Art and language: is music a simple metaphor or an alternative way of speaking?

Arte e linguagem: a música é uma simples metáfora ou uma forma alternativa de falar?

Arte y lenguaje: ¿es la música una simple metáfora o una forma alternativa de hablar?

ABSTRACT
Music, along with painting, literature and architecture, is considered one of the “fine arts”, which means that it is to be appreciated primarily or even solely by means of their imaginative, aesthetic or intellectual content. Mainly in Western Music, we can see an ensemble of musicians in an orchestra performing an artistic miracle of playing together the same piece, each one with their respective instruments, with almost no flaws whatsoever. However, we might ask ourselves: are they really flawless as we say about a correct result of a mathematical equation? and/or are they really following a strictly precise rule that allow them to “prove” all the steps to a same exact “true” performance? And, about its expressive capacities, can music convey actual fixed concepts as do words, paragraphs or entire books? On this paper, we are going to analyze how far Music Theory is able to determine accurateness in a musical score, regarding reading for a performance and the aesthetic effects in a hearer, highlighting the conceptual problems behind numbers and linguistic reference and how music is actually much more similar to the cognitive socio-linguistic phenomenon of the pragmatics of language than to an exact science.

Keywords: Music. Accurateness. Theory. Expressivity. Language.

RESUMO
A música, juntamente com a pintura, a literatura e a arquitectura, é considerada uma das “belas artes”, o que significa que deve ser apreciada principalmente ou mesmo apenas através do seu conteúdo imaginativo, estético ou intelectual. Principalmente na música ocidental, podemos ver um conjunto de músicos em
uma orquestra realizando o milagre artístico de tocar juntos a mesma peça, cada um com seus respectivos instrumentos, quase sem nenhum defeito. No entanto, podemos perguntar-nos: serão realmente perfeitos como dizemos sobre um resultado correcto de uma equação matemática? e/ou eles estão realmente seguindo uma regra estritamente precisa que lhes permite “provar” todas as etapas para um mesmo desempenho “verdadeiro”? E, no que diz respeito às suas capacidades expressivas, pode a música transmitir conceitos fixos reais como fazem palavras, parágrafos ou livros inteiros? Neste artigo, analisaremos até que ponto a Teoria Musical é capaz de determinar a precisão de uma partitura musical, no que diz respeito à leitura para uma performance e aos efeitos estéticos em um ouvinte, destacando os problemas conceituais por trás dos números e da referência linguística e como a música é realmente muito mais semelhante ao fenômeno sociolingüístico cognitivo da pragmática da linguagem do que a uma ciência exata.

waste too much time with us, annoying children, and simply put us to be distracted by something loud, cheerful and singable. It’s easy, as well, to see that we are always relying on metaphors when we try to describe what the effects of hearing music are on us. One thing is certain: these metaphors are not by chance. They are not just the best resource to describe music, at least, for a non-initiated person in Music Theory, but they are, actually, the only linguistic way of doing so, because theory, scientifically speaking, is much more concerned on stating how things work in order to promote (or to predict) the repetition of an event. But, have we created music only to try to repeat or predict it? Is music even useful for our primary needs, like eating and housing, as Geometry or Biology are for measuring terrains, capacities and our blood taxes? It seems that we haven’t had to formulate or use music in the first place! Unless we start to consider that the phenomenon of Art is something intrinsic to the human evolution, mainly when we consider that language\textsuperscript{1} is what characterizes us the most.

So, on this paper, we are going to occupy ourselves answering, primarily, the following questions: 1. is music, as a type of art, just an alternative way of speaking (a prettier way of speaking)? 2. If so, can we determine a semantics, syntax and pragmatics of it? 3. Do we use metaphors to say that a piece of music is (or expresses adjectives as) cheerful, sad, melancholic, exciting, depressive, strong, empty, spatial etc. because music is merely an imitation of language or they both are tools of a same primitive purpose of communication? 4. Related to the question 2, can we develop a grammar and a dictionary of music? 5. In general, regarding the preciseness of how people speaking the same language don’t have any significant problems whatsoever to communicate, how can we compare the tools of Arithmetics (an exact science) to the rules of Grammar? 6. And, finally, is music writing and reading (and performing) as accurate as it seems

\textsuperscript{1} As John Searle has defined in his Speech Acts (1969), language is an institutional fact (SEARLE, 1969, p. 50), which means that this has a human and, therefore, a compromise with the creation of social conventions and institutions in order to build the complexity of people interaction. For us, in here, music is the same thing. This will avoid us to consider that a singing of a bird, for the bird, or a repetitive noise provoked by an object, for a given object, are music. For more, read my other work in Portuguese called “A Filosofia por trás do som: ontologia, expressividade e epistemologia da música”, 2023, p. 209.
to be considering the fact that it uses a lot of notions from Geometry, Mathematics and Physics (alleged “exact” sciences as well)?

It is clear, therefore, that we are going to feast on debates concerning Philosophy of Language, Art and Science, not necessarily in this particular order. Problems of linguistic reference (Bertrand Russell, Saul Kripke, Peter Strawson) as well as contextual aspects of pragmatics (John Searle, Daniel Vanderveken) will be noted in that first branch mentioned above. For the second branch, specifically the Philosophy of Music is going to take part (Deryck Cooke, Peter Kivy). And, for the last, we will deal with the science behind the Music Theory (Alfred Blatter, Daniel Wilkerson) and philosophical problems concerning science in general (Karl Popper). Our objective is to answer the six questions above based on a critical analysis of these author’s works.

