ON THE FLEXIBILITY OF THE PERCEPTION AND COGNITION OF BROKEN HARMONIC EXPECTATIONS

Carlota Maria Pagès-Portabella

TESI DOCTORAL UPF / 2023

DIRECTOR DE LA TESI

Dr. Juan Manuel Toro, Center for Brain & Cognition

DEPARTAMENT DE MEDICINA I CIÈNCIES DE LA VIDA

Universitat Pompeu Fabra Barcelona

A tota la gent valenta que es dedica a la ciència,

estem fets d'una altra pasta

(bueno, i als meus gats).

To all these brave people who devote their lives to science,

we are a different breed

(well, and to my cats too).

Acknowledgements

I would like to take this opportunity to express my gratitude to my supervisor, Juan Manuel. Under Juan's supervision I have blossomed as a researcher; thanks to him, the dream of investigating the musical brain has come true. Juan, your endless kindness, your open-mindedness toward my ideas, your patience and dedication are priceless. Thank you for bringing me back to earth when research and writing became too hard. Thank you too for supporting me unconditionally in my darkest times and for making me feel a valued member of the research group.

I also want to thank all the members of the LCC group. Irene, Júlia in the early years, Ferran in the last years, thank you for all your help. Thanks Mila, Miriam, Evangelia, my master's students whose effort made possible our studies. But especially Sander, for the endless hours that we worked side by side, for all that you've taught me, for your laughter and energy and for the beautiful friendship we build. And of course, thanks Paola, for all your support and for being "a rock" for the group, a reference of perseverance.

I need to thank David Quiroga for all the help, methodological guidance, stimulating discussions and scientific advice that he's shared with me selflessly, simply for the sake of shared research interests.

También quiero agradecer a mis amigos del CBC, con quién hemos crecido juntos en el difícil mundo de la academia. Irene, Chiara, Camille, Anira, Daniela, Marc Gimeno, Alice. Xavi i Sílvia, gràcies per fer possible la recerca sempre amb la millor disposició, ens feu sentir a casa. Gracias Ege y Gonzalo por vuestro cariño y apoyo. Gracias Ane por ser

una referencia sobre cómo tomarnos las cosas con calma y escuchándonos primero a nosotros mismos. Y, last but not least, gracias Konstantina por construir una amistad para toda la vida (con Dimitris ahí, también!). Codo con codo iniciamos la aventura de empezar un doctorado y de independizarnos juntas, con los miedos y retos que eso conlleva, pero contigo siempre tuve la confianza que todo iría bien.

Gràcies als que sempre hi heu estat i sempre hi sereu. Gràcies als meus pares, que m'han donat tot el seu suport incondicional en aquesta empresa napoleònica que mai han acabat d'entendre. Gràcies Josep per donar-me el do de la música i insistir-me en què valia la pena dedicar-m'hi. Gràcies Carmina per fer-me confiar en mi i en les meves capacitats quan jo no n'era capaç. Gràcies als meus amics que puc comptar amb una mà perquè no me'n calen més. Gràcies Sabina, per ser la germana que mai vaig tenir. Gràcies Silvi, per ajudar-me sempre a veure el *meu* camí. Gràcies Nat, per ser el meu company de camí i la meva llar, per fer-me creure de nou en la bondat i per fer-me il·lusionar pels projectes (musicals i no) futurs junts.

I, per acabar, com no pot ser d'una altra manera, gràcies a la música, que m'ha fet sobreviure a una tesi doctoral.

Abstract

Music is present in all cultures. Despite shared auditory abilities, the experience of music changes depending on cultural context, variations in day-to-day exposure or musical expertise. Such factors shape listeners' preferences and even manifest at the neurophysiological level. One of the main topics in music cognition is the study of harmonic expectations arising from the hierarchical relationships between tones. Tension and release patterns lead most listeners to expect musical resolution. If expectations are not fulfilled, our brain responds automatically and elicit a set of well-studied neural responses. Importantly, experience might lead to different harmonic expectations. For example, a trained jazz musician might find certain harmonic combinations acceptable and pleasant while a trained classical musician or a non-musician might not. The present dissertation aims to explore the effect of musical training on the perception and cognition of harmonic unexpectedness, as well as the shaping power that sensory and psychophysical factors may have. In a series of behavioural and neurophysiological studies, we found that musical training shapes the affective and cognitive preferences of the listener toward harmonic unexpected resolutions as well as the underlying neural responses. More specifically, we show that musical training strengthens and refine the event-related potentials toward musical irregularities, which in turn interact with the physical and contextual properties of the musical stimuli. Together, these findings extend our understanding on how musical expertise shapes the perception and cognition of musical unexpectedness.

Resum

La música està present en totes les cultures. Tot i tenir les mateixes habilitats auditives, la manera en la què experimentem la música depèn del context cultural, la música que escoltem en el dia a dia o si hem rebut entrenament musical. Aquests factors modulen les preferències dels oients i, fins i tot, es manifesten a nivell neurofisiològic. Un dels temes centrals en recerca sobre cognició musical és l'estudi de les expectatives harmòniques que sorgeixen de les relacions jeràrquiques entre notes. Els patrons de tensió i resolució porten a la majoria dels oients a esperar una resolució. Si això no es compleix, el nostre cervell respon automàticament i genera un seguit de respostes neuronals molt ben caracteritzades. Cal destacar que l'experiència pot modificar aquestes expectatives harmòniques. Per exemple, els músics de jazz probablement trobaran acceptables i agradables certes combinacions harmòniques que un músic clàssic o una persona sense entrenament no acceptarien. Aquesta tesi pretén explorar l'efecte de l'entrenament musical en la percepció i la cognició de resolucions inesperades d'expectatives harmòniques, així com també l'efecte de factors psicofísics i sensorials. En un seguit d'estudis conductuals i neurofisiològics, hem trobat que l'entrenament musical pot modular les preferències afectives i cognitives dels oients respecte a irregularitats harmòniques així com les respostes neuronals subjacents. En concret, mostrem com l'entrenament musical reforça i refina els potencials evocats per irregularitats harmòniques i que aquestes respostes neuronals interactuen amb les propietats físiques i contextuals dels estímuls musicals. En conjunt, els nostres resultats amplien la nostra comprensió sobre com l'expertesa musical dóna forma a la percepció i la cognició musical.

Preface

I have been interested in music for as long as my memory goes. As a child, one of my favourite games would be to play by ear on the (toy) piano the songs that I loved from TV shows. Back then, my dad used to play the guitar in a punk-rock band and, at some point, he wanted to teach me some basic chords. But by then I was a teenager and, of course, "I wasn't interested in music anymore". Fortunately, I changed my mind and started learning guitar and studying music. At the same time, I became fascinated by the brain when I learned in high school how electrical signals were transmitted from one neuron to another. I now realize that my scientific curiosity was born in those days. From there, questions about how the brain worked and how it allowed all animals to understand the world that surround us only grew bigger. It was a matter of time that I ended up in the Centre for Brain and Cognition.

However, even today, one of the comments that I get the most when I tell what I do for a living (as a PhD fellow in music cognition) is: "Fascinating! You studied biology and music, such two separate worlds, they have nothing to do with each other!". This was too my believe during my early years as a student (Biology by day and music by night). Fortunately, a teacher from uni told me that one could do research on any topic and that it existed a world in which disciplines fused with each other. That day I discovered that my passions could coexist together through scientific research.

At that time, I was in the last years of my training in jazz and modern music, where I learned from experience about plasticity in music: the rules of music can be bent on purpose, music perception changes depending on the listener, training (and culture) changes the way we understand sounds and music, and many other examples. I was a very In a way, I became obsessed with the fact that, an experience as universal as listening to music, could be so different for each of us. I learned that, even if humans are born with the same auditory wiring and abilities, experience plays a huge role how we experience (listen, play, get moved by, compose, share, dance to) music. Therefore, the moment I was given the opportunity to pursue a PhD in music cognition, I decided to test some of these ideas out on my own.

The aim of this dissertation was to understand to what extent the perception of musical unexpectedness is plastic. I chose to study musical unexpectedness because it is one of the musical events that stimulates me more intellectually as a musician. Thus, we aimed to understand to what extent musicians (as opposed to non-musicians) would be engaged by it, characterize the neural mechanisms that govern its processing and identify which factors can shape that processing. To do so, we first reviewed the current theories on music perception and cognition, preferences, emotions, and the role of experience in them (including cross-cultural and training-derived differences). To extend that knowledge, we performed a series of experiments where trained and untrained listeners were tested in different scenarios. In the first experimental section, we evaluated listeners preferences towards different types of unexpected endings of harmonic sequences, based on affective and cognitive evaluations. In the second experimental section, we ran three electrophysiological studies where we investigated the neural correlates behind the processing of harmonic unexpectedness. In each of these studies, we focused on one factors that might modulate the processing of unexpectedness, such as the type of unexpected event, the amount of expectation that is broken and the accumulated previous evidence of upcoming unexpected events. After the experimental section, we discuss the main findings of our studies and connect them to the current literature and some possible future lines of research. The last chapter of this dissertation closes the thesis by reviewing the main conclusions from our work.

1. INTRODUCTION

The study of music processing is an ideal tool for investigating the human brain because it is a process that engages a large amount of perceptual, cognitive and affective processes (Koelsch, 2011). It is not surprising that music has long been a topic of interest for psychologists and neuroscientists. One of the most popular questions around music processing are training-induced effects, because they provide a means to examine the different roles of nature and nurture on the encoding of music (Bidelman, 2013). Moreover, the variety in musically acquired skills is very useful to study the effects of training on brain functioning, as musicians represent a unique model in which to study plastic changes in the human brain. On the other hand, studying brain organization and functioning may reveal to what extent and at what level music processing recruits neural networks distinct from those involved in other auditory-vocal functions, such as language (Peretz & Zatorre, 2005).

1.1 Fundamental musical concepts

To study music cognition and perception, it is essential to understand the building blocks of music, which will be reviewed below. Music is based in changes of sounds across multiple dimensions such as pitch, timbre, rhythm and intensity. Musical events acquire meaning through their relationship to other events to become part of a larger structure (Jentschke, 2007). For the scope of the present thesis, we will mainly focus on the physical dimension of pitch and the more complex musical dimensions that derive from pitch: melody and harmony.

1.1.1 Pitch and intervals

Pitch is the perceptual correlate of the periodicity of sound waves. Each pitch corresponds to a fundamental frequency (F0) and a series of harmonics whose frequencies are integer multiples of the F0 (Figure 1). Musical pitches are tuned to certain frequencies that are specific of each musical context and are called *notes* or *tones*. The relative pitch between successive notes relate to each other – for instance, whether a note is higher or lower in pitch than the previous note (McDermott & Oxenham, 2008). That distance between pitches (either sounding at the same time or sequentially) defines *musical intervals*. For instance, the interval of an octave is the distance between two pitches where one has half of frequency of the other. The intervals that appear in the harmonic series of any F0 (Figure 1, in respect to the F0 they appear in this order: octave, perfect fifth, major third, perfect fifth, minor seventh, etc.) originally served as musical anchor points that eventually lead to the standard western musical scale (Gill & Purves, 2009).

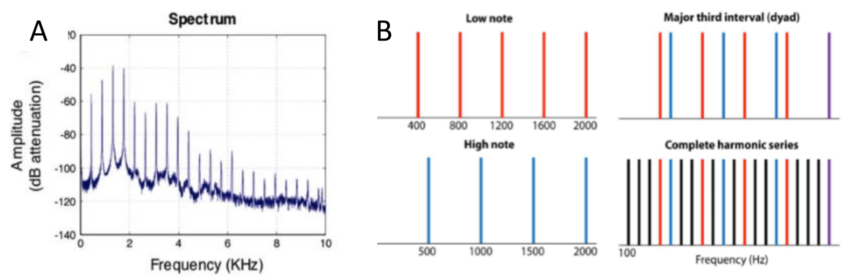


Figure 1. (A) Spectrum of the note A440 played on an oboe, where the peaks are at integer multiples of the F0, which is characteristic of a periodic sound. The frequencies in the note are all integer multiples of the fundamental frequency of 440 Hz, and as a result are regularly spaced along the frequency axis (adapted from McDermott & Oxenham, 2008). (B) Harmonic structure of a musical interval compared to a harmonic series (adapted from from Gill & Purves, 2009).

1.1.2 Scales and the hierarchy of stability of tones

In the 17th century, the need for a universal tuning for all instruments in western culture established the current musical standard scale (Hagenow, 2005): the equal-tempered scale. The tempered scale is based on twelve equal divisions of an octave, called *semitones* (referred to with the labels C, C# or Db, D, D# or Eb, E, F, F# or Gb, G, G# or Ab, A, A# or Bb, and B). From these twelve available tones in the tempered scale, different subsets of 7 tones determine different scales (Brattico et al., 2006). For instance, the major scale contains the intervals of major second and third, perfect fourth and fifth, and major sixth and seventh (in C major, C, D, E, F, G, A, B, respectively). In any scale, a central tone attracts all the other tones and constitutes the *tonal centre,* while the combination of tones surrounding it, constitute the *tonality* or *key*. The tones within a key form the *diatonic scale* and are organized according to the first level of a *hierarchy of stability* (Bharucha & Krumhansl, 1983): certain tones are more stable than others depending on how strongly related they are to the tonal centre (Figure 2). Moreover, tones outside the tonality are the least stable ones and tend to demand resolution to more stable tones, normally, in-key tones (Figure 2). Thus, when tones are combined into sequences to create *melodies*, each note forms harmonic and rhythmic relationships with the earlier ones (Minati et al., 2008). In this context, some have a more central musical function and are often placed at the beginning and the end of musical pieces, because they serve better as completions of musical phrases (Bharucha & Krumhansl, 1983). Melodies constitute the horizontal dimension of music. But, in tonal music, notes normally sound simultaneously. That constitutes the vertical dimension of music, broadly known as *harmony* (Tramo et al., 2001).

