Routledge

Handel, S. (1995). Timbre Perception and Auditory Object Identification. In B. Moore (Ed.) *Hearing*. (pp. 425-459) San Diego: Academic Press 1995

Jensenius, A., Wanderley, M., Godøy, R. & Leman, M. (2010). Musical gestures. In R. Godøy & M. Leman (Eds.) *Musical gestures: sound, movement and meaning* (pp. 12-35). Routledge

Johnson, M. & Larson, S. (2003). "Something in the way she moves"-Metaphors of musical motion. *METAPHOR AND SYMBOL, 18(2), 63-84*

Järveläinen, H., (2010) Musiikin ryhmittyminen ja kuulemma-analyysi. In J., Louhivuori & S., Saarikallio (Eds.) *Musiikkipsykologia* (pp. 67–82) Atena 2010

Killick, A. (2021). Representing timbre with a spectrogram.
http://globalnotation.org.uk/representing-timbre-with-a-spectrogram/

Lartillot, O. (2019) *MIRtoolbox 1.7.2. User's manual.* University of Oslo, Norway

Leman, M. & Godøy, R. (2010). Why study musical gestures? In R. Godøy & M. Leman (Eds.) *Musical gestures: sound, movement and meaning* (pp. 3-11). Routledge

Lhommet, M. & Marsella, S. (2015). Expressing emotion through posture and gesture. In R. Calvo, S. D'Mello, J. Gratch & A. Kappas (Eds.) *The Oxford handbook of afective computing* (pp. 273-285) Oxford University Press

Nummenmaa L., Holopainen M. & Pulkkinen P. (2019) *Tilastollisten menetelmien perusteet.* Sanoma Pro Oy

Shapiro, L. (2007). The embodied cognition research programme. *Philosophy Compass 2/2 (2007): 338–346, 10.1111/j.1747-9991.2007.00064.x*

Slawson, W. (1980). *Sound color*. University of California Press. Berkley and Los Angeles, California

Wallmark, Z. (2014). *Appraising timbre: embodiment and affect at the threshold of music and noise* (Doctoral thesis, university of California) https://escholarship.org/content/qt99t2t939/qt99t2t939_noSplash_85095c136e24c47adeb1b9b9e0523e67.pdf?t=nrzlzh

Wallmark, Z., Deblieck, C. & Iacoboni, M. (2018). Embodied listening and timbre: perceptual, acoustical and neural correlates. *Music perception February 2018 DOI: 10.1525/mp.2018.35.3.332*

Wallmark, Z. & Kendall, R. (2021) Describing Sound. In E. Dolan & A. Rehding (Eds.) *The Oxford Handbook of Timbre*. (pp. 579-608). New York, NY: Oxford University Press.

Williamon, A., Ginsborg, J., Perkins, R. & Waddell, G. (2021). Performing Music Research: Methods in Music Education, Psychology, and Performance Science. Oxford, New York: Oxford University Press.

Wilson, R. & Foglia, L. (2015). Embodied cognition. https://plato.stanford.edu/entries/embodied-cognition/