2 A PHILOSOPHY FOR MUSIC THEORY

In order to state some fundamental problems behind the music theory, let’s remind us (or learn) some principles of music understanding. We won’t occupy the reader with a list of tiresome definitions (like we see in music theory books), but with a rise of several debates concerning some of them, first, because this is a philosophical work, and second, because the core of these definitions are in check here, so, when we start to doubt some fundamental things, the rest of them, which are dependent on the foundations, might fall. We will admit also that the reader has some familiarity with music theory.

Music is a temporal art (Blatter, 2007, p. 1) that organizes sound waves in a specific and pleasant way for the human ear. We are going to deal here mainly with the pure music². Basically, it is divided in melody (a sequence of notes/pitches, i.e., sounds), harmony (how these pitches relate proportionally at the same time) and, for many authors, counterpoint (nothing more than the melodic and harmonic relationship of musical lines. It doesn’t actually seem to be

---

² Peter Kivy defines “pure music” in his Music Alone: Philosophical Reflections on the Purely Musical Experience (1990) as “just plain music; music unaccompanied by text, title, subject, program, or plot; in other words, music alone” (p. ix).
something new). But, what are those things in these definitions exactly? In order to organize their conceptual differences, it was necessary to formulate a bunch of symbols that characterize the notation for music writing and reading, which, of course, has suffered several changes during history. Let’s take a look at them.

The Pitch is our perception of the relative highness or lowness of a sound in terms of frequency (Hertz). It is different from its Loudness, which determines the potency (Watts) of a pitch (about what we will talk later). Since music is temporal, similar to the writing of a natural language in the west part of the world, we need guidelines to put the symbols of the sounds, from left to right, in order to organize them according to the premises of a “good” organization (like Logic does to the human thinking). In western music, the main pitches are seven in number and they are named by A, B, C, D, E, F, G. In the following image, we have the Staff (five lines, a pentagram, whose lines and spaces between lines are used to place a pitch, counting bottom-up). We also have the Clef to establish where the pitches are on the staff. On the example, we have the G-clef (there are others), because its most-to-the-right point (in the “belly” of the symbol) decides that we can find the G pitch on that line.

![Figure 1 - The pitches on the staff](image)

Duration is the relative length of a pitch (measured in terms of seconds). Before we discuss the symbols of duration, we should make our first philosophical inquire about music theory. We start to see one big similarity between music and language, despite of the fact that they essentially are about sound. These three basic aspects of music, pitch, loudness and duration are all relatives. What does that mean? The word “house” means “a building for human habitation”, but not in every language we use this word to denote the same object (and there are a lot of ways to define what is, properly, a human habitation). Also, what could be

---

3 All figures have been created with the help of the Windows application called MuseScore4.
something “fast” to someone, might not be for someone else. In our natural language, we have arbitrated that the length of word writing and speaking should not be “so long”, so we could speak about a decent number of things in a timely manner. Same thing for music, although, as we will see soon about expressivity of music (section 4), the pitches are not semantically equivalent to words in a natural language, they must have a given duration, accordingly to the number of beats (pulse, BPMs, beats per minute) that you want your composition to have. This number also defines what we call a time signature, which is a fractioned number representing the number of pitches (numerator) and their lengths (denominator). A measure is merely an arbitrary division of the pitches to help the reader identify better the symbols of duration in a musical score.

The symbols (rhythm figures) for the duration of the pitches are relative to one another proportionally using division by the number 2, but they don’t have an absolute value until the composer define the pulse. They are the following:

Consider a time signature of 4/4. 24bpm means that, if you have 6 measures within a minute and each one of them has 4 quarter-notes (1/4 of a whole-note duration), you have 4 notes per 10s. So, each quarter-note has 2.5s. This means that a whole-note has 2.5s x 4 = 10s in this song.
All of this seems pretty accurate, but this is not the only way of defining the duration of a pitch and of a song, as a whole. Actually, in musical scores for orchestral performances, it is not uncommon to have several *imprecise* (subjective) ways of communicate the beat of a piece of music, using what we call *tempo marks*. Some of them are (names in Italian): grave (extremely slow), adagio (comfortably slow), andante (walking speed), vivace (lively), presto (very fast) etc. We can also talk about the *sense* of speed, using words like animato (with life), bravura (boldly), dolore (sadness), fuco (fiery), grazioso (gracefully) with the help of some prepositions and adverbs, like a (by), ma (but, however), molto (very much), non (not), poco (little) etc. Also, totally related to the sense of speed is the loudness of the sound, which can be measured by electronic equipment. However, during a performance, these things have to be communicated by the *conductor* using specific hand and baton movements, as well as they have to be written in the musical score, as the tempo marks above. Some of them are: triple forte (as loud as possible), forte (loud), piano (soft), pianissimo (very soft) etc. This aspect is what we call the *dynamics* of the music. We also have gradual increases (crescendo) and decreases (decrescendo) in music loudness, which are very imprecise as well.

The point is that, despite of the fact that the orchestra seems to be very precise in their notes, be about speed or intensity, *there is no other way* of dealing with so many differences in music performance than using terms of natural language itself. This subjectiveness could be a problem to the immaculate

---

4 Blatter highlights (2007, p. 38) how the conductor is the most natural speaker of all the orchestra, literally relying on an almost ordinary language to orient the performers in a very metaphorical way of utterances: “The primary responsibilities of the conductor are to set tempos, facilitate ensemble, communicate style, and adjust dynamic balances. Since speaking during a performance obviously would adversely affect the music, conductors communicate through gestures, eye contact, and body language”. There is a pretty interesting video showing how this actually works. It is a video called “What does a conductor really do on the stage?”, made by the YouTube Channel Vox. Link: https://www.youtube.com/watch?v=z_yIn8V3UcU&t=100s
accurateness of music, but this is not the case. Although the other fine arts have a science describing their rules as much as music has, the latter has this arrogant\(^5\) vibe of having its own theory (is much easier to find books called “music theory” than the others) while the former doesn’t seem like it, maybe because they are not so easily accessed (paintings and architecture worth of being named as “art” are often found in museums or in preserved places in a specific part of a city). But, thing is, even in music, mathematics has found some problems concerning the linguistic problem of reference.