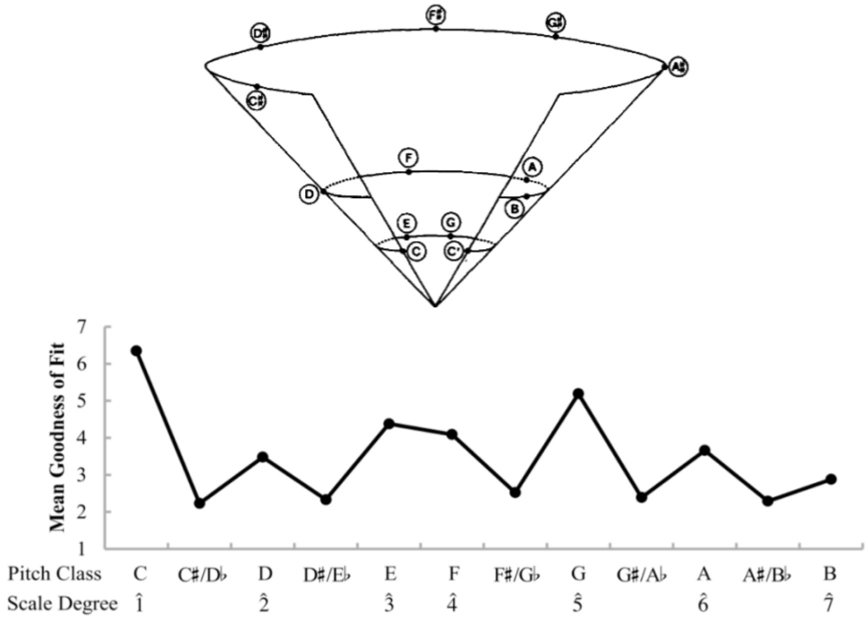


Figure 2. (Top) Idealized three-dimensional conical representation of tonal centre that depicts attracting forces toward the surrounding notes (adapted from Krumhansl, 1979). (Bottom) Krumhansl & Kessler's (1982) major-key profile, which represents the doofness of fit ratings of the 12 pitch classes following an authentic cadences, a good representation of the hierarchy of stability (adapted from Sears et al., 2018).

1.1.3 Chords

The combination of three or more tones played simultaneously form *chords* (Figure 3), that are usually heard as a single fused auditory entity (termed homophony, Tramo et al., 2001). To simplify, combining a major third (2 tones) and a perfect fifth conforms a major chord while combining a minor third (1 & ½ tones) and a fifth conforms a minor chord. Over each tone of the diatonic scale, chords are build to form a diatonic sequence of major of minor chords. For example, in the C major key, the chord built on the first degree of the scale consists of the tones *c, e* and *g* (and form the C major chord), the chord built on the fourth degree consists of the tones *f*, *a* and *c* (and form the F major chord), while the chord built on the sixth degree consists of the tones *a*, *c* and *e* (and is the A minor chord), where the C chord represents the tonal centre. Although single tones do not belong unambiguously to a specific chord, each tone in a chord hierarchically relates to each other (e.g, in the triad of C, the tone *c* is more stable than the other tones *e* and *g*), which represents the second level of hierarchy of stability. Moreover, each chord within the tonality is more or less stable depending on how much harmonically related they are to the tonal centre (which, in turn, depends on how many tones they share). Thus, in C major, the C chord is more stable than the F chord which, in turn, is more stable than the G chord. This represents the third level of hierarchy of stability and is known as *harmonic hierarchy* (Bharucha & Krumhansl, 1983). Harmonic hierarchy determines the function that each chord performs within the key. For instance, the chord C serves as the referential point within a musical passage and performs the function of *tonic* (commonly noted with the Roman numeral "I"). The other chords of the tonality have a well-defined relationship to the tonic that determines their *tonal function* (Krumhansl, 1979): for instance, the chord F performs the function of subdominant (noted as IV) and the chord G represents the dominant (noted as V; Bigand et al., 2003). The function of the subdominant is usually to lead to the dominant, which is the more unstable chord of the tonality and demands resolution to the tonal centre. Other chords (such as the III and the VII) are less essential for establishing the key (Figure 3).

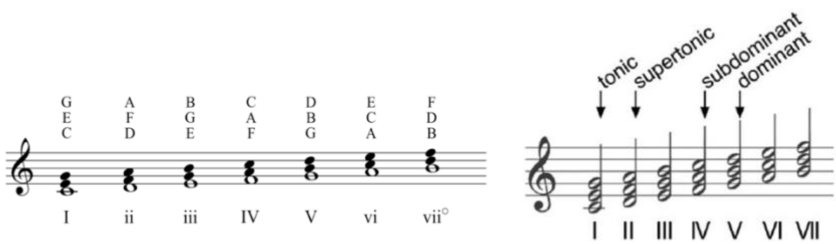


Figure 3. (Left) The seven diatonic triads for the key of C major, with pitch classes above. Open (whole) notes indicate the degrees of the scale (adapted from Sears et al., 2018). (Right) Illustration of tonal chord functions denoted with Roman numerals (adapted from Koelsch & Jentshchke, 2008).

1.1.4 Combination of chords into sequences

When chords are combined into sequences (known as *chord progressions*) the hierarchy of stability becomes evident. Western tonal-harmonic music essentially relies on the conventionalized use of certain chord progressions or *cadences* (Jentschke, 2007), where the alternation of different stable and unstable chords creates patterns of tension and release. Each tone or chord transition has a relative probability that organizes musical events (and the subsequent expectations) across time. As they hear tone or chord successions, listeners make predictions about the likely tones or chords to follow based on their relative stability within the key (Figure 4). First, the tonal centre is extracted and, in relation to it, the harmonic distance of subsequent chord functions is calculated, which determines the relative stability of each chord. Musical context primes those chords that are more related to the context. It is not surprising that most cadences begin and end on the tonic, which is an unambiguous expression of tonality (Krumhansl, 1979). Based on tonal hierarchy, the probabilities of each chord transition are calculated. Then, based on these transition probabilities and tonal representations stored in long-term memory, listeners make predictions of chords to follow (Bharucha & Stoeckig, 1986, 1987; Koelsch, 2009a; Koelsch et al., 2000). However, harmonic expectancies do not only occur sequentially from chord to chord, but also depend upon the harmonic function of the chord in the extended temporal context (Regnault et al., 2001). Some authors suggest that the probability for each chord transition emerges from mathematical principles that might represent abstract, rather than physical (or acoustical) features (Woolhouse & Cross, 2010). The interplay between expectancies, as they unfold over time, and the varying degrees to which they are fulfilled or violated are fundamental for music composition and experience (Koelsch et al., 2000).

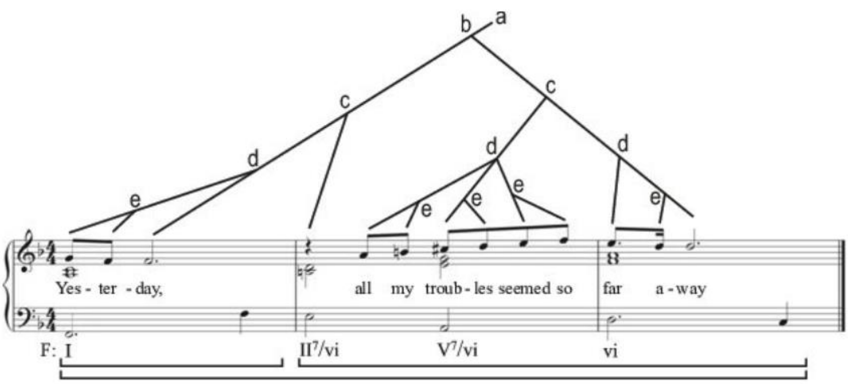
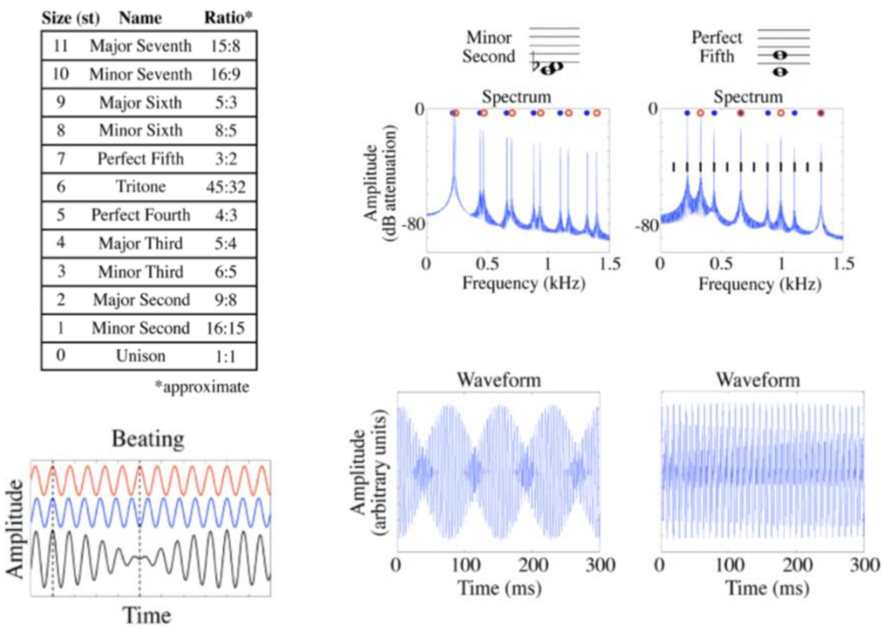


Figure 4. Prolongational tree for the first musical phrase of Yesterday, where relative stabilities and dependencies are represented. At the bottom, tonal functions are indicated with Roman numerals (adapted from Lehrdal, 2013).

1.1.5 Consonance-dissonance

Besides the hierarchical relationships established *between* chords, of high importance in music are the relationships that tones establish *within* chords. Depending on that relation, chords will be considered more or less consonant or dissonant. Consonance and dissonance have been a question of interest for centuries and mathematical, physical, psychoacoustic and neurobiological theories have contributed to their understanding.

Dissonance was initially defined in the 5th century B.C. by Pythagoras as intervals bound by complex frequency ratios (16:15), while consonant combinations were related by simple ratios (1:1, 2:1, 3:2). However, strict acoustic theories were not enough to explain consonance in later musical practice, as aesthetically consonant chords could be mathematically dissonant, such as the perfect fifth (442:295) of the tempered scale (Bidelman, 2013). Von Helmholtz (1877) associated the feeling of dissonance with *roughness* and of consonance, with the absence of it. Roughness arises when simultaneously sounding tones are very close in frequency, so that the resulting wave has amplitude variations (called *beats*; Figure 5). When that occurs, each tone interferes with the other's perception by auditory masking because their amplitude envelopes overlap in the basilar membrane and are thus difficult to be encoded into the auditory nerve (Deutsch, 1982; Plomp & Levelt, 1965). The smallest frequency difference between two tones such that they can be heard separately is known as critical bandwidth and a quarter of it corresponds to the most dissonant interval, the minor second (16:15). In his famous theory, Terhardt highlighted the distinction between *sensory dissonance* and *harmony* (Terhardt, 1984). He defined sensory consonance as the absence of annoying features such as roughness, which reflects the built-in constraints of a poor spatial resolution of the basilar membrane (Krumhansl, 2000; Peretz & Zatorre, 2005). In a musical context, though, the term *sensory dissonance* can be used in a broader sense to refer to learned associations and to low pitch commonality between successive chords (Koelsch et al., 2007; Parncutt, 1989). Meanwhile, harmony would be the music-specific component of consonance that governs pitch relationships (Virtala & Tervaniemi, 2017a).



Size (st)	Name	Ratio*
11	Major Seventh	15:8
10	Minor Seventh	16:9
9	Major Sixth	5:3
8	Minor Sixth	8:5
7	Perfect Fifth	3:2
6	Tritone	45:32
5	Perfect Fourth	4:3
4	Major Third	5:4
3	Minor Third	6:5
2	Major Second	9:8
1	Minor Second	16:15
0	Unison	1:1

*approximate

Figure 5. (Top left) Intervals with their size in semitones, their common names and their corresponding frequency ratios. (Bottom left) Two sinusoids of different frequencies are plotted in red and blue; their superposition (in black) contains amplitude modulation known as "beating". (Top right) Spectra for the minor second and perfect fifth as sinusoidal tones. Red and blue circles denote the frequencies belonging to each note. The frequencies of the fifth are approximately harmonically related (black lines denote harmonic series). (Bottom right) Waveforms of the minor second and the perfect fifth. Amplitude modulation (from beating) is evident in the waveform of the minor second, but not the fifth (adapted from McDermott, 2010.

Thus, sensory consonance could be considered a low-level psychoacoustic phenomenon that is not specific to music but that serves as the neurocognitive basis for harmony.

In other words, sensory and musical dissonance reflect built-in auditory constraints and learned associations, respectively (Peretz & Zatorre, 2005). However, consonance cannot be explained simply as the absence of roughness, because some intervals wider than the critical band are also perceived as dissonant. More recently, dissonance has been defined as the lack of harmonicity (Figure 5), which is how much the F0s and harmonics of all the tones in a chord belong to the same harmonic series (McDermott et al., 2010). This theory could explain dissonant percepts in dichotic listening, which cannot be explained by inner ear mechanisms related to roughness (Bidelman & Krishnan, 2009). Recent studies investigated the neurobiological basis of consonance, which will be addressed in the following sections of the introduction ("*1.3 From sound to music in the brain"*).

1.2 The perception of music

1.2.1 Acquisition of the implicit knowledge of music

All listeners possess an implicit and cultural-specific knowledge about the regularities of their musical context. The acquisition of the knowledge about tonal regularities (and the subsequent processing of musical information) is the result of the combination of the initial predisposition of the human brain for music processing together with the passive exposure to musical stimuli in everyday life (Bigand, 2003;

Tillmann et al., 2000). On top of that, there is musical training, which amplifies implicit learning via the acquisition of an explicit set of rules.

Humans have an exceptional capacity for implicit learning, as seen in the fact that new-borns show a preference for musical pieces that they heard during late pregnancy (Jentschke, 2007). The sensitivity to universal aspects of spectral and temporal structures (such as the sensitivity to consonance) emerge as early in development as 4 months of age (Zentner & Kagan, 1998). However, before they fully acquire the regularities of their own musical culture, infants sometimes show general abilities unaffected by culture. For instance, 8-month-old infants detect equally well the introduction of in-key and out-of-key tones in melodies, while adults more readily detect out-of-key tones (Trainor & Trehub, 1992). Also, 9 to 11 months old infants detect equally well changes in a melody based either major or augmented triads, whereas adults discriminat better the changes on the melody based on major triads, i.e. the more prototypical for Western melody (Trainor & Trehub, 1993). Then, during the first years of life, due to everyday exposure infants rapidly acquire the sophisticated musical regularities of their own context, through statistical learning in a way similar to the acquisition of language (Saffran, 2003). In that learning process, innate predispositions allow young listeners to sort through the noise to find the signal, where the perceptual system weights some cues more highly than others. Five-month-old infants easily detect deviations in melodic contour (Trehub et al., 1984) and 7- to 10-month-old infants more readily detect semitone changes in typical Western melodies compared to a change in non-typical melodies (Trehub et al., 1990). These results suggest a very early sensitivity to phrase structures, which may help them to acquire tonal structure. During the first year of life, they acquire the interval structure of musical scales (or, in other words, the reference system), around which they can build regularities of the musical context (Lynch & Eilers, 1992). Finally, implicit knowledge of key membership is present in 5-year-olds (Trainor & Trehub, 1994) and by 7-year-old (but not 5-year-old), children show superior memory for melodies conforming to the rules of tonal music (Sloboda, 1985). Thus, system-specific responses emerge later as a result of enculturation: first, knowledge of key membership appears and knowledge of harmony is observed last (Hannon & Trainor, 2007; Trainor, 2005). It could be said that increasing infants' exposure to the music of their own culture amplifies the influence of culture-specific factors and attenuates the effects of culture-general factors, which is similar to perceptual narrowing in language acquisition.