Let’s learn something about the *tuplets*, more specifically, a *triplet*. Try to see a problem in the following image (let’s also remember the example of Fig.3 used above explaining time signatures):

![Figure 4 - A conventional way to represent a triplet](image)

Source: own authorship (2024)

\(^5\) According to Wilkerson (2012) in his *Harmony Explained*, the way people still learn about music theory “is like reading a medieval medical textbook: such books are full of unjustified superstition, non-reasoning, and funny symbols glorified by Latin phrases” and that “music theory is not a scientific theory of anything” because it fails to contain any of these five aspects of a scientific theory: It has to be *mechanical* (something has to at least appear to work in a pattern); *deterministic* (we have to be bold to say that something will happen according to the rules); *universal* (it has to apply to every object predicted by the theory); *minimal* (the Occam’s razor: the best explanation is the shortest one) and finally, the theory has to be *factorable*, that means, although simple, it has to have simple parts in its explanation. Well, Wilkerson tried to be convincing, but it seems that he has fallen in a *petitio principi* (and here it is another Latin phrase that he hates!). Is he making a theory of a what makes a theory? To assume that a theory has to present a mechanism is already a way of saying that we know what mechanism this is (the simplest that it can be) and, thus, having a theory already, *before* we know that something is, in fact, a theory. But, developing a theory is trying to get a pattern on the event in the first place (not assuming it)! Either this might compromise you to say that everything has a science (and, of course, music would have a lot) or to say that there is no satisfactory explanation for anything (what we would call “scientific”), because any minimum pattern would be enough to develop a theory and we will never know the “perfect” numbers of patterns to justify something as sufficiently (and scientifically) predictable. E.g. if I drop an apple, it falls down. The first step is to see that this is a pattern and, thus, might be considered a theory already (it is also arbitrarily deterministic and universal, given a certain scenario chosen by the scientist. Being minimal or factorable are just numerical and arbitrary aspects too). So, why would we bother, after the first step, in continuing the other steps in order to *still* determine whether this is a theory or not (something that we already know)?
According to the proportion that we have seen above in Fig. 2, a new notation is telling us that, instead of putting 2 quarter-notes (each one 2.5s) to be length-equivalent to a half-note, we have now 3 quarter-notes, indicated by the line above them. This is supposed to represent that the performer has to play “3 quarter-notes within the time of 2” (???). As we have established, once we have the time signature and the number of beats, all the symbols will have their respective durations, exactly, according to our convention of time measuring of seconds. If someone tells you that you can compress the length of 3 quarter-notes within the length of 2, you are in front of a paradox. You have two choices: either you say that the quarter-note has 2 types of lengths at the same time, discarding its symbolic and referential value, or you could actually put any symbol below the line saying “3”, because the only thing that matters is the new number of beats (3) within 2.5s x 2 = 5s. Each “new” quarter-note (or whatever symbol you put!) would have 5/3 = 1.666…s.

Another strange thing that is allowed in the notation of a music score is that you can say that the time signature is 4/4, which means that you can find within two bars (a measure) 4 quarter-notes, can also comport 2 quarter-notes and 1 half-note. In fact, we have the same length in duration, but we don’t have the same number of beats pre-decided by the pulse of the music. With this change, we would play 3 pitches instead of 4, being the last one (the half-note) longer than the first 2 pitches (the quarter-notes). We could solve this creating new time signatures, say, 2/4 and 1/2 for each measure, what would be another paradox, because is in the definition of a measure to have a given number of

---

6 In a practical point of view, of course, musicians “know what they’re doing”. According to Blatter on p. 26, footnote 21: “Although throughout this discussion it has been assumed that notes with equal durations should be performed with equal lengths, that assumption is not always true. Historically, some musical styles have required unequal durations (although the extent of the inequalities is debated). In current performance practices, jazz or swing notation assumes unequal durations. In fact, in swing notation one must specify “even eights” when unequal durations are not wanted”. Ok, but this doesn’t remove the fact that, in a philosophical point of view, we are dealing with the same symbol denoting two different things (a homonym), which is essentially a paradox, although accepted in the natural (not rigorous) linguistic interaction. Nevertheless, analytical philosophy is interested in rigor. This reminds us the debate between the Descriptivism (Bertrand Russell and Peter Strawson) and the position of Saul Kripke in his Naming and Necessity (1980), in which he tells us that there is no description to an object that can always determine it as a fix reference in every possible world. Only a proper name would suffice and resist this modal test. But, is it even possible to have a “void” proper name referring to an object without any minimum fixed description whatsoever (in other words, a proper name referring to nothing)?
beats (the numerator) with their respective length values (the denominator), otherwise it would be useless to create the time signature in the first place! Look how strange this is:

Figure 5 - The use of half-note instead of a quarter-note

Looking at the second measure, we ask ourselves: what is the meaning of the numerator “4” to it? We are all counting 3 pitches and 3 beats. Although we know that the music is not going to take longer to finish, when we look to that half-note, we also ask ourselves the meaning of the denominator “4”, because it is not a quarter-note. “You have to only care about the length equivalence”, they say. However, all the rhythm figures are related to one another by proportion, so we could put any correct number of equivalent symbols in place of another, although not more than the size of 4 pitches of quarter-notes per measure. But, this would alter the number of beats, which should alter the time signature as well, because it is determined by the number of beats first established. So, instead of 24 bpm, we would have 23bpm. At the same time, however, we know that we have 24bpm. If we start to say that a longer pitch has its length equivalent to 2 times of its previous lesser pitch (like a half-note has 2 quarter-notes) and still say that the first has “2 beats” (actually, it only sounds without a pause during the time of these 2 lesser pitches), we would end up in the problem of the infinite division of a time interval, saying that this also has infinite beats, changing the pulse to ∞bpm just because of this weird convention. A beat should be a distinguishable sound of a pitch, not divisions within an interval determined by the length of a given pitch (which can be divided indefinitely, sounding as a unison to the human ear).

There is also this strange concept of pause, that counts as a beat with the same equivalences of the pitches proportions (see Fig.2), although we still have the sensation of a pause not formally counted by these pause symbology. This sensation appears between the sound of the pitches:
That third symbol is a pause (rest) that takes the same time as the quarter-note to finish. But, what about the pauses between the pitches? This is divided, once again, by the pulse of the music, but only by the clock. So, we couldn’t actually say that, in the example of Fig.3, the quarter note has exactly 2.5s, but a little less, because we have to consider the “informal” pauses and respect the length of 1min and the 24bpm (otherwise, we would hear only a constant pitch without noticing the 4 quarter-notes). It’s amazing how intuitive is for the performer to respect the “exact” moment of finishing a pitch to start another without any formal symbol for this. Is this intuition the same as that we use to wait someone finishing his or her line to start our turn of conversation? The same as when we speak accordingly to someone else’s voice volume and mood? Is it related to politeness and good sense? All of these things are imprecise, showing that music is not so mathematical as we have thought. Our fundamental questions 5 and 6 can be answered already: music and language are not so precise.

We could talk about more metaphors in music related to natural languages. A legato (heritage), for example, is a way of extend the pitch as much as you find necessary until, generally, a different one, to create a specific feeling in your song (tied notes do the same thing, but the first and last notes are the same). As we will discuss, again (in section 4), it wouldn’t be plausible to compare pitches with words, but a specific group of them, according to well accepted semantic patterns in music. But, should be the legato and staccato (sudden and “dry” interruption of a note) similar to the linguistic events when someone speaks too much or is impolitely stopped? Notations as Da Capo and Dal Segno work in a music sheet as “relative pronouns”, signing to the performer that some parts of the music should be repeated from the beginning or from a certain point of the music, respectively.
According to Blatter (2007, p. 29), even some of the symbols of our current notation had their origin in a similarity with the language, more specifically, with the art of poetry (the *poetic meters*). Instead of talking about instruments, we used to deal with the *sung* form of the words. Here is the scheme:

Figure 7 - Notation comparison between poetic meters and rhythm figures

![Notation comparison between poetic meters and rhythm figures](source: own authorship (2024))

These *poetic foot* names work this way: Iamb has two syllables, the first of which is unstressed and the latter of which is stressed (e.g., “today”); Trochee has two syllables, the first of which is stressed and the second of which is unstressed (e.g., “matter”); Spondee has two consecutive stressed syllables (e.g., “A.I.”); Anapest has three syllables, the first two of which are unstressed and the last of which is stressed (e.g., “unaware”) and, finally, we have Dactyl, which has three syllables, the first stressed and the following two unstressed (e.g., “Waverly”). We have also their lengths, like monometer: one foot, one beat per line, dimeter: two feet, two beats per line and so on (with Greek prefixes) until the number eight. An example of an *lambic tetrameter* would be that of William Shakespeare: “Shall I compare thee to a summer’s day? (Sonnet 18):” The

---

7 For more details, see the webpage: [https://poets.org/glossary/meter#:~:text=The%20word%20comes%20from%20the%20as%20accented%20with%20unstressed%20syllables.](https://poets.org/glossary/meter#:~:text=The%20word%20comes%20from%20the%20as%20accented%20with%20unstressed%20syllables.) There is also a good video showing how these things actually sound. It is called “What is Meter in Poetry? : A Literary Guide for English Students and Teachers”. It is a content made by the Oregon State University and it has been provided by the YouTube Channel OSU School of Writing, Literature and Film: [https://www.youtube.com/watch?v=S13Tg3RAUW4.](https://www.youtube.com/watch?v=S13Tg3RAUW4.)
highlighted parts of the words show that the stressed parts are always after an unstressed one and it happens four times. It is easy to see, therefore, that our firsts notions of pitches and durations have come from the very act of speaking, and a lot of metaphors used to allude to screams, cries, smiles, anger, tranquility, tiresomeness, courage (all things represented by the human voice) are used in music expression.

Speaking more about these metaphors, we can talk about the different “steps” that we have between pitches.