1.2.2 Processing and judging harmonic hierarchies

Western adults, even those without formal musical training, develop sophisticated musical abilities (Bigand, 2003). For instance, western listeners tend to interpret the first chord or tone of a sequence as the tonic and perceive the rest of the chords depending on the distance to the tonic (Jentschke, 2007; Krumhansl et al., 1982; Tillmann et al., 2000). After establishing a key context, diatonic tones are rated as better fit than non-diatonic tones, the highest ratings being for the tonic, third and fifth degrees, respectively (Krumhansl & Kessler, 1982). The same applies for chords, which are interpreted depending on their harmonic stability and the function that they perform within the context (Krumhansl et al., 1982).

Naïve listeners are able to perceive musical tensions and relaxations, anticipate musical events based on syntactic features (Bigand, 2006) and detect when they are irregular, even if these are subtle harmonic changes, such as a cadence resolving in the subdominant (Bigand, 2003). Such implicit processing of harmonic rules is done independently of attention or any conscious evaluations of musical stimuli (Minati et al., 2008). Importantly, perceived harmonic stability is negatively correlated with perceived musical tension. In general, the most important chords in the hierarchy create weaker musical tension, and thus the tonic is judged as less tensioning than dominant and subdominant chords. Moving away from the tonic centre (for instance, the movement from the supertonic to the dominant) is perceived as tensioning and demands resolution by going back to that tonic, which releases that tension (Koelsch, 2009a; Koelsch et al., 2000). Musical fragments ending on very stable chords, indicate that the musical process has reached some point of arrival, whereas ending on unstable chords evoke the feeling that there will be a continuation of the sequence (Bigand & Parncutt, 1999) and are perceived as unexpected (or erroneous, depending on the listener), as if there was no sense of completion or finality (Tillmann et al., 2000). The stronger the violation that a chord introduces, the easier that listeners detect it (Koelsch et al., 2000; Koelsch & Friederici, 2003). Tension judgements are further influenced by sensory dissonance (where minor and seventh chords are rated as inducing higher tension than major chords), pitch commonality between successive chords (the weaker it is, the greater the perceived tension) and horizontal motion (the larger the melodic interval in each voice, the greater the perceived tension; Bigand et al., 1996), which is related to melodic anchoring (the smaller the pitch distance between an unstable and a stable chord, the better the resolution of the tension created by the unstable event). Moreover, when a temporary change in key - a *modulation* (Bharucha & Krumhansl, 1983) - occurs, a return to the main key is perceived as a departure from the local tonic, generating tension rather than relaxation. Apparently, a single harmonically unrelated chord is sufficient to disrupt the influence of the global context of a piece (Bigand & Parncutt, 1999).

One of the most popular methodologies to explore how a previous context influences the processing of a target chord is harmonic priming. Prime contexts activate the listener's knowledge of tonal hierarchies and lead them to anticipate events belonging to the same key (Jentschke, 2007). Different studies investigated the influence of priming with single chords or long chord sequences and the mutual influence of global and local contexts (Tekman & Bharucha, 1998; Tillmann et al., 1998, 2003). According to the literature, western listeners are faster and more accurate in judging a feature about a target (e.g.: whether they are in-tune/out-of-tune or consonant/dissonant) when it is harmonically related to the prime, either by local (Bharucha & Stoeckig, 1986, 1987; Bigand et al., 1999) or global (Bigand & Pineau, 1997) context relationships. Response-time patterns reflect chord ranking according to tonal structure, with faster processing for tonic chords, followed by the dominant and the subdominant (Park et al., 2018; Tillmann et al., 2000). Processing the tonic chord is always faster than processing harmonically less related chords, even if they are specifically primed in the context, which demonstrates the predominance of harmonic over sensory priming (Bigand et al., 2003). If priming relied on sensory mechanisms, processing the primed chord would be facilitated over the most expected tonic (Jentschke, 2007). Harmonic priming performance is also affected by attention and musicianship. When given an orthogonal task (to focus only on melodies), the performance of musicians is slower when melodies are accompanied by unexpected chords whereas the performance of non-musicians is unaffected. In this setting, musical training may facilitate the responses to harmonically expected chords by enhancing the automatic formation of harmonic expectations, all independently of attention. Meanwhile in non-musicians, the detection of violation of harmonic expectation depends on attention (Loui & Wessel, 2007).

To summarize, listeners possess an implicit knowledge of hierarchy of stability, which also affects how they perceive the rest of musical features of western music. However, aesthetic judgements of musical structure do not necessarily translate directly to aesthetic emotions and aesthetic preferences.

1.2.3 Emotions in music

Traditionally, in Western music the perceived mood of a certain piece is related to its *mode*. The mode of a musical piece is defined by a specific combination of chords in a tonality, where a major mode starts with a major tonic and a minor mode starts with a minor tonic. Pieces in major key are usually perceived as happier than pieces in minor key. Moreover, tempo, rhythm and intensity (which accentuates certain events) further influences the perception of the affective value of musical pieces (Jentschke, 2007). The minor mode's affective connotation appears early in life (around 5yo) and may have its origin on psychoacoustic features - such as higher sensory dissonance (Parncutt, 2014) or lower-than-expected pitch (linked to the lower pitch of sad speech) - together with learned arbitrary associations build through familiarity and exposure (Virtala & Tervaniemi, 2017b). Importantly, the perception of negative basic emotions (such as recognizing the sad affective connotation of the minor mode) does not necessarily correspond with the induction of negative emotions or with judging it unpleasant or disliked (Brattico & Pearce, 2013). On the contrary, sad music is often liked, considered beautiful and induces the positive aesthetic emotion of enjoyment (Brattico et al., 2016; Virtala & Tervaniemi, 2017b). And vice versa, sad listeners do not necessarily show a preference for major music. Moreover, personal experiences further influence which emotions are evoked by music (Brattico et al., 2016).

Musical aesthetic emotions (such as wonder, transcendence, entrainment and awe) are complex, multi-faceted and qualitatively different from common goal-oriented and utilitarian emotions. Musical emotions start with early affective pleasant or unpleasant sensations (namely sensory pleasure) that derive from perceptual features (Brattico, 2015). Musical enjoyment (or *chills*) is associated with increased subjective emotion and physiological arousal (reflected in an increase of skin conductance), and it is experienced as highly pleasurable. Then, evaluative conditioning induces emotion through association with aversive or rewarding stimuli, which is further influenced by emotional contagion, visual imagery of external referents and episodic memory (Brattico & Pearce, 2013). A peculiar case of discrepancy between sensory and conscious pleasure occurs when a musical stimulus originates unpleasant sensations at peripheral organs (which could be associated with withdrawal reactions), but at the same time produce strong feelings of pleasure mediated by higher-level mechanisms, by activating an association with positive past events or by understanding and identifying the neatness of the construction. Thus, to become a conscious emotion of enjoyment and engender a liking judgement for a musical piece, sensory pleasure must be followed by value attributed mediated by personal associations, knowledge, social constructs and other top-down processes (Brattico, 2015). These processes eventually lead to the development of musical preferences (which, in turn, influence back the perceptual processes, see below).

1.2.3.1 Musical emotions in the brain

Brattico (2015) proposed that the aesthetic enjoyment of music is governed by a bidirectional neural network. A bottom-up route is based on perceptual features (like sensory dissonance or consonance) that innately trigger early sensory pleasure or hatred. The coupling between the auditory cortex and the amygdala is responsible for the aversive, unpleasant sensory experience deriving from the acoustic features of the music, such as dissonance, which is related with withdrawal behaviour and negative affect areas (Brattico & Pearce, 2013; Virtala & Tervaniemi, 2017). Meanwhile, the striatum (right nucleus accumbens) and the caudate nucleus are important for experiencing pleasurable sensations from music, such as chills (Brattico, 2015). In fact, the soothing sensation of consonance may involve reward centres (Sammler et al., 2007). The inverted-U shaped interaction between uncertainty and surprise is even reflected in the amygdala, anterior hippocampus and auditory cortex (Cheung et al., 2019). At the same time, a top-down route influences sensory and cognitive processing in an implicit way by personal and social constructs (such as preferences), which are controlled by prefrontal and associative cortices and increased connectivity between the nucleus accumbens and the inferior PFC. For instance, listening to a disliked musical genre elicits smaller responses than listening to a liked musical genre during descriptive tasks (Brattico, 2015). Moreover, listening to liked music elicits neural activity in emotional and motivational structures of the brain, whereas listening to happy music (as contrasted to sad music), elicits activity in sensory areas Interestingly, chord incongruities elicit different responses in the limbic system in musicians and non-musicians, likely because musicians are trained to express emotions through their playing (Brattico et al., 2016).

1.2.3.2 Musical emotions across cultures

Music is able to express and induce similar emotions in individuals of all ages and cultural backgrounds (Peretz et al., 2013) that are recognized cross-culturally (Fritz et al., 2009). Although listeners find the music of their own culture more arousing and recognize the basic emotions with more accuracy (Virtala & Tervaniemi, 2017b), the affective connotations of the features of music may be similar across cultures (Mehr et al., 2019). Members of the African Mafa tribe (who have no exposure to western music) successfully recognize musical excerpts with happy, sad and scared connotations, although with more variability than Western listeners. Importantly, they rely on tempo (faster=happier) and mode (major=happy / indefinite=sad / minor=scared), although tempo is a more effective affective cue than the association with major-minor modes (Fritz et al., 2009). On the other hand, Western listeners are able to successfully infer non-western song's most common behavioural contexts from their relational acoustic properties such as accent, meter and interval structure (Mehr et al., 2019). Tonal relationships in western music are also very similar to south Indian music, with corresponding affective connotations, similarly to each culture's language vocalizations (Bowling et al., 2012). In short, music across cultures is the product of underlying psychological faculties that make certain sounds (rhythmic or melodic patterning) appropriate to certain moods, desires, social and emotional circumstances (Mehr et al., 2019).

1.2.4 Musical preferences

Our musical preferences are built - at least to some extent - on emotional arousal. The more we are moved by music (independently of it eliciting positive or negative emotions), the more we like it. Especially, when the gap between the expressed and induced emotion is minimal. The formation of music preferences is further affected by personality, age and music training: for instance, openness to experience correlates with preference for complex music, as people with this personality trait is more willing to create unusual associations, especially if they are musicians (Brattico et al., 2016; Vuoskoski et al., 2012).

1.2.4.1 Emotional responses and the development of preferences for consonance and/or dissonance

Sensory consonance and dissonance have been long used by composers in all cultures to manipulate aesthetic responses. Traditionally, in western culture there has been a preference for consonance, which is usually described perceptually as resolved, stable, euphonious, beautiful; whereas dissonance is described as unpleasant, unresolved, unstable, discordant or rough (Bidelman, 2013; Bones et al., 2014). Western adults typically assign consonant intervals with higher status in hierarchical rankings of harmony (Bidelman & Krishnan, 2011; Plantinga & Trehub, 2014; Roberts, 1983). These affective connotations carried by the consonance-dissonance continuum partly rely on psychophysiological cues. For instance, sensory responses to dissonant sounds tend to be coupled with an affective experience of irritation (Brattico & Pearce, 2013), which is reflected in a stronger decrease in heart rate compared to consonant music, especially if musical excerpts are highly arousing (Sammler et al., 2007).

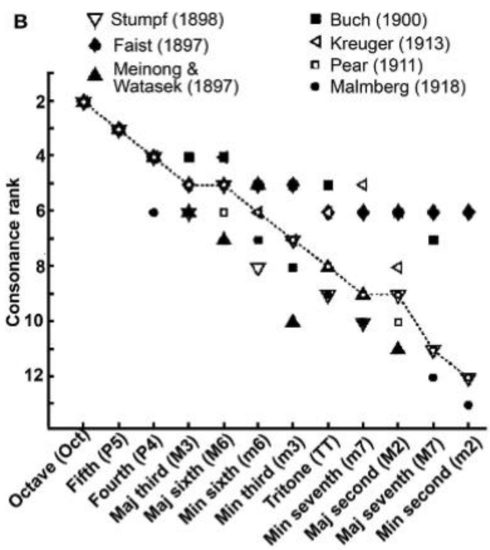


Figure 6. Rank order of musical interval consonance ratings reported across psychophysical studies (adapted from Bidelman et al., 2013).

However, in the last decades it has been a matter of debate how fixed the aesthetic preferences for consonance are (Weiss et al., 2020). In fact, the conception of consonance is an always-changing continuum (Virtala & Tervaniemi, 2017b), as a result of the change in the degree of exposure to dissonance (Weiss et al., 2020). Until the 18th century, seventh chords were considered the origins of all dissonance but then became commonplace in classical music (Johnson-laird et al., 2012). Currently, western musical practice exhibits a variety of sonorities and consonant/dissonant distinctions no longer appear straightforward (Popescu et al., 2019). Harmonic relationships of the consonance-dissonance continuum are not encoded in a strict binary manner but rather are processed differentially based on their degree of perceptual consonance (McDermott et al., 2010). For instance, musicians and non-musicians rank as more consonant and pleasant major chords, followed by minor, diminished and augmented chords (Bidelman & Krishnan, 2011; Roberts, 1983). Pure consonance can even be considered uninteresting in a music context rather than pleasant. In fact, musicians and non-musicians rate as more pleasant intervals such as major and minor sixths/minor thirds (Plantinga & Trehub, 2014)/minor ninth and major seventh chords (Lahdelma & Eerola, 2016; Schön et al., 2005) than perfectly consonant chords such as the octave and perfect fifth. Moreover, despite the negative affective reactions toward dissonance, its role among music genres varies drastically and the associated conscious evaluations, too. Listeners can experience very positive feelings from listening to death metal, which is very dissonant and loud.