Figure 8 - All 12 pitches (C-B) in a section of a keyboard

![Figure 8](image)

Source: own authorship (2024)

Every time that we have a black key between two white pitches, we say that this is a “whole step” (whole tone or, simply, a tone), otherwise it is a “half-step” (semitone). From C to B we count 7 pitches (with black keys, 12). We say that two pitches are (octave) equivalents if one has twice as many sound vibrations as the other. The symbols in the black keys are the sharps and flats, obtained by accidentals on the music staff. These are symbols in front of a pitch, working like this: Sharp ♯ raises associated pitch by one semitone; Flat ♭ lowers associated pitch by one semitone; Double sharp (♯ ♯) raises the associated pitch by two semitones; Double flat ♭♭ lowers the associated pitch by two semitones; Natural ♮ removes effect of a sharp, a double sharp, a flat, or a double flat. So, we have a total of 12 pitches (the chromatic scale). This counting will be important when we talk about the “dictionary of music” (section 4). Here is an example: the last half-note is “suffering” an accident in the second space, bottom-up, dropping by 1 semitone, because of the symbol ♭. It has gone from A to A♭ or G♯ (just see the black keys in the figure above):
But, comparing this to natural language, wouldn’t these accidentals be similar to when we expect a word or expression to mean something but, in the end, it means another? Like a homonym, a difficult of hearing or a canceling of a implicature? In fact, in a music staff, everything is predicted by the standard symbols, but we can put at the right of the clef something to warn the hearer that a change is about to happen (like we naturally expect to not understand someone in 100% of what one is saying). Although it doesn’t make sense to compare pitches to words (and we will see why), these “steps” are sufficiently clear to represent that we need minimum differences in sound to create any different meanings whatsoever (not only in the word level).

But, what about the physical interaction with the instrument? If we were really orthodox about how a piece of music should be played, and this is very important, mainly in rock bands that, for example, change their iconic guitar player, the articulations of a song are something crucial to consider. Changes in the brand of an instrument and how the performer wastes (or not) energy to interact with the public, if he prefers to alter some notes (improvisation) or even if he uses fingers or guitar picks to “attack” or “release” the strings might make a band win or lose fans! This starts to make us realize something about the Ontology of music and what is music (and a musician)\(^8\) in the first place. Is it the music score, a specific performance (maybe, the first or the “best” one) or the set of all performances (which are, necessarily, different from one another)? Changing the tone color (the timbre) using different instruments or ways of

\(^8\) We can really put some thought on this concept. Is the musician only that person who is able to write and read a music score? Or those who can actually play or sing? What about the sound engineers who understand a lot of things about acoustics, mix of sounds and equalization? Should we mention those producers who know better than anyone, even better than the artists themselves, about the ideal interaction between a certain kind of music and their public? Let’s not forget the technicians of a concert, who are always ready to help the performer in the stage providing whatever equipment (new strings, drum sticks, replug cables, change instruments) that they need when something goes wrong, which is something that demands knowledge in music. So, is the musician that person who simply knows about music (and the philosopher of music)?
playing something are not things strictly predicted by a music score (although they can actually specify the type of instruments that a performer should use). But, according to Blatter:

Performers make many changes in tone color as a normal part of performing on their instruments without any particular notation so indicating. These subtle, but expressive, changes are almost never specified, and one certainly finds no commonly accepted set of terms or symbols for them. Standard notation for these effects simply does not exist.” (Blatter, 2007, p. 34).

In the theory of *Speech Acts* (TAD), when someone wants to perform an *Illocutionary Act*, he or she has to take into account the (Vanderveken, 1990, p. 113) *preparatory conditions* (e.g. for someone to fire a worker from a company (i.e., *declaring* this worker as “fired”), one must have the proper authority for doing so). Could we compare the differences in music articulations with those preparatory conditions? When a performer changes an electric guitar with 6 for one with 7 strings, isn’t he or she preparing him or herself (and the public) to a new type of (“stronger”, “heavier”) song? Even when they change their clothes or start to use a new accessory, aren’t they gathering information for their interlocutor to understand that they are up to something else in the performance? Surprising someone with new techniques during a live performance of a familiar song is an act of showing that the musician is *prepared* for giving us an unseen perspective of the song (that might divide people’s opinion, by the way). Blatter completes:

Indications of tone quality are frequently limited to the specifying of particular instruments or specific voice qualities. The indication to use or remove mutes, add or remove vibrato, pluck or bow the strings, and so on, provide changes in tone color. In the last one hundred years, composers have called for electronic and mechanical devices, altered vowels, and consonants and alternative bowing, tapping, and plucking locations to expand the range of tone colors (Blatter, 2007, p. 35).

In *harmony*, often referred as the “vertical” study of the music, we deal with the *consecutive combination* of sounds. Differently from melody, in which we deal with pitches one after another (thus, “horizontal”), we need to analyze which notes are acoustically fit to another one. Once again, we deal with metaphors, now
saying that consonant combinations are pleasant to the ears, while dissonant are not. It is really incredible how we have reached this dimension of pleasure in a combination of sounds! These sounds are not, for themselves, conscious. They don't have any compromise or reason to makes us happy or not. However, it is natural for us to always do this personification, granting less abstractness and transforming the process in a more linguistic (human) way. The most famous dissonant interval is the tritone, or the only interval between two pitches in which you can find 3 black keys (thus, 3 whole tones), that is, the F-B interval (see the keyboard figure above in Fig.8). It has been called as Diabolus in Musica⁹ (devil in music or devilish music), because how F and B sound together and/or in a sequence isn’t pleasant, creating an unsettling feeling of chaos, disorder and despair. This dissonant effect has been avoided and prohibited for years by the catholic church and, in the modern times, the heavy metal⁹ (also a metaphor) band Black Sabbath has used this interval by purpose to provoke the Christ believers in a song called as the same name of the band in 1970.