There are many factors that further modulate the preference for consonance or dissonance. Although adding more tones to a chord generally increases roughness, many three- and four-tone chords are perceived as more consonant than two-tone chords. That could be related to the fact that parsing complex auditory signals generates a greater dopaminergic reward because tone combinations imply vocal cooperation and social cohesion (Bowling & Purves, 2015). Moreover, as the emphasis on harmony became more prominent across history, the quality of a chord was increasingly based on the role of the chord within a harmonic context instead of its fundamental frequency ratios. Currently, one same chord might be considered consonant or dissonant (and more or less pleasant) depending on the harmonic context in which it is embedded regardless of the presence of beating (Parncutt, 1989; Plantinga & Trehub, 2014; Steinbeis et al., 2006). In other words, with the appropriate surroundings, a dissonant interval can be pleasurable and serve important musical functions (McDermott & Oxenham, 2008). These results can be, at least partly, rationalized by the arousal theory of aesthetics, which follows an inverted U-shaped function that links the arousal potential with its hedonic value (Berlyne, 1971). According to that hypothesis, pleasure is maximal at intermediate degrees of complexity because too simple stimuli are boring and too complex stimuli are distressing. For relatively low degrees of dissonance, preference increases with increasing dissonance; for relatively high degrees of dissonance, preference decreases with increasing dissonance (Gordon & Gridley, 2013; Parncutt, 1989).

1.2.4.2 Is the preference for consonance innate?

A long-standing question in music cognition research is whether (or to what extent) the preference for consonance is innate. As reviewed above, there is a clear neurophysiological basis to dissonance perception, but it further depends on how psychoacoustics interacts with acquired implicit knowledge, enculturation and musical experience (Popescu et al., 2019). There are divergent perspectives about the origins of the preference for consonance, some emphasizing biological factors and others emphasizing experiential factors, but they are not mutually exclusive.

Some authors show that the preference for consonance is present even in the absence of enculturation or exposure, maybe because harmonic tone combinations are somehow similar to human vocalizations (Bowling & Purves, 2015). The perception of pitch structure may thus develop from domain-general capabilities and constraints of the auditory system (Bidelman, 2013; Tramo et al., 2001). fMRI evidence shows that the subcortical auditory system presents physiological sensitivity to dissonance already present at birth (Popescu et al., 2019). Moreover, phase-locked activity in the brainstem predicts well the relative ordering of consonance reported in behavioural studies: consonant intervals elicit more robust and synchronous phase-locking than dissonant intervals. Some authors hold that the preference for consonance may simply be a by-product of the more effective processing and reduced computational load that simple periodic information requires (Bidelman & Krishnan, 2011). In this line of thought, the sensory dissonance (some authors prefer sensory irritation) account proposed by Helmholtz holds that unpleasant sensations arise from roughness: beating is typically considered unpleasant and designated as dissonant by western listeners and is used to modulate tension in music (McDermott et al., 2010; Plantinga & Trehub, 2014; Popescu et al., 2019). However, the preference for consonance is not related to the aversion to roughness: although the presence of roughness negatively affects pleasantness ratings, even for listeners without exposure to western music (McDermott et al., 2016) and listeners with amusia (Cousineau et al., 2012), by itself is not enough to account for listeners' preferences for some intervals over others (Bowling & Purves, 2015). Meanwhile, the preference for harmonicity predicts well the preference for consonance: the stronger the preference for harmonicity, the stronger the preference for consonance (McDermott et al., 2010, 2016). However, listeners without exposure to western music (McDermott et al., 2016) and listeners with amusia (Cousineau et al., 2012) do not show any preference for consonance or harmonicity. Moreover, while the preferences for harmonicity and consonance correlates with time of exposure to music (at 6 years of age start increasing through development) and years of musical training, the preference for non-beating stimuli does not (McDermott et al., 2010; Weiss et al., 2020). Therefore, the aversion to beating may represent an aesthetic evaluation orthogonal to the preference for consonance.

In contrast, the role of experience in the emergence of consonance preference is demonstrated by developmental studies on infants, cross-cultural studies, and comparisons of listeners with different degrees of musical experience (Weiss et al., 2020). Evidence on infants' sensitivity toward consonance has been used to argue in favour of the universality of the preference for consonance. For instance, new-born's brain responses distinguish consonant from dissonant chords and they listen and look longer to consonant than to dissonant musical excerpts (Trainor et al., 2002; Virtala & Tervaniemi, 2017b; Weiss et al., 2020). However, looking times only represent interest/attention but are uninformative about aesthetic preferences (Plantinga & Trehub, 2014). This evidence on infants' research is neither robust nor consistent across ages and their results might be highly influenced by familiarity, as infants have been shown to have preference for familiar tones heard during pregnancy. When 6-month-old infants are exposed to consonant melodies accompanied by consonant chords and dissonant melodies in dissonant contexts, they always look longer to the one they are familiarized with (Plantinga & Trehub, 2014). As they grow older, western children tend to rate dissonant intervals more favourably than adults, maybe because of a relatively flexible template of consonance (Weiss et al., 2020). Proper preferences for consonant over dissonant intervals do not appear until 9 years of age and become adult-like around 12 years, which is accelerated by musical training (Plantinga & Trehub, 2014).

Cross-cultural studies show that consonant intervals with small integer ratios (such as the octave and the fifth) are structurally more important than intervals with more complex frequency relations (such as the tritone) across cultures, which lead some authors to argue for biological constraints. Non-western listeners (such as the African Mafa tribe) prefer consonance over dissonance, although the interpretation of that study (Fritz et al., 2009) is unclear because dissonant excerpts are also spectrally more complex. Later, McDermott and colleagues (2016), showed that the Tsimane' tribe are indifferent to consonance, although they still prefer non-rough stimuli (McDermott et al., 2016). Moreover, Indian listeners judge dissonant sounds to be less "in need of resolution" than western listeners (Virtala & Tervaniemi, 2017b). Thus, literature suggests that all cultures hear roughness, but what diverges across cultures is the preference for consonance and aversion to dissonance. For instance, in Middle Eastern, North Indian and Bosnian musical cultures beating is evaluated neutrally or even favourably and they often use dissonant intervals or tuning systems. Indonesian gamelan orchestras are tuned to produce beats, resulting in music considered "lively and full". Although these differences in attitudes towards beating are acknowledged in ethnomusicology, they are largely ignored in psychoacoustics and music cognition research (Plantinga & Trehub, 2014). Therefore, positive aesthetic responses to consonance may emerge from experience with consonant intervals. As an example of that, while the Tsimane' community showed no preference for consonant intervals, these preferences increased in other close communities as a function of how much they were exposed to western music (McDermott et al., 2016).

Research with musically trained listeners show that musical training strongly modulates the preferences for consonance and dissonance in complex ways. For instance, some studies show that musicians possess stronger preference for consonance and a strong negative response towards dissonance (McDermott et al., 2010). That preference might derive from an experience-dependent refinement in the internalized templates for complex harmonic sounds (Bidelman, 2013) and musicians being highly familiar with the conventional affective connotations of different musical features. Other studies show that musicians are also more familiar with dissonant chords (especially in jazz and modern genres) and tend to rate all types of chords as more consonant and preferred (Virtala & Tervaniemi, 2017b). In other words, musicians object less to dissonance than non-musicians and even prefer mildly dissonant chords over pure consonance. For a trained listener, dissonance does not imply a decrease in a chord's pleasantness (Popescu et al., 2019).

Importantly, musicians make a clear distinction between pleasantness and consonance, while for non-musicians the two concepts are barely distinguishable (Lahdelma & Eerola, 2016; Plomp & Levelt, 1965; Popescu et al., 2019; Roberts, 1983). The relationship between consonance and pleasantness is further affected by musical style (and the aesthetic ideals associated with it). For instance, jazz, classical and avant-garde musicians have different pleasantness-dissonance profiles that depend on a combination of a listener's expertise and on the music's harmonic characteristics. Each style employs different mechanisms to trade-off consonance for pleasantness. In jazz, dissonance can be used to elicit pleasantness (chords are dissonant yet pleasant) while for classical musicians, the former occurs at the expense of the latter. Although consonance is almost never associated with unpleasantness (although it can become uninteresting), the higher the musical sophistication of the listener the stronger is the decoupling between the absence of roughness (smoothness) and pleasantness. Importantly, more experienced listeners could derive pleasure from music not merely based on acoustic properties, but from dimensions such as structural cues, culturally or autobiographically relevant connotations (Lahdelma & Eerola, 2016).

The fact that the status of intervals has not stayed constant over time and varies across cultures and musical experiences, suggests that they are linked with subjective cultural dimensions. Musical experience may enhance an initial innate bias for harmonic sounds and lead listeners to learn to like harmonicity, rather than learning that specific arbitrary chords are pleasing (McDermott et al., 2010). Thus, the perception of consonance and dissonance may be grounded in psychoacoustics but mediated (and possibly overridden) by culturally acquired preferences (Popescu et al., 2019). Zatorre posits that "rather than mapping onto a simple pleasantness dimension, dissonance and consonance may be better thought of as ways to manipulate sound expressively and engender emotions" (Popescu et al., 2019; Zatorre, 2016).

1.2.4.3 Musical unexpectedness

One of the factors that makes music enjoyable is the mental satisfaction that the listener finds in continuously following and anticipating the composer's designs, sometimes by being confirmed in the expectations and sometimes being by led astray. Generation and violation of expectations induce experiences of tension, release, surprise and uncertainty (Brattico & Pearce, 2013), independently of our familiarity with the musical piece. Veridical expectations (based on explicit knowledge of a specific piece) and schematic expectations (based on years of implicit learning through exposure; Bharucha & Stoeckig, 1987) further interact to shape how we experience expectations: even if we know what is about to happen, it is difficult to switch off schematic expectations, which allow us to keep enjoying a familiar, expected, piece of music (Brattico & Pearce, 2013). According to Huron (2018), our brain derives aesthetic responses by evaluating whether musical events conform to prior expectations. Anticipatory success can derive in emotions with positive valence, leading to a preference for predictable events. In contrast, surprising events may elicit affective reactions with negative valence and stress, derived from maladaptive anticipatory failure, which provides negative feedback for the learning process that generated the prediction (Huron, 2018; Meyer, 1956). For instance, musicians and non-musicians' have been shown to provide negative affective ratings chords for violations of harmonic expectations (Loui & Wessel, 2007).

Nevertheless, surprise can be enjoyable. Chills can be elicited by violation of expectations, such as unexpected harmonies, sudden dynamic or textural changes (Brattico & Pearce, 2013), which is reflected in increased emotionality and physiological arousal (EDA) as a response to unexpected events (Steinbeis et al., 2005). Cognitive dissonance (not to be confounded with musical/sensory dissonance) might mediate pleasure by encouraging listeners update their generative model of the environment to match the output, so the brain minimizes prediction errors and avoid negative affective responses resulting from cognitive conflict (Brattico, 2015; Cheung et al., 2019). In fact, tension ratings and physiological arousal increase as a function of harmonic unexpectedness (Bharucha, 1984). Cheung and colleagues proposed that the expectation of a chord (based solely on its conditional probability of occurrence and independently from its acoustic characteristics) can evoke pleasure through two temporally dissociable states: the uncertainty when anticipating what chord could be before it occurs (or how precise are our predictions) and the surprise elicited when the actual chord deviates from expectations (which depends on the probability of each chord given the tonal harmonic context). When uncertainty of the harmonic context is low (toward the end of a cadence), chords with higher surprise are rated as more pleasant than less surprising chords. When uncertainty is high (in atypical chord progressions), less-surprising chords are rated as more pleasant (Cheung et al., 2019). These results are reminiscent of the inverted U-relationship (but more multifaceted) between expectation and affective arousal curve (Berlyne, 1971; Meyer, 1956), according to which music that slightly deviates from expectation is experienced as most emotionally arousing (and, therefore, most pleasurable), whereas highly expected or highly unexpected music is less interesting or arousing. Indeed, unstable tones serve important musical functions such as embellishments and ornaments and, more importantly, they have a dynamic quality because they can induce in the listener a need for resolution to a consonant event (Bharucha, 1984).

In addition to culture, genre and style, the musical expertise of the listener shapes their internal model of the statistical regularities of chords in a progression and affects how surprising a chord is and with how much precision it can be expected (Cheung et al., 2019). For instance, tension judgements are more pronounced in musicians, who are sensitive to roughness and harmonic effects. Meanwhile, non-musicians base their responses on the most easily perceivable features, such as melodic surface (Bigand et al., 1996). When making aesthetic judgements, musicians rely more on cognitive strategies, although they are known for appreciating mixed emotions in music (Virtala & Tervaniemi, 2017b), whereas naïve listeners show enhanced emotion-related processing (Brattico & Pearce, 2013). The training of musicians in understanding formal structures might surpass the training in recognizing and conveying emotions via music (Brattico et al., 2016). In line with that idea, some authors argue that the valence of an aesthetic response is determined by the ease and speed with which a stimulus can be processed: the more fluent the processing, the more pleasant the experience (Reber et al., 2004). According to that perceptual fluency hypothesis, because musicians possess increased auditory processing abilities, they might show a preference for more complex musical styles (Brattico et al., 2009; Brattico & Pearce, 2013). For instance, when comparing highly expected endings of chord sequences, medium expected chords at the middle of the cadence and unexpected endings, non-musicians prefer highly expected endings over medium expected and unexpected chords. Classical musicians have an even stronger preference for expected and medium expected over unexpected endings. Meanwhile, jazz musicians show higher ratings for the medium expectation condition, suggesting an inverse relationship between jazz training and preference for the expected. Indeed, jazz musicians are more exposed to novel, unexpected harmonies and create novel auditory-motor sequences in real time that are aesthetically and emotionally rewarding. This fact may eventually discourage sounds that are too expected or ordinary and encourage a higher tolerance, or relative preference, for more unexpected or complex stimuli. These results are also in accordance with Berlyne's theory inverted U (Przysinda et al., 2017).

1.3 From sound to music in the brain

1.3.1 The processing of sound in the auditory system

The primary acoustic circuit starts at the outer ear, where sound vibrations are collected and transduced into neural signal in the cochlea by the action of the auditory receptor cells of the basilar membrane. In the cochlea, sound is broken down according to its frequency content. Thus, physical periodicity is transformed to neural periodicity, which provides a frequency-to-place, or "tonotopic" mapping of sound into neural signals. The signal is carried by the auditory nerve (AN) up a series of interconnected nuclei toward the auditory cortex, where musical percepts are generated and controlled (Bidelman, 2013; Peterson et al., 2021). That tonotopic mapping of frequencies is maintained throughout the auditory system, although the precision deteriorates at each successive stage (McDermott & Oxenham, 2008). At the primary auditory cortex, the basic acoustic features of music (such as pitch, timbre, intensity and roughness; (Boso et al., 2006; Koelsch, 2011; Peretz & Zatorre, 2005) are extracted. At the right secondary cortex, early stages of melodic analysis take place and pitch relationships such as contour (defined by pitch directions of sequentially presented tones) and intervals (defined by frequency ratios between simultaneously presented tones) are integrated. Furthermore, posterior areas of the secondary cortex process pitch height and anterior regions process pitch chroma (Peretz & Zatorre, 2005).