3 A SCIENCE FOR MUSIC THEORY?

Wilkerson tries to develop a “true” music theory in his Harmony Explained (2012), putting, this time, a lot of scientific and computational justification for all things. First, mentioning Michael O’Donnell (a famous music teacher in London) and closing the argument about this matter, he says in section 4.4:

I think that most “Music Theory” as taught in music departments is intended more as a development of descriptive terminology and notation than as an explanatory theory. I think that a lot of humanists don’t understand the difference between description and explanation (…) Amen to that. But what makes me angry is that music teachers do not even MENTION that there is a whole science of the perception of sound and if you would like to know how it really works, then you should talk to them. They basically lie by omission, which is how it ended up taking me several decades to figure all of this out.

---

⁹ You can hear this interval and know more about the historical religious problems involved in this YouTube video called “The Devil In Music” from the channel pianoTV. https://www.youtube.com/watch?v=17YGQ2arlEg

¹⁰ You can understand more about music genres in my other work in Portuguese, called “Meta-melopeia: o som como fonte de conhecimento”, 2020, p. 192.
However, this attempt fails in two dimensions: 1. The music theory that we have nowadays, which is practically the same during centuries, by purpose try not to mention any of the mathematical and physical properties of the sound because these are not aesthetic values. The big difference between art and science is that we make art to feel something while it is executed. Although we can observe some science in it, this is not the primary objective. 2. And, for that, trying to explain music as a mere undulatory and algorithmic phenomenon is not only a demerit to the aesthetic properties of sound, but puts us in the same problem that Wilkerson has tried to solve: musicians, or, simply, people interested in do music, don’t understand (and don’t want to) anything about complex equations, formulas and calculations. While he was unsatisfied with the lack of “exact” explanations (considering that physics, for example, is a satisfactory theory) to the sounds phenomena, he will see that, if we depended on a full explanation about anything that we wanted to learn, we would not even learn how to sum up numbers at the elementary school. Of course it is nice to understand things, but time is limited, and not everyone wants to waste theirs with such justifications just because these explanations are supposed to be “better” than the other ones (of course, someone interested in music might not like this paper). Is the exact science that good? Will it help you to really get a better music perception and turn you into a very good artist who can fill up an entire stadium of fans? This is what music is about!

Let’s take a look at one of his attempts to explain harmony, the concept that we just saw above. Trying to sound modern and not academic, he starts to explain this with an analogy about two people holding a long jump-rope and bouncing it. The movements of the rope are supposed to match our idea of how waves work in physics, but the schemes that we see in books representing them are very unreal and abstract themselves. We should struggle to believe whether these waves in physics are really what they say they are, put aside understand this analogy! The problem is that he inevitably starts to mention a lot of scientific concepts like frequency, numbers, logarithms and measure units, which puts him, again, in the same problem that he wanted to solve: avoid talking about things
that have not been explained yet. Talking about timbre in music, he, once again, starts with an analogy, saying that this concept tells us what is this feeling of a same note sounding differently in different instruments and that this is the thing that distinguishes the *grow* of a lion from the *purr* of a cat (by the way, many languages appeal to onomatopoeias to describe animal sounds, suggesting that these are not actual language sounds, but simple *noises*). But, eventually, he inevitably has to invoke mathematical concepts of “integer multiples of the fundamental”. We don’t even need to go look for these explanations somewhere else. We are simply not interested. This won’t definitely make someone a better artist!

The reader must be asking him or herself: but, if you don’t want rigor, what is the point of this paper, anyway? This is exactly the point: to show that music is not so “scientific” as people say or wanted this to be, based on its apparent exactness (question 6 in the introduction). Actually, it is much more related to our imprecise linguistic interaction than anything else and, despite of all the numbers, Wilkerson highlights at least two important parallels between music and language and how we could make a better theory of music if we respected that. On section 7.2 “The Role of Narrative Generally”, he says that, thanks to the predictive properties of the brain, we can feel that the proportions among notes in melody and harmony help us to see music as a *story* and, even when some notes in a composition don’t make much acoustic sense, it can be justified by the upcoming note, like in a joke, a *punchline*¹¹ (the sentence that actually makes us laugh) only makes sense when we know the whole story before it. Right after in section 7.3 “Embodiment and Emotion”, he tries to promote a kind of *enactive*¹² justification

¹¹ You can understand more about this in my other work in Portuguese, called: “A Filosofia por trás da piada: argumentos lógicos e linguísticos que fundamentam o que nos faz rir dela”, 2019, p. 79.

¹² Enactivism: a theory about how mind and language are dependent on the body relations with the environment. The introduction of the term “enaction” in this context is first linked to Francisco Varela, Evan Thompson, and Eleanor Rosch in *The Embodied Mind* (1992). They have proposed the name to “emphasize the growing conviction that cognition is not the representation of a pre-given world by a pre-given mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs” (p. 9). In this paper, we say that this is half-true, because, although this explains a lot about the common interaction promoted by speaking, considering biological aspects to justify the philosophy behind language would be, in our view, a methodological inversion (science has to come after philosophy), since problems of proper names, reference and definition of a term are more fundamental than
between music and language, listing several “findings” to prove his point: concrete words (onomatopoeias, for example) as blast, blink, freeze, frizzle, crack, yawn and some other things that could resemble the simplicity or even the rigidness or the texture of some things, like the word knee (very “solid” sound) and the word blob (a reference to something very… bloby?) or slime (the “b” to “l” and “s” to “l” sounds allude to something very bouncing and slippery, respectively, as you see your tongue doing very abrupt movements); spatial relations, like the words “through” (the “r” resembles a “trespassing”. Again, another word with “r” to explain the previous word) and “around” (the movement of the mouth saying this makes a circle); and that abstract concepts like harmony and melody are derived, respectively, from the human voice (not by chance, we call the music lines as “voices”) and the general physical movement of things (ascending, descending etc.). Which means, if we were different, with different bodies, the abstract concepts would be too. He completes:

Thoughts and language are not disembodied symbol systems that happen to be realized in the human brain through its computational properties. Instead, thought and language are inherently embodied. They reflect the structure of human bodies and have the inherent properties of neural systems as well as the external physical and social environment (…) Since much human relationship is expressed through voice and movement, is music therefore recalling to us experiences of vocally and physically relating to others and therefore also recalling the associated emotions? (Wilkerson, 2012).