1.3.2 Neural processing of dissonance and musical expectations

Neural representations of music are emergent before cortical involvement at pre-attentive stages of audition. In our auditory system, sound is transformed into musical percepts. To do so, incoming acoustic input is first separated into sound sources, then the extraction of sound features allows the representation of sounds as auditory objects with specific characteristics.

As we ascend in the auditory pathway, the central auditory system distinguishes pitch relationships according to their consonance by exploiting the harmonicity of sound. Consonant intervals trigger neuronal firing at precise, harmonically related pitch periods. In fact, the activity at the AN correlates well with perceptual judgements of consonance whereas dissonant relations produce multiple, more irregular neural periodicities. If dissonant intervals are separated by less than a critical bandwidth (and generate roughness), their amplitude envelopes overlap in the basilar membrane and are difficult to encode into the AN (Plomp & Levelt, 1965). Therefore, consonant intervals may be more compatible with pitch templates and provide a more robust, unambiguous cue for pitch. That reduces computational load and require fewer brain resources, therefore processing consonance is computationally more efficient than processing dissonance (Bidelman, 2013). Specifically, the more robust and synchronous phase-locking of consonance manifests early in subcortical areas (such as the AN and midbrain) and in the cortex. Moreover, the frequency following responses (FFR) of consonant intervals arising from the brainstem yield more harmonicity than the FFR of dissonant intervals (Bidelman, 2013). Consonant chords also elicit differential hemodynamic responses in inferior and middle frontal gyri compared to dissonant chords. Meanwhile, dissonance is computed bilaterally in the superior temporal gyri: neural populations of Heschl's gyrus exhibit phase-locking to roughness for dissonant (but not for consonant) chords (Minati et al., 2009a).

Moving on to the generation of musical predictions: it involves interactions between sensory and low-level predictions (data-driven perception), acquired style-specific syntactic or schematic knowledge (harmonic rules) and veridical knowledge of the present piece as well as non-sensory structures acquired during a piece through processes of online-learning (Rohrmeier & Koelsch, 2012). These processes take place at different levels of brain processing, involving associative auditory sensory memory and more abstract representations of musical syntax (Kalda & Minati, 2012). These processes follow specific steps:

once sound sources are separated and auditory objects are created, regularities inherent to the sequential presentation of each event are detected and integrated into a model of the acoustic environment. Then, predictions about forthcoming auditory events are derived from the model. Finally, representations of the incoming sound and the sound predicted by the model are compared (Koelsch, 2009a). It is important to keep in mind that, although the creation of melodic or harmonic expectations follows similar paths, chord progressions imply processing at various levels at the same time because they include multiple pitch relationships. Thus, the analysis of sequential musical features (such as pitch changes) are processed in the primary auditory cortex, which activates temporal regions often in the right hemisphere (McDermott & Oxenham, 2008). But the analysis of harmonic relationships engages a whole network: acoustic features are extracted in the anterior superior temporal gyrus (STG) while abstract relationships are processed in the inferior frontal gyrus (IFG; which establishes the hierarchical relationships between sounds) and the lateral premotor cortex (which is key for the short-term prediction of upcoming events), and from there, passed down again to the auditory cortex. Then, detecting music-syntactic irregularities involves the pars opercularis in the IFLC and the vlPMC. It has been even suggested that there might be an immediate link between the prediction of upcoming events and a representation of corresponding motor schemas in the PMC to enable an immediate mapping of perception onto action. Moreover, working memory – which holds each sound in memory to compare it to the next event (Gaab & Schlaug, 2003) - engages the auditory cortex as well as frontal areas. More specifically, the detection of deviations from harmonic expectancies engages inferior fronto-lateral areas (the frontal operculum) bilaterally, which correspond to Broca's area on the left hemisphere (Koelsch, Gunter, et al., 2002; Maess et al., 2001; Peretz & Zatorre, 2005; Tillmann et al., 2003).

1.3.3 Electrophysiological signatures of music processing

The processes of auditory analysis reviewed above generally result in stereotyped electrophysiological responses. Via electroencephalography (EEG) we are able to visualize these responses and study the dynamics of information processing. One of the most common ways to visualize electrophysiological changes is to represent voltage changes (measured in μV) time-locked to the latency of appearance of stimuli (measured in msec), called event-related potentials (ERPs) (Boso et al., 2006). Because of their excellent temporal resolution, ERPs allow to study how perceptual and cognitive processes unfold in time. Particularly, by determining at which point in time the ERPs elicited in two experimental conditions start to diverge, we can infer the temporal correlates of the two underlying processes (Schön et al., 2005). Determining the latency at which two processes differ is really useful to infer at which stage of processing the brain has started reacting to the stimuli. For instance, early ERPs (around latencies of 100-200ms after the presentation of the stimuli) usually reflect perceptual processes (e.g. N1, P2, N2, which can be grouped as auditory evoked potentials or AEP) while later ERPs (from 300-400ms and beyond) reflect more complex cognitive processes (P3, N5, LPC). The characteristics and implications of the relevant ERPs will be briefly reviewed next.

1.3.3.1 ERPs to consonance-dissonance

The ERPS of consonant and dissonant sounds have been investigated with different paradigms, most of them with chords presented in isolation. For instance, Proverbio et al. (2016) examined chords with different degrees of dissonance. They found that consonant chords elicit an early auditory N1 in comparison to dissonant chords (see also Regnault et al., 2001). The N1 is a well-studied response that represents the first cortical response to sounds and is associated with the processing of physical attributes of stimulus (Näätänen & Picton, 1987; Regnault et al., 2001). The authors suggested that, since consonant chords elicit a stronger N1 than dissonant chords, the universality of sensory consonance may be an emergent property of the nervous system (Proverbio et al., 2016). Consonant chords also elicit a stronger P1 (Minati et al., 2009b) and P2 than dissonant chords (Itoh et al., 2003; Kung et al., 2014). P1 indexes early auditory processing and its amplitude modulations result from the phase-locking of oscillatory activity of the primary auditory cortex to the degree of consonance. P2 is part of the auditory N1/P2 complex and is thought to reflect a rough evaluation and classification of stimuli (known as "first rough stimulus appraisal"; García-Larrea et al., 1992). P2 is also linked to associative processes such as holistic or gestalt grouping of pitches, which is relevant in the perception of consonance. Thus, the P2 reflects the determination and representation of the consonance of intervals based on whether they are easily grouped together to represent a single entity (Itoh et al., 2003).

Multiple studies found that the degree of dissonance of intervals presented in isolation is reflected in an increase in the amplitude of an auditory N2 (Kung et al., 2014; Minati et al., 2009 but see Schön et al., 2005). The N2 has been suggested to reflect sensory discrimination and classification processes. The larger N2 observed for dissonant chords suggests their categorization according to acoustic characteristics. Importantly, most of the studies mentioned above used harmonic complex tones (by using sounds from musical instruments like the piano), which can elicit roughness. Other studies have used sinusoidal tones with intervals wider than the critical bandwidth to rule out roughness from contributing to dissonance. These studies found that the N2 reflects the degree of dissonance even in the absence of roughness, therefore it might reflect a categorization based on factors other than sensory roughness (Itoh et al., 2003, 2010).

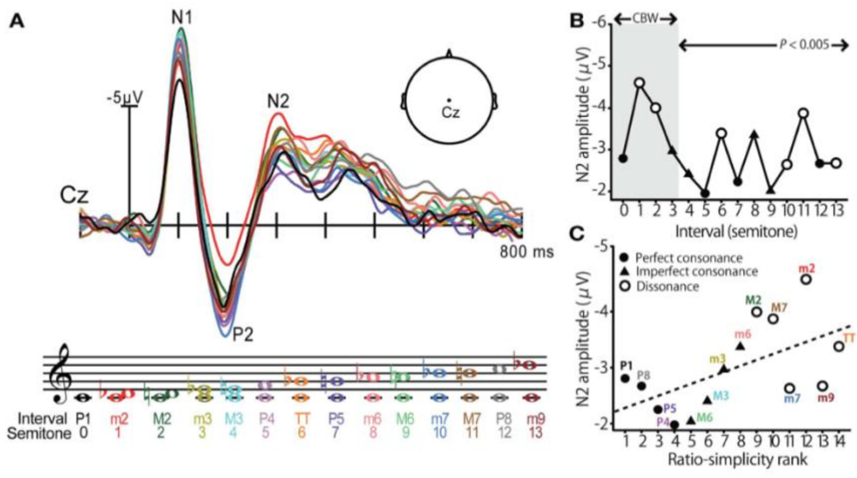


Figure 7. (A) Cortical event-related potentials elicited by chromatic musical intervals. (B) Cortical N2 response magnitude modulated by the degree of consonance. The shaded region demarcates the critical bandwidth (CBW). (C) Response magnitude of the N2 as a function of the ratio simplicity (adapted from Bidelman et al., 2013).

When embedded in oddball paradigms, dissonant chords elicit a larger MMN than consonant chords or other acoustic irregularities, such as sound intensity or mistuning (Brattico et al., 2009; Crespo-Bojorque et al., 2018; Virtala et al., 2011a). This result also argues in favour of the idea that listeners are able to make a qualitative distinction between consonance and dissonance at early auditory stages. Moreover, in musicians, consonant chords elicit a larger P3 than dissonant chords. That musicians show a larger P3 for consonant chords reflects their ability on detecting tonal and atonal relationships (Proverbio et al., 2016).

Some studies investigated the neural responses to dissonant chords within a tonal context by embedding them in chord cadences. Strongly dissonant chords placed as cadence closure elicit a very large N5 (Koelsch et al., 2000), suggesting that they require a large effort to integrate them into the previous musical context. That response parallels to the N400 of semantic irregularities in language, elicited by non-words (which are comparable to dissonant chords within a western musical context). Finally, different chord functions rendered dissonant, elicit a late positive component (LPC), independently of their harmonic expectedness (Regnault et al., 2001). In that study, the authors suggested that such a late response to dissonance may derive from the decision-making processes on whether the chords are dissonant or not, and to what extent they fit the context.

1.3.3.2 ERPs to sound violations and broken harmonic expectations

When listeners are presented with streams of sounds, they build representations of the regularities of local inter-sound relationships extracted on-line from the acoustic environment (Koelsch, 2009a). The auditory sensory memory forms representations of the repetitive aspects of auditory stimulation or memory traces (Yu et al., 2015). If an oddball (an infrequent deviation) is introduced in a stream of repeated or familiar events (frequent standards) by changing any auditory dimension, the brain automatically detects it pre-attentively (Garrido et al., 2009). That cognitive operation elicits a well-studied auditory ERP: the mismatch negativity (MMN), which is represented as a difference wave obtained by subtracting the ERP of the oddball events from the standards (Näätänen et al., 2007). The MMN is elicited by stimuli ranging from simple deviations such as timbre, pitch or beat (which are physical deviations and elicit a *phMMN*) to complex violations in abstract features (which elicits the *afMMN*, Saarinen et al., 1992), abstract rules (such as higher-order grammar violations) or speech sounds (Näätänen et al., 2007). However, music involves changes across multiple dimensions at the same time. Studies using oddball paradigms investigated the specific error-related response to diverse music features such as changes in intensity, frequency, duration, stimulus omission, timbre or pitch contour or information specific to melodic structures, such as changes in contour and interval direction (Näätänen et al., 2004; Vuust et al., 2011). More specifically, musical scale incongruities such as out-of-tune and out-of-key chords elicit a MMN that is larger for mistuned tones (Brattico et al., 2006), although they likely reflect their perceptual salience, rather than the violation of abstract scale rules.

Temporal structure is a key aspect of music because it is constituted by complex architectures based on hierarchical relationships between scale tones. In a structured tonal context, musical events (and the subsequent expectations) are organized across time by their relative probabilities and stability within the key. Melodies convey enough syntactic information within one voice to create expectations for tones to follow. For instance, scale violations in melody-like structured sequences elicit an early negativity strongly reminiscent of the MMN termed the *early right anterior negativity* (or ERAN) (Kalda & Minati, 2012; Miranda & Ullman, 2007). The ERAN has been proposed to reflect the processing of acoustic information structured according to complex and abstract regularities (Yu et al., 2015). Moreover, in chord sequences, each chord is built by several voices that together convey syntactic information about its function. Thus, the subsequent hierarchy of stability is more complex than the hierarchy of stability of tones (Bharucha & Krumhansl, 1983; Krumhansl, 1979), but elicit similar ERPs. For instance, out-of-key chords placed at the end of a chord cadence elicit an ERAN (Koelsch et al., 2000) because they are harmonically distantly related to the preceding harmonic context and therefore introduce a tonal-syntactic irregularity. The amplitude of the ERAN, but not of the MMN to mistuned chords (Leino et al., 2007), increases with tonality establishment across positions in a chord cadence (Koelsch et al., 2000; Leino et al., 2007) and when unexpected chords appear instead of highly expected chords. Therefore, the ERAN has been suggested to reflect the processing music-specific syntactic rules (such as tonality establishment) but, more importantly, the rules of chord succession, which determine the order of chords within a cadence (Brattico et al., 2006).

There is a current debate over the roles that auditory sensory memory and cognitive music-syntactic processes have in detecting musical irregularities (Kalda & Minati, 2012; Yu et al., 2015). In many studies investigating the ERPs related to musical violations, music-syntactic regularities co-occur with acoustic similarity. For instance, Neapolitan chords used in the study by Koelsch and colleagues (2000) introduced pitches that had not been presented in the previous harmonic context and, thus, the ERAN may overlap with a phMMN (Koelsch, 2009a). There is, in fact, a strong resemblance between the ERAN and MMN. They have similar time-courses and scalp distribution and both their amplitudes increase with the amount of violation. Both correlate with behavioural performance and can be elicited pre-attentively (Koelsch et al., 2001). In many studies, the ERAN is even referred to as the music-syntactic MMN (Koelsch, Grossmann, et al., 2003; Koelsch, Gunter, et al., 2002, 2003; Koelsch, Maess, et al., 2003), not only due to its resemblance with the MMN but because the term early *right* anterior negativity falls short when the effect elicited by irregular chords is not significantly lateralized. Thus, the use of the term ERAN is more related to its proposed functional significance than its scalp distribution.