We should not forget about the influence of culture in music and how different behaviors, values and beliefs are reflected in how music affects a certain group of people. What to say about one of the most famous differences between western and eastern music? Mitali Joshi has said in How Does Eastern Music Differ from Western Music? (2012) that eastern music is much more concerned about putting someone in a different state of mind with meditation, choosing for more complexes melodies, which are, naturally, more difficult to memorize, with the use of the microtones (even more divided than the usual 12 notes of western music). Because of this, creating chords (at least 3 notes in good harmony,
consonant), is much more difficult and uncommon in Chinese, Arabic and Japanese music, for example. Western music is oriented around written music, while the eastern is about oral. All of this could mean that people from west are not so interested in aspects of the soul and peace of mind as, at least, are people from east. The relationship between some cultural decisions and how we interact with the world concerning our habits and beliefs and the creation/consuming of music in certain places and times of the history reveals a lot about how language, music and behavior help each other in their meanings.

Could we be talking about a Sociology of music? Tia DeNora brought a very interesting definition about how music can actually be defined as a “way of being”:

I have employed the concept of affordance to describe music’s abilities to, as I put it, ‘get into the action’, its mediating role in relation to social action and experience. To suggest that music may come to afford modes of being and doing, however, is by no means to suggest that music will cause such modes. Rather – and here it is important to remember the lesson from Willis concerning the need to explore connections between music and action from within particular locations – affordances only exist if they are real for some social actor(s). The sociology of music, then, can focus on how affordances are created, how links between music and social life/social experience are forged. As Daniel Cavicchi has put it (in a discussion of my work but throughout his own as well), we need to focus on what music will ‘do for different people’ (De Nora, 2003, p. 170)

And, for a last consideration about science, according to the Internet Encyclopedia of Philosophy on Karl Popper’s article, in fact, it’s not because music theory doesn’t seem to be so scientific, after all we have seen here, that this cannot be the case in the future. In the section “Popper on Physics and Psychoanalysis”, Popper doesn’t like so much the explanations of psychologisms, because it was too dependent on unconscious and unpredictable behaviors, but: “Popper allows that there are often legitimate purposes for positing non-scientific theories, and he argues that theories which start out as non-scientific can later become scientific, as we determine methods for generating and testing specific predictions based on these theories.”, which means that even naive and imprecise theories can be turned into scientific ones. However, does music really need this? As we’ve noticed above, music is much
more interested in making us feel rather than explain itself and, although this could be useful for a systematical composition of songs, this “scientification” would also transform the music into a mere computational product, with no soul, no subjectivity, no background story, no true (human, linguistic, lived) expression, as we have been seeing in A.I. songs. This serves as an answer to our question 1 in the introduction section: music has been created to help us express ourselves differently, more aesthetically, rather than exposing pretentious “exact” concepts.

Knowing that, we will see now the proposal of Deryck Cooke in creating a kind music dictionary, not according to physical formulas, but rather related to the metaphors we can draw from certain sound of notes and scales (sequence of notes).

4 EXPRESSIVITY IN MUSIC

In order to understand the semantics that Deryck Cooke has created in his *The Language Of Music* (1959), we have to consider the chromatic scale that we have seen above and, now, pay attention to some new names that we can give to all these 12 pitches in terms of scale degrees and what emotions Cooke has attributed to them according to some metaphorical criteria, like for example, that the last note of an octave gives us the feeling of “conclusion”, while the middle one gives us a feeling of “stability”, answering our question 3 of the introduction section (Cooke, 1959, p. 89-90):

1. **Tonic**: emotionally neutral; context of finality.
2. **Supertonic**:
   - Minor Second: semitonal tension down to the tonic, in a minor context; spiritless anguish; context of finality.
   - Major Second: as a passing note, emotionally neutral; as a whole-tone tension down to the tonic, in a major context, pleasurably longing, context of finality.
3. **Mediant**:
   - Minor Third: Concord, but a depression of a natural third: stoic acceptance, tragedy.
- Major Third: Concord, natural third: joy.

4. Subdominant

- Normal Fourth: As a passing note, emotionally neutral. As a semitonal tension down to the major third, pathos.
- Sharp Fourth: as modulating note to the dominant key, active aspiration. As “augmented fourth”, pure and simple, devilish and inimical forces.


6. Submediant:

- Minor Sixth: semitonal tension down to the dominant, in a minor context: Active anguish in a context of flux.
- Major Sixth: as a passing note, emotionally neutral. As a whole-note tension down to the dominant, in a major context, pleasurable longing in a context of flux.

7. Leading Tone:

- Minor Seventh: semitonal tension down to major sixth, or whole-tone tension down to the minor sixth, both unsatisfactory, resolving again down to the dominant: ‘lost’ note, mournfulness.
- Major Seventh: as a passing note, emotionally neutral. As a semitonal tension up to the tonic, violent longing, aspiration, in a context of finality.