However, most literature insists in the importance of not confounding them. For instance, the ERAN is elicited by out-of-key chords (Koelsch & Friederici, 2003; Koelsch & Sammler, 2008), which could lead to argue that it reflects the detection of pitch deviance. However, it is also elicited when controlling for sensory-novelty (for instance, by introducing the deviant notes previously in the cadence) and by syntactically incorrect chords (such as the supertonic as a cadence closure), that belong to the tonality and do not introduce any kind of physical deviance or, even, are acoustically more similar to the context

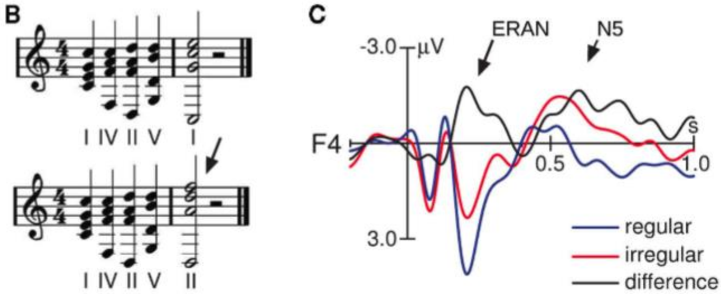


Figure 8. (Left) The dominant-tonic progression represents a regular ending of a harmonic sequence (top), the dominant-supertonic progression is less regular and unacceptable an end of a harmonic progression (bottom). (Right) Brain responses to irregular chords (best to be seen in the black difference wave, regular subtracted from irregular chords), ERAN is maximal around 200ms and is followed by an N5 (adapted from Koelsch, 2011).

(Koelsch et al., 2007; Koelsch & Sammler, 2008; Koelsch & Siebel, 2005; Leman, 2000; Poulin-Charronnat et al., 2006). Therefore, the ERAN has been suggested to be sensitive not only to violations of the tonal key but also to violations of music-specific syntactic regularities within a tonal key and to work independently of the activation of auditory sensory memory (Kalda & Minati, 2012).

The processes underlying the generation of the MMN and the ERAN are pretty similar, but the comparison of the auditory input to the internal models differ. In the case of the MMN, the internal model of sound regularities is extracted on-line. Meanwhile, for the ERAN, incoming sounds are integrated into a cognitive model of the representations of regularities stored in long-term memory and acquired via passive exposure to western music (Bigand, 2003; Koelsch, 2009a). The generators of the MMN are located in temporal areas (specifically, the superior temporal gyrus or STG) involved in auditory processing, with additional contributions from frontal areas (Näätänen et al., 2007).

Meanwhile, the ERAN has its neural generators in inferior frontal regions, which are involved in the processing of syntactic rules of language (Maess et al., 2001). Together the MMN, ERAN and ELAN (the equivalent of the ERAN for linguistic syntactic irregularities) belong to a family of peri-sylvian negativities that reflect the processing of irregularities of auditory input. One view proposes that they reflect stages on a continuum from rather simple (physical) to complex (syntactic) auditory feature processing (Koelsch et al., 2001). Importantly, both these mechanisms are bidirectionally related. Tonal hierarchies and music-syntactic regularities stored in memory are partially grounded on acoustic similarities because auditory sensory memory provides the basis for learning more complex rules. For instance, chords related to a previous context also have more component tones in common than a chord less related to the context (Bigand, 2006). Processes such as the formation of representations of the standards, the detection and separation of auditory objects and their subsequent sequential organization in memory allow the extraction and memorization of statistical probabilities (Koelsch, 2009a). At the same time, the learned hierarchically structured representations support more abstract monitoring of ongoing note streams by activating specific context-dependent expectations (Kalda & Minati, 2012; Koelsch, 2009a).

Violating harmonic expectancies also elicit other ERPs of interest. For instance, ending a cadence in an unexpected chord (such as the subdominant, which belongs to the tonality and does not introduce a frequency deviation) elicits a larger P3 than ending in an expected chord (Janata, 1995; Regnault et al., 2001). The P3 reflects top-down influences in perceptual stages of processing and is usually related to the unintentional switch of attention to the novel or unexpected event. The P3 has two subcomponents: a frontal P3a, related to attention orientation to no-go stimuli and a parietal P3b, normally elicited by voluntary attention toward target stimuli. Moreover, out-of-key deviant notes in melodies and irregular chords (such as Neapolitan endings of chord cadences) elicit an N5. The N5 is a bilateral late negativity suggested to reflect processes of melodic and harmonic integration (Miranda & Ullman, 2007). It has been suggested that the amplitude of the N5 indexes the amount of effort invested in integrating the irregular musical event into the preceding context (Koelsch et al., 2000; Koelsch, Gunter, et al., 2003). Finally, unexpected endings of chord cadences have also been reported to elicit an LPC or P600 (A. D. Patel et al., 1998), which is also found in syntactic irregularities in language. The P600 has been related to decisional processes regarding syntactic relationships by indexing the difficulty of fitting a given chord into the established context.

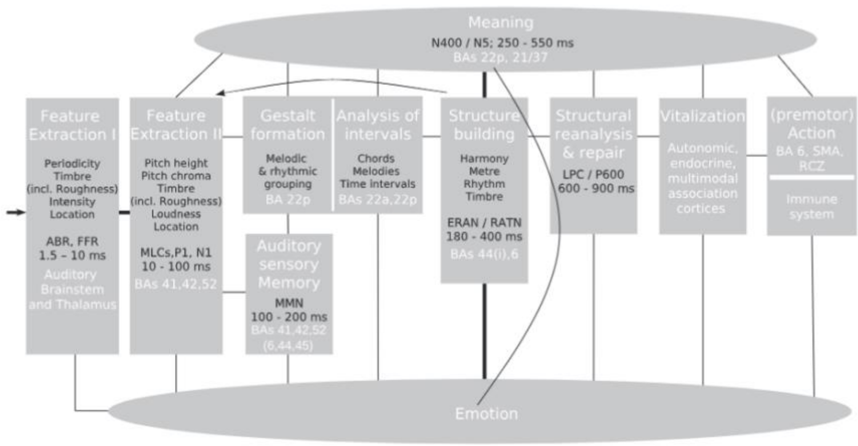


Figure 9. Neurocognitive model of music perception proposed by Koelsch, 2011.

1.3.4 Brain plasticity as a result to long-term exposure to music

The acquisition of implicit and cultural-specific knowledge about the regularities of tonal music results from the long-term exposure to music in everyday life (Tillmann et al., 2000). These developmental processes are reflected in the neural responses of western infants and children. The ERP responses of young infants are dominated by slow positive waves; but by 3 to 4 months of age, faster negative components are apparent in response to unexpected sound features (Hannon & Trainor, 2007). These components increase in amplitude with age, reaching a maximum around 10 to 12 years, and diminish to adult levels by 18 years of age. More specifically, the ERAN and the N5 as a response to Neapolitan endings are, in fact, observed already in 30-month-old children (Jentschke, 2007; Jentschke & Koelsch, 2009). By 5 years of age, the ERAN is present as a response to deviations at the end, but not in the middle of chord cadences, likely because, at this age, infants still might have less specific representations of musical regularities than adults (Koelsch, Grossmann, et al., 2003). Interestingly, musical training accelerates this development, as seen in 4- to 6-year-old children studying music, who show larger N1 and P2 responses than children not undergoing musical training (Shahin et al., 2003). Moreover, 10-year-old children show a larger ERAN than untrained children. Meanwhile, no difference is found at that age for the N5, as well as between adult musicians and non-musicians (Jentschke, 2007; Koelsch et al., 2005).

A very relevant case of brain plasticity caused by exposure to music is musical training. In a way, musical training simply amplifies a human musical capacity rooted in innate predispositions (Trehub, 2003). Via an extensive learning of perceptual, cognitive and motor skills, musical training facilitates the encoding and memory for musical structures. That is further reinforced by the use of labels of musical events, eventually leading to the acquisition of an explicit musical knowledge (Bigand, 2006). As a result, musicians display a variety of changes in their brain structures and functioning. At a general level, they show structural and functional enhancement of the motor cortex, as well as changes in motor-related areas such as the corpus callosum and the cerebellum (Peretz & Zatorre, 2005). They also show an increase in grey matter in the auditory cortex (Schneider et al., 2002). Moreover, the IFLC (pars opercularis) and the right STG are activated more strongly in musicians than in non-musicians, in both adults and children (Koelsch et al., 2005). Importantly, there is a complex interplay between structural changes and neural responses. For instance, string players show enlarged cortical representation of the left hand (Elbert et al., 1995) although professional pianists show decreased activity in motor areas (Hund-Georgiadis & Yves Von Cramon, 1999). A greater volume of tissue may reflect a reorganization at the structural level, which may manifest as recruitment of fewer neurons, different synchronization of firing patterns or even different connectivity. That suggests that musicians may recruit more neural tissue or use it more efficiently. However, the nature of that reorganization is still under investigation.

Musicians also possess a general enhancement of the auditory processing of sound stimuli in comparison to non-musicians. Musicians display increased neural responses to piano tones, which are more pronounced for their own type of instrument, but also for pure tones, which do not exist in the traditional musical environment (Peretz & Zatorre, 2005). More specifically, musicians possess an enhanced and temporally more precise encoding of pitch-relevant information. Therefore, they have better performance in detecting small pitch differences and structuring rhythms than non-musicians. Moreover, their more robust hierarchically-structured representations activates specific context-dependent expectations, which supports more abstract monitoring of ongoing note streams (Kalda & Minati, 2012). Consequently, pitch information can be decoded with higher resolution by musicians, leading to a stronger phMMN than non-musicians (Koelsch, 2009a). In fact, chords deviating by very small differences in pitch are enough to elicit MMN in musicians, whereas non-musicians require a much larger pitch deviation. In addition, rhythmic deviations also elicit a stronger and earlier MMN (Yu et al., 2015). Also, an N2 is elicited only in musicians in response to chords with quartertones, supporting a greater pitch sensitivity (Proverbio et al., 2016). Together, these results suggest that musicians are more sensitive to acoustic stimuli in general (which could be explained by better auditory abilities) but they also are more sensitive to music-specific deviations. For instance, musicians (but not non-musicians) display a MMNm for contour and interval changes (Näätänen et al., 2007; Tervaniemi et al., 2001), while both show a MMNm for frequency changes in single tone patterns (Pantev et al., 2003). Reversed order of consecutive tones elicits a larger MMN in musicians than in non-musicians (Brattico et al., 2002; Tervaniemi et al., 1997), suggesting that they have more accurate neural representations for temporal information than non-musicians. These neural differences are moreover modulated by the length of musical training and their specific practice strategies (Musacchia et al., 2007), especially for those who began to play their instruments earlier (Elbert et al., 1995). For instance, the MMN is larger in musicians who need to intone while playing (such as violinists and singers) and also as a response to contour deviations in musicians who perform without a score in comparison to musicians that follow a score (Yu et al., 2015).

The processing of consonance and dissonance is also modulated by musicianship, both at the level of the brain areas that are engaged and the ERPs that are elicited. For instance, the involvement of inferior and middle frontal gyri for consonant chords is more distributed in musicians (Bidelman, 2013; Minati et al., 2009b). Moreover, the processing of quartertones elicits an asymmetry in musicians, who show the engagement of the left temporal cortex, while the right side is engaged in non-musicians. Regarding the ERPs, musicians show a larger N1 for consonant chords, likely because they are strongly reactive to physical attributes of musical stimuli, allowing them to react to the distinction between consonance and dissonance at very early stages of processing (Regnault et al., 2001). However, more recently, Bidelman and colleagues (2013) also found that the P1/N1 complex is similar in musicians and non-musicians, while there's a distinct variation of later waves (N2) nearly exclusively in musicians. The fact that N1/P1 is similar regardless of musicianship, but N2 is more modulated in musicians suggests that musical training might exert more changes on later, endogenous mechanisms (N2) than on earlier, exogenous processing (P1/N1). Dissonant intervals in streams of unstructured consonant chords elicit a clear MMN, but consonant chords in streams of dissonant chords only elicit a MMN in musicians (Crespo-Bojorque et al., 2018). Moreover, in musicians, N2 follows the degree of consonance accepted in western music tradition and not only the degree of absence of roughness (Itoh et al., 2010). That result is supported by the fact that, for distinguishing consonant from dissonant intervals, musicians rely on pitch-intervals whereas non-musicians rely on roughness (Kung et al., 2014).

Regarding the processing of harmonic expectancies, very few studies have directly compared musicians and non-musicians. Although they are equipped with the same set tools for sound analysis than non-musicians, musicians develop stronger implicit representations of syntactic regularities and explicit knowledge of tonal regularities, which allow them to build harmonic expectations automatically (Koelsch, 2009b). That enhancement is reflected in their neural responses to musical irregularities. For instance, musicians show an ERAN as a response to scale deviants both in structured and scrambled melodies. Meanwhile, the ERAN in non-musicians is only present in structured melodies (Kalda & Minati, 2012). Moreover, the ERAN is larger in musicians than in non-musicians as a response to Neapolitan (Koelsch, Schmidt, et al., 2002) and dissonant endings of chord cadences (Pagès-Portabella et al., 2021; Pagès-Portabella & Toro, 2019). The P3 is also increased in musicians as a response to harmonic violations, to the point that it is often observed only in musicians, while the ERAN is present regardless of musical training. The larger P3 in musicians has been related to better perceptual learning, stronger and faster involuntary attention switching (Guo & Koelsch, 2016; Seppänen et al., 2012), enhanced memory matching in musical context-updating processes and structural analysis in rule-governed sequences (Polich, 2007; Steinbeis et al., 2006). These results suggest that the neuro-cognitive mechanisms of attention allocation and confirmation of expectations reflected in the P3 are influenced more strongly by musical expertise than the processes of music-syntactic analysis reflected in the ERAN (Guo & Koelsch, 2016). Finally, harmonic irregularities at the end of cadences elicit a larger LPC in musicians than in non-musicians (Besson & Faïta, 1995). The LPC is related to the conscious categorization of chords as irregular and the decision-making process (Regnault et al., 2001).

As suggested by some literature, it is important to understand whether the brain differences observed between musicians and non-musicians result from training or correspond to genetic differences that predispose some individuals to become musicians (Bigand, 2006). For instance, some studies (Bigand, 2003, 2006) challenge the anatomical and neurophysiological evidence derived from musical training (reported above) and suggests that it would be wrong to conclude that these differences have any repercussion for the general cognitive structure that allows musical processing in all its complexity (Bigand, 2003). According to Bigand (2006), these changes in the brain are relatively small compared to the amount of training received. Moreover, Bigand argues that the perception of music is an infinitely rich experience that is difficult to reduce to a sequence of rudimentary qualities of musical sound such as pitch and timbre. Therefore, we must be aware that experimental tasks of this nature tell us more about the auditory abilities of listeners than their strictly musical abilities (Bigand, 2003, 2006).

2. EXPERIMENTAL SECTION 1: BEHAVIORAL STUDY

The brain of western listeners is hard-wired and strongly sensitive to musical unexpectedness and dissonance. It is unclear to what extent the processing of unexpected musical events may be plastic, influenced by other factors such as the type of unexpectedness and musical training (Peretz & Zatorre, 2005), which will be the main focus of the present thesis. We addressed this issue in a series of studies, taking advantage of behavioural and electrophysiological methodologies. While behavioural assessments allow us to study conscious evaluations of musical preferences, electroencephalographic recordings reveal the signature changes on electrical activity in the brain, allowing us to better understand the processes underlying perceptual and cognitive responses. Thus, in a behavioural study we investigated whether western listeners share universal (low-level) preferences toward unexpected violations of different musical dimensions (syntax, tonality, consonance) or whether these may be susceptible to change as a result to long-term musical training. Moreover, because preference is a multifaceted concept that involve cognitive and affective processes, we also investigated whether both types of listeners similarly prefer unexpected events in terms of correctness or pleasurableness or whether they dissociate (and whether there is an interaction with musicianship).

2.1 Introduction

In this behavioural experiment we analysed the evaluations of musicians and non-musicians of irregular endings of chord cadences. A conscious evaluation of endings of chord sequences can be made based on different parameters. On one hand, listeners can analyse the goodness of fit/correctness/congruity of a target chord with the previous musical context, a decision that is made based on their implicit (or explicit, in the case of musicians) knowledge of musical rules. These cognitive judgments tend to be consistent across western music tradition, although they can vary according to musical practice (for instance, the functional interpretation and tolerance to the dissonance of some chords vary between classical and jazz musical tradition). We will refer to this type of evaluations as "correctness" judgements. On the other hand, listeners can also evaluate to what extent they like the endings of chord sequences (regardless of their correctness), which is based on the evaluation of their affective reactions to these target chords. Affective judgements are a more subjective measure which may be bound to higher inter-subject variability due to personal musical taste and experience. We will refer to this type of evaluation as "pleasurableness" judgements.

Our aim was twofold: first, we ought to test to what extent correctness and pleasurableness judgements of a spectrum of chords (ranging from very consonant to very dissonant chords) are modulated by musicianship. In musicians, we could expect to have a dissociation between their cognitive (correctness) and affective (pleasurableness) ratings (especially for chords with ambiguous interpretations depending on musical practice). For instance, chords that are theoretically not the optimal resolution may not be necessarily disliked. In fact, musicians are known to appreciate unexpectedness and mild dissonance (Virtala & Tervaniemi, 2017). Meanwhile, in non-musicians, given that they do not have an explicit knowledge of musical rules, we expect them to guide their ratings based on their implicit knowledge of music and emotion-based strategies (Brattico & Pearce, 2013). In other words, as non-musicians have no explicit way of telling apart what is musically correct or incorrect in theory, they have to trust their 'gut instinct'. Therefore, we would expect them to provide similar judgements in both scales of pleasurableness and correctness. Our hypotheses were: (1) there might be a dissociation between pleasurableness and correctness mediated by musicianship and (2) musicians and non-musicians may differently rate chords depending on the musical context.

Our second aim was more specific. We aimed to assess whether the neural responses registered in the electrophysiological studies (that will be reported in the next sections) were qualitatively good predictors of the pleasurableness and correctness ratings of the present experiment. In those experiments we did not ask participants to consciously evaluate the pleasurableness or correctness of chords because they were not informed about the presence of musical irregularities (see Experimental section 2 for further details). Asking them would have not allowed us to observe the neural responses of interest without the overlapping of attentional and task-relevance effects. Thus, in the present experiment we aimed to evaluate how musicians and non-musicians would rate Neapolitan sixths and dissonant chords at different positions of a chord cadence. We expected listeners to similarly rate tonic and dissonant chords (which are at opposite extremes of congruity and dissonance) but differently rate Neapolitan chords, as musicians might be more familiar with their use in musical pieces. Moreover, we were interested in testing if the position that these chords occupied within the cadence was reflected in their ratings.

2.2 Methods

2.2.1 Participants

39 volunteers participated in the study. 19 of them were non-musicians (10 identified as cisgender females, mean age 23.7 ± 4.47) and 20 of them were musicians (8 identified as cisgender females, mean age 24.2 ± 4.31). Musicians either finished or were studying Advanced Studies in Western music in Spain. They started their formal musical training when they were 6.47 ± 2.67 on average and had been musically active for 16.05 ± 3.64 on average. All musicians were musically active at the time of the experiment. 15 out of the 20 musicians have had training in classical and contemporary music. The rest specialized in jazz and modern music, pedagogy, composition, and other fields of study.

2.2.2 Stimuli

We chose a range of chords that could have an ambiguous interpretation or, at least, more than one interpretation depending on musical style or practice. These chords may not very common as cadence closure and, in terms of congruity may not be the optimal resolution (in comparison to ending with the tonic triad, as recommended by traditional music theory), but they can still be liked by listeners as they create a suspended sensation and introduce a pleasurable degree of novelty (Brattico et al., 2010). Thus, the strict correlation between musical-theoretical congruity and pleasurableness was meant to be avoided. In order to minimize effects of familiarity, the musical stimuli were purposely composed for the experiment. Therefore, they were completely novel to the subjects. Stimuli were chosen to consist of isochronous (i.e., with the same rhythm) sequences or cadences of chords following the rules of Western tonal music but otherwise without a direct association with a specific musical genre (like classical or popular music) to avoid retrieval of personal and social attitudes or association (Brattico et al., 2010).

We created unique sequences with the context chords of C-G-Dm-G (which perform the functions of T-D-sT-D), which create an expectation of resolution. Starting from this reference sequence we created 42 different versions of it. In some versions we introduced a dissonant or a Neapolitan chord at the third position. Other versions could either end on the tonic, on the Neapolitan or on the dissonant cluster. These versions were presented 12 times each. The other versions included a spectrum of major and minor chords, triads and quatriads, with and without musical tensions (that belong to the superstructure of chords) and covered a variety of degrees of dissonance. More specifically, sequences could end on Dm E7, D7, CMaj7, Cm7, Cmaj9, Cm9, CDim, Cm7b5 and Caug. All the sequences were composed in C major key for simplicity. Chords 1 to 4 lasted 800ms and the last chord lasted 1600ms. These sequences lasted longer (4800ms) than the ones used in our EEG experiments (3600ms) so participants could have enough time to consciously evaluate the chords online, especially for those versions that required them to evaluate a chord in the middle on the cadence.

2.2.3 Task and experimental procedure

The task was performed embedded in an experimental session where participants were doing an EEG experiment with an unrelated task that is out of the scope of the present dissertation. Participants performed the behavioural task at the beginning, middle and end of such EEG experimental session. Thus, participants were presented three times with the 42 different versions of the sequence. In each of these times, the different versions of the sequence were presented in a random order. Participants were asked to rate a target chord in a seven-point Likert scale (1 for very incorrect/very unpleasant, 7 for very correct/very pleasant). Before the beginning of the task, participants were presented with some examples of the sequences so they would get used to the task and we made sure that, especially non-musicians, understood the meaning of "each chord". Before each sequence a text appeared on the screen with the instructions, which told them which chord they had to focus on in the following sequence (either the third or fifth chord) and which question they had to answer (to give answer to the judgements of correctness or pleasurableness). Some examples of instructions would be "*Rate from 1 to 7 how CORRECT is the THIRD chord that you will hear*" or "*Rate from 1 to 7 how PLEASURABLE is the FIFTH chord that you will hear*". Participants could choose whether the instructions were in Catalan, Spanish or English, to ensure their understanding. The type of judgement was randomly assigned to each chord sequence and presented in a random order. Then, a cross appeared on the screen indicating the beginning of the sequence. After hearing the sequence, they had to enter their response on the number pad of the keyboard.

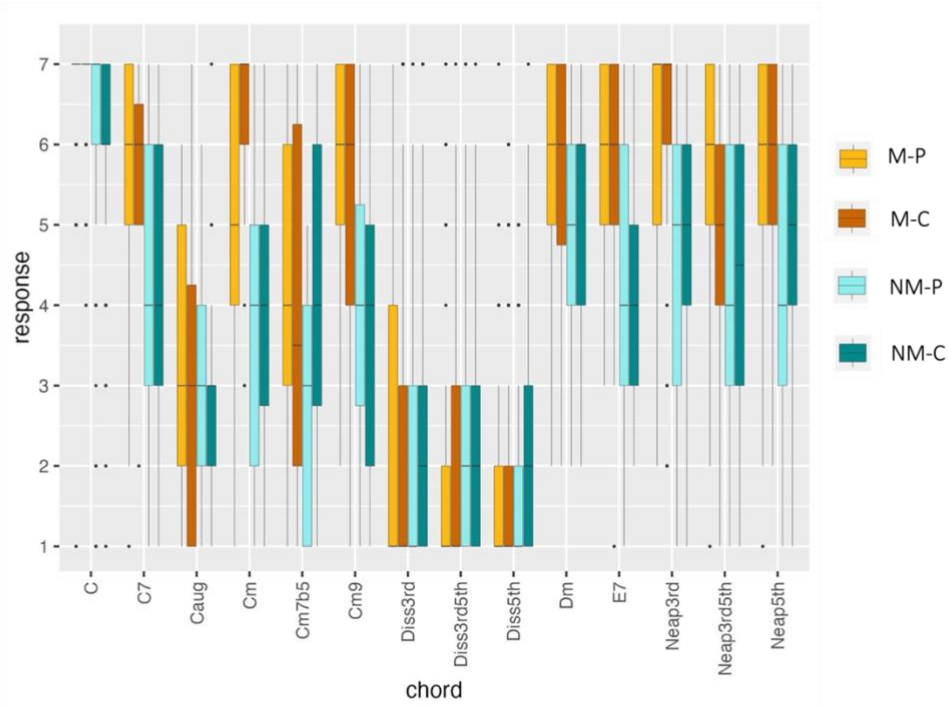


Figure 10. Distribution of responses toward different types of ending chords for the evaluation of correctness (C) and pleasurableness (P) done my musicians (M) and non-musicians (NM). Labels on the x axis depict the different types of endings, from left to right (C, C7, Caug, Cm7b5, a dissonant chord at the third position, a dissonant chord both at the third and fifth position, a dissonant chord at the ending position, Dm, E7, a Neapolitan chord at the third, third and fifth or only fifth position. Error bars are depicted.

2.3 Results

We performed a linear mixed-effect model (LME) to analyse our data. To do so, we used R statistical language (R Core Team, 2012) and the lme4 package (Bates, Maechler & Bolker, 2012). The LME allowed us to assess the relationship between the evaluation of chords and our factors of interest: type of rating (called "Condition" for simplicity), musicianship (called "Group") and chord type (called "Chord").

Figure 10 shows the tendency that musicians tend to provide higher ratings to all chords. Note that the variability in musicians' ratings of the tonic chord C is very small, indicating a high consistency within group. In some cases, the boxplot could suggest that pleasurableness ratings differ from correctness ratings in different chords for the different groups (for instance, for the chord Cm in musicians and the chord Cmb5 in non-musicians). We analysed in depth these interactions in the following sections.

2.3.1 Analysis 1: do affective and cognitive judgements dissociate as a function of musicianship?

In order to test our first hypothesis, that is, whether there is a dissociation between pleasurableness and correctness ratings of chords and whether that is modulated by musicianship, we performed a linear mixed model estimated using REML and nloptwrap optimizer (where t-tests used Satterthwaite's method). As fixed effects, we entered condition, group and their interaction. As random effects, we had intercepts for each subject and for each chord type and random slopes for the by-participant effect of condition and the by-chord effect of condition and group (Table 1, formula: response ~ group + condition + group:condition + (1 + condition | participant) + (1 + condition + group | chord)). Since we were not interested in the effect of chord itself (namely, the differences in ratings between chords) we did not enter chord as a fixed effect, but we had to control for it because we could expect different chords to elicit different effects on the ratings.

95% Confidence Intervals (CIs) were computed using a Wald t-distribution approximation and pvalues were calculated with Kenward-Roger's method (see Table 1). The model's total explanatory power was substantial (conditional R2 = 0.60) and the part related to the fixed effects alone (marginal R2) is of 0.04.

Formula of the model	**npar**	**AIC**	**BIC**	**logLik**	**deviance**	**Chisq**	**Df**	**Pr (>Chisq)**
response ~ 1 + (1 | participant)	3	21176	21195	-10584.9	21170			
response ~ group + (1 | participant)	4	21666	21192	-10578.9	21158	11.894	1	p < .001
response ~ condition + (1 + condition | participant)	6	21001	21040	-10494.6	20989	168.670	2	p < .001
response ~ group + condition + (1 + condition | participant)	7	20989	21035	-10487.7	20975	13.867	1	p < .001
response ~ group + condition + group:condition + (1 + condition | participant)	8	20981	21033	-10482.3	20965	10.686	1	0.001
response ~ group + condition + group:condition + (1 + condition | participant) + (1 + condition + group | chord)	12	17242	17333	-8606.8	17214	3.750.958	6	p < .001

Table 1. List of models created where fixed and random factors were increasingly included to evaluate goodness of fit.

Once the most optimal model was chosen, we aimed to assess the goodness of fit of each fixed factor (or, in other words their significance). To do so, we performed Likelihood Ratio Tests (or LRTs) by comparing compared the goodness of fit based on the ratio likelihoods of our full model versus each null model. We report the results obtained by using Kenward-Roger's method, which is a more conservative approximation than LRTs (see Table 2).

	Sum Sq	Mean Sq	Num DF	Den DF	F value	Pr (<F)
group	16.679	16.679	1	40.105	9.089	p < .001
condition	1.020	1.020	1	18.685	0.556	0.465
group:condition	3.518	3.518	1	40.269	1.917	0.174

Table 2. Output of Kenward-Roger's method to estimate the goodness of fit of fixed factors of the linear mixed model for group and condition.

The results in Table 2 suggests that the effect of Group ($F = 9.089$, $p < 0.001$) was significant. However, the effect of Condition and the interaction between Group and Condition were non-significant. We suggest that the responses were not different neither between conditions (across all groups), nor between groups (across all conditions).

The model's intercept corresponding to Group = musicians and Condition = correctness, was at 4.709 (95% CI [3.75, 5.66], t(4907) = 9.896, p < .001, see Table 3). Within that model, the effect of group was statistically significant and negative (beta = -0.91, 95% CI [-1.43, -0.37], t(4907) = -3.326, p = .002). This suggests that musicians rated chords significantly higher than non-musicians. That might simply be an effect of being sure or more confident about their answers, while non-musicians may stick to intermediate values. The effect of Condition and the interaction between of Condition and Group were statistically non-significant. These results suggest that we do not have evidence of a dissociation between the judgements of pleasurableness and correctness of chords. Maybe that is related to differences in the ratings depending on the type of chord, because some are more ambiguous than others.