And, knowing this, we can talk about the 16 fundamental scales that could actually convey some basic emotions and the composers should inspire themselves on these “words” (a sequence of notes according to the chromatic scale) to create the type of music they want. These are some basic terms of the music vocabulary:

Ascending 1-2-3-4-5 (major): growing up from tonic to dominant by major third (interval between two notes formed by two whole tones) is to express an extroverted, active and assertive emotion of joy (Id. Ibid. p. 115).

According to the author, there are even synonymous scales, like the words in natural language. Following his classification:

Ascending 5-1-2-3 (major): jump from dominant to tonic, then to the major third, with or without the intervening second, is equally expressive of an extroverted emotion of joy (Id. Ibid. p. 119).
Ascending 1-2-3-4-5 (minor): when we replace the major third with the minor third, we should expect to find the resulting phrase expressive of an extroverted feeling of pain, an assertion of lament, complaint, a protest against bad luck (Id. Ibid. p. 122).

Ascending 5-1-2-3 (minor): the minor version expresses pure tragedy, because it focuses on the minor third. And, moving up steadily and decisively from the lower dominant, through the tonic, to the minor third, provides a great sense of courage, in which one is aware of the existence of the tragedy and jumps towards it (Id. Ibid. p. 124-125).

Descending 5-4-3-2-1 (major): if decreasing in height is expressing an eminent emotion (into the listener), diminishing from dominant superior to the resting point, the tonic, by the major third, will naturally communicate a sense of experiencing joy passively, that is, accepting or well receiving blessings, relief, consolation, reassurance or completeness, along with a sense of “having arrived home” (Id. Ibid. p. 130).

Descending 5-4-3-2-1 (minor): replacing the major with the minor third in the descending progression, we have a phrase that has been used a lot to express an “eminent” painful emotion, in a context of finality: acceptance or giving in to grief; discouragement and depression; passive suffering; and despair connected with death (Id. Ibid. p. 133).

1-2-3-2-1 (minor): grounding a theme on the tonic, only moving away to the minor third, and returning immediately, is “looking at the dark side of things” in a context of immobility, neither growing to protest, nor going back to acceptance. (Id. Ibid. p. 140).

5-6-5 (major): the major sixth is often used as a melodic base stone in the rise from the dominant to the tonic, in which case its expressive identity is mixed into a general pattern with the “optimistic” major seventh (…) and the effect is a simple assertion of joy (Id. Ibid. p. 143).

5-6-5 (minor): the main and almost only expressive function of the sixth minor is to act as a bridge on the dominant, giving an effect of an explosion of anguish. (Id. Ibid. p. 146)
The author also highlights that some scales could be used to express very specific concepts, although this goes against the general rule of expressing only basic emotions:

1-2-3-4-5-6-5 (major): it is almost always applied to express the innocence and purity of angels and children or some natural phenomenon that has the same qualities in the eyes of man (Id. Ibid. p. 153-154).

1-2-3-4-5-6-5 (minor): the minor version (…) clearly expresses a powerful assertion of fundamental unhappiness – the 1-3-5 “protest” extended to the “anguish” of 5-6-5 (Id. Ibid. p. 156).

8-7-6-5 (major): descending from the (octave) tonic to the dominant, passing through the optimists seventh and sixth major, it is to express an eminent emotion of joy, an acceptance or good reception of comfort, consolation or completeness (Id. Ibid. p. 159).

8-7-6-5 (minor): descending from the tonic to the dominant, passing through the “sad” minor seventh and “distressing” minor sixth, it is clearly expressing an eminently painful emotion, an acceptance of or yielding to mourning; passive suffering; and despair connected with death (Id. Ibid. p. 162-163).

This answers our question 4. It is natural to imagine, though, that these rules couldn’t actually convince a lot of people, maybe because the concept of art itself doesn’t involve any restriction in its definition. So, to answer the question 2 of the introduction, if we can barely create some semantics, it would be much worse to have any kind of pragmatics of music and we could say that music theory, not so precise as we have thought, could represent a syntax to music.

What does make a work of art anyway? Tired of rules, the human being has created the art to express him or herself without any constraints of the physics, chemistry, mathematics and other exact sciences. And, if we realize that, for some reason, we would need to transform music into a bunch of rules, which is totally not adequate to the fact that, actually, anyone can enjoy it, the idea of being an artist and, maybe, the idea of being human itself, is going to disappear.
5 CONCLUSION

We have, now, good evidences of a greater similarity between music and natural languages, more than to an exact science. These metaphors, however, are not here to prove that music is just an imitation of human actions and language, but a manifestation of human natural will of expressing themselves, regardless of the way that they’re doing it so. Of course that we could start to do the opposite and speak about natural languages in terms of music theory concepts and symbols, although it would not be necessarily useful. Music and Language are, in fact, instances of a same purpose and of the inherent tendency of the human mind’s intentionality. Music is a total legit mean of expressing ideas, but in a “prettier” way, so to say. It is able to say less concepts, maybe two or three during a whole 5-10min song, but in a very eccentric, new, addictive, sparkling, contagious, heroic, depressive, sad, compound, majestic, that is, whatever-metaphors-you-want-to-use way.

We hope that this paper might have clarified a lot of academic issues concerning the philosophy of music and might have inspired teachers in the field of music teaching to avoid simply expose the content without proper justifications. This work, however, was not sufficient to demonstrate that we cannot have a pragmatics for music and we haven’t showed precisely how would it be the reverse job of expressing natural languages in terms of music notation, according to the conclusion that music and language are supposed to be interchangeable. Future works should occupy themselves in developing a kind of “music in use” and propose a kind of “music acts theory” and an alleged musical illocutionary act.
REFERENCES