	Coeff	Std. Err	df	t	Pr(>|t|)
Intercept (M, Cond1)	4.709	0.476	15.421	9.896	p < .001
Group2 (NM, Cond1)	-0.901	0.271	40.233	-3.326	0.002
Condition2 (M, Cond2)	-0.009	0.120	29.419	-0.072	0.943
Group2:Condition2 (NM, Cond2)	0.171	0.124	40.269	1.385	0.174

Table 3. Linear mixed model's intercept and coefficients of interest of the combination of the levels of the two fixed factors Condition (Condition 1, Pleasurableness & Condition2, Correctness) and Group (Musicians, M & Non-musicians, NM)

2.3.2 Analysis 2: do listeners judge chord types differently as a function of musicianship?

With this analysis, we aimed to test our second hypothesis, that is, whether musicians and non-musicians differently rated our chords of interest (those used in the electrophysiological studies). These were a selection of all the chords to which they were presented. We fitted a linear mixed model (estimated using REML and nloptwrap optimizer). In this analysis we were not interested in the effect of condition, although it might be a good predictor of our data. As fixed effects we entered the model with the factors Chord, Group, Condition and the interaction between Chord and Group, Condition and Group, and the three-way interaction between Condition, Chord and Group. As random effects, we entered intercepts for subjects as well as by-subject random slopes for the effect of chord type (Table 4, formula: response ~ chord + group + condition + condition:group + condition:group:chord + chord:group + (1 + chord | participant)). The model's total explanatory power is substantial (conditional R2 = 0.77) and the part related to the fixed effects alone (marginal R2) is of 0.63. 95% Confidence Intervals (CIs) were computed using a Wald t-distribution approximation and p-values were computed using Kenward-Roger's method.

The model's intercept, corresponding to chord = C, group = 1 and condition = 1, is at 6.79643 (95% CI [6.46, 7.13], t(2298) = 39.91, p < .001).

Formula of the model	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr (>Chisq)
response ~ 1 + (1 | participant)	3	10441,43	10458,69	-5217,72	10435,43			
response ~ group + (1 | participant)	4	10434,07	10457,07	-5213,03	10426,07	9,362	1	0,002
response ~ group + chord + (1 + chord | participant)	22	7517,051	7643,593	-3736,53	7473,051	2953,02	18	<.0001
response ~ chord + group + chord:group + (1 + chord | participant)	26	7506,508	7656,058	-3727,25	7454,508	18,543	4	<.0001
response ~ chord + group + condition + condition:group + chord:group + (1 + chord | participant)	28	7503,377	7664,43	-3723,69	7447,377	7,131	2	0,028
chord + group + condition + condition:group + chord:group + condition:group:chord + (1+chord|participant).	36	7496,26	7703,329	-3712,13	7424,26	23,117	8	0,003

Table 4. List of linear mixed models used, where fixed and random factors were increasingly included to evaluate goodness of fit with the data.

The Kenward-Roger method to assess the goodness of fit of each fixed effect revealed that the effect of Chord (F = 116.790, $p < 0.001$), the effect of Group (F = 13.451, $p < 0.001$), the effect of Condition (F = 4.284, p = 0.04) and the interaction between Chord and Group (F = 4.117, p = 0.005) and the three-way interaction between Chord, Group and Condition (F = 2.866, p = 0.004) were significant (see Table 5).

	Sum Sq	Mean Sq	Num DF	Den DF	F value	Pr(>F)
Chord	625,674	156,419	4	36,256	116,79	<.0001
Group	16,639	16,639	1	45,718	13,451	<.0001
Condition	5,3	5,3	1	2190,389	4,284	0,039
Group:Condition	0,006	0,006	1	1905,084	0,005	0,945
Chord:Group	21,338	5,335	4	62,188	4,117	0,005
Chord:Group:Condition	28,362	3,545	8	2171,645	2,866	0,004

Table 5. Output of Kenward-Roger's method to estimate the goodness of fit of fixed factors for the group and condition analysis.

We then performed pairwise comparison of marginal means for the fixed effects Chord, Group and Condition (Figure 11). For simplicity, only pvalues are reported in text, for more details on the statistics, see tables 6, 7 & 8. Comparison of marginal means showed that the evaluation of each pair of chords behaved similarly across conditions and groups for most chords. For instance, both groups consistently rated Neapolitans higher than dissonant chords at corresponding positions (Neap3rd versus Diss3rd and Neap5th versus Diss5th) for both conditions. Both groups rated the tonic as more pleasurable and correct than dissonant chords at both the third (p < .001) and fifth (p

< .001) position (although the difference may be larger in musicians). However, there were some differences between groups depending on the chord and the rating condition, that will be reviewed below.

NON-MUSICIANS					
Pleasurableness					
	estim.	**SE**	**df**	**t.ratio**	**p.value**
C - Diss3rd	3,885	0,31	61,93	12,395	<.0001
C - Diss5th	4,281	0,33	53,22	12,893	<.0001
C - Neap3rd	1,608	0,23	67,27	7,01	<.0001
C - Neap5th	1,86	0,26	51,78	7,287	<.0001
Diss3rd - Diss5th	0,396	0,25	114,1	1,608	0,111
Diss3rd - Neap3rd	-2,28	0,31	64,41	-7,401	<.0001
Diss3rd - Neap5th	-2,03	0,31	56,93	-6,489	<.0001
Diss5th - Neap3rd	-2,67	0,33	54,23	-8,136	<.0001
Diss5th - Neap5th	-2,42	0,32	50,84	-7,643	<.0001
Neap3rd - Neap5th	0,252	0,2	67,36	1,237	0,2203
Correctness					
C - Diss3rd	3,706	0,29	46,27	12,723	<.0001
C - Diss5th	3,75	0,32	46,05	11,718	<.0001
C - Neap3rd	1,154	0,22	60,8	5,171	<.0001
C - Neap5th	1,427	0,27	63,13	5,315	<.0001
Diss3rd - Diss5th	0,044	0,2	50,24	0,221	0,826
Diss3rd - Neap3rd	-2,55	0,28	44,37	-9,114	<.0001
Diss3rd - Neap5th	-2,28	0,3	49,64	-7,56	<.0001
Diss5th - Neap3rd	-2,6	0,31	44,28	-8,324	<.0001
Diss5th - Neap5th	-2,32	0,32	50,1	-7,368	<.0001
Neap3rd - Neap5th	0,273	0,21	80,98	1,278	0,205

Table 6. Output of the pairwise comparisons of marginal means for all the combinations of Condition and Chord within the group of non-musicians. Degrees-of-freedom method: Kenward-Roger.

On one hand, non-musicians rated the tonic chord as both more pleasurable and more correct than Neapolitan chords at both the third ($p < .001$) and fifth ($p < .001$) position. They moreover rated both Neapolitan and dissonant chords as similarly pleasurable and correct between the third and fifth position (Table 6).

MUSICIANS					
Pleasurableness					
Contrast	**estim.**	**SE**	**df**	**t.ratio**	**p.value**
C - Diss3rd	4,267	0,3	57,08	14,288	<.0001
C - Diss5th	5,2	0,32	52,1	16,183	<.0001
C - Neap3rd	0,833	0,21	55,81	3,954	0.0002
C - Neap5th	1,182	0,24	44,43	4,956	<.0001
Diss3rd - Diss5th	0,933	0,25	126,6	3,809	0.0002
Diss3rd - Neap3rd	-3,43	0,3	62,05	-11,57	<.0001
Diss3rd - Neap5th	-3,09	0,3	55,53	-10,22	<.0001
Diss5th - Neap3rd	-4,37	0,32	55,38	-13,59	<.0001
Diss5th - Neap5th	-4,02	0,31	52,62	-12,92	<.0001
Neap3rd - Neap5th	0,348	0,19	59,16	1,82	0,0738
Correctness					
C - Diss3rd	4,569	0,29	51,54	15,676	<.0001
C - Diss5th	4,845	0,32	49,19	15,287	<.0001
C - Neap3rd	0,329	0,23	76,33	1,437	1,55E-01
C - Neap5th	0,994	0,28	77,05	3,613	<.0001
Diss3rd - Diss5th	0,276	0,19	46,47	1,457	1,52E-01
Diss3rd - Neap3rd	-4,24	0,27	44,76	-15,51	<.0001
Diss3rd - Neap5th	-3,58	0,3	52,4	-12,01	<.0001
Diss5th - Neap3rd	-4,52	0,3	43,48	-14,94	<.0001
Diss5th - Neap5th	-3,85	0,31	51,42	-12,46	<.0001
Neap3rd - Neap5th	0,665	0,22	93,36	3,088	0,0027

Table 7. Output of the pairwise comparisons of marginal means within the group of musicians for all the combinations of Condition and Chord. Degrees-of-freedom method: Kenward-Roger.

On the other hand, musicians exhibited a similar pattern of responses, but with some interesting exceptions (Table 7). For instance, musicians rated the tonic as more pleasurable than Neapolitans at both the third ($p = .0002$) and fifth ($p < .001$) position, and as more correct than Neapolitan endings ($p < .001$). However, unlike non-musicians, musicians rated Neapolitan chords at the third position as similarly correct as the tonic ending ($p = .155$). Unlike non-musicians, musicians rated Neapolitan chords at the third position as more correct than at the fifth position ($p = .003$), although that difference did not reach significance for judgements of pleasurableness ($p = .074$). Interestingly, while musicians rated dissonant chords as similarly correct across positions ($p = .152$), they rated them as more pleasurable at the third position ($p = .0002$).

The pairwise comparison of musicians versus non-musicians per each Condition and Chord (see Table 8) showed that musicians rated tonic chords as more pleasurable ($p = .010$) and correct (.012) than non-musicians did. Musicians consistently rated Neapolitans as more pleasurable ($p < .001$) and correct ($p < .001$) than non-musicians did, regardless of their position within the cadence. However, the difference between groups was non-significant for dissonant chords for both conditions, although non-musicians tended to rate them slightly higher than musicians.

PLEASURABLENESS					
Chord	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
C	0,707	0,26	49,5	2,695	0,01
Diss3rd	0,324	0,36	62,13	0,895	0,374
Diss5th	-0,213	0,38	53,3	-0,55	0,582
Neap3rd	1,481	0,3	50,38	4,932	<.0001
Neap5th	1,385	0,32	41,51	4,376	<.0001
CORRECTNESS					
C	0,723	0,28	61,87	2,6	0,012
Diss3rd	-0,139	0,32	39,91	-0,43	0,67
Diss5th	-0,371	0,36	40,54	-1,04	0,306
Neap3rd	1,549	0,3	46,82	5,255	<.0001
Neap5th	1,157	0,34	56,73	3,377	0,001

Table 8. Output of the pairwise comparison of groups (musicians vs non-musicians) within each condition and per each chord type. Degrees-of-freedom method: Kenward-Roger.

Pairwise comparison between conditions for each group and chord type (Table 9) suggests that the consistency between pleasurableness and correctness depended on the chord, but also on musicianship. For instance, both groups rated the tonic chord and Neapolitan endings similarly across conditions. Musicians and non-musicians agreed in that Neapolitans at the third position were more correct than pleasurable (although that difference between conditions was bigger in musicians than in non-musicians). Surprisingly, the evaluations for dissonant chords diverged in opposite directions for dissonant chords: while musicians rated dissonant chords at the third position higher in the pleasurableness than in the correctness scale, the opposite was found for non-musicians rating dissonant endings higher. All in all, these differences were subtle.

MUSICIANS					
Chord	**estimate**	**SE**	**df**	**t.ratio**	**p.value**
C	0,119	0,16	2101,04	0,736	0,462
Diss3rd	0,421	0,16	2134,16	2,563	0,01
Diss5th	-0,236	0,16	2156,24	-1,44	0,149
Neap3rd	-0,385	0,15	2152,87	-2,64	0,008
Neap5th	-0,069	0,16	2151,99	-0,43	0,668
NON-MUSICIANS					
	estimate	**SE**	**df**	**t.ratio**	**p.value**
C	0,136	0,17	1660,13	0,796	0,426
Diss3rd	-0,043	0,17	2143,43	-0,25	0,802
Diss5th	-0,395	0,17	2100,77	-2,34	0,019
Neap3rd	-0,317	0,16	2072,44	-2,02	0,044
Neap5th	-0,297	0,16	2157,84	-1,85	0,064

Table 9. Output of the pairwise comparison of conditions (pleasurableness vs correctness) within each group and per each chord type.

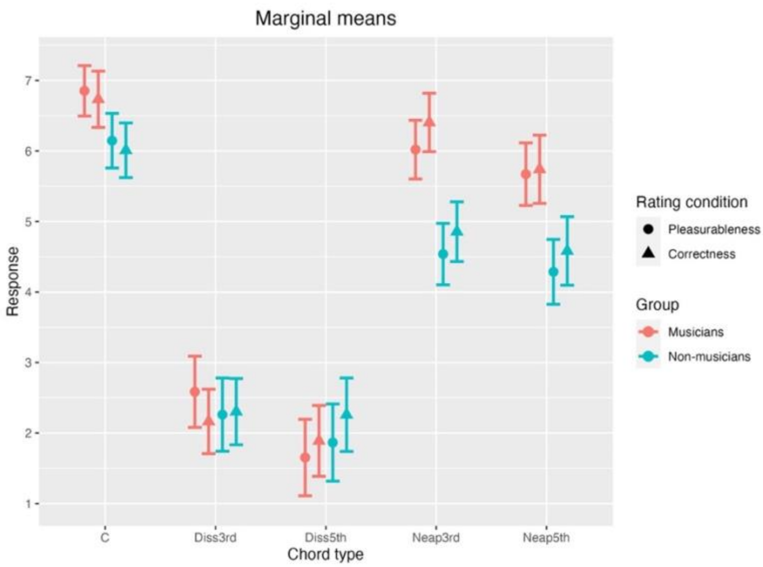


Figure 11. Marginal means of the pleasurableness and correctness ratings of musicians and non-musicians. Error bars are displayed.

2.4 Discussion

The tonic chord as a cadence ending was highly rated by both musicians and non-musicians in both the pleasurableness and correctness scales (McDermott et al., 2010). That provides further evidence on how the preference for consonance is deeply rooted in western musicians (Bidelman 2013; Bones et al., 2014). However, musicians rated the tonic significantly higher than non-musicians, supporting previous research that shows that preference for expected resolutions is enhanced by musical training, because their expectations are stronger (Koelsch, Schmidt & Kansok, 2002). Results on dissonance were also consistent across groups, as both musicians and non-musicians similarly evaluated dissonant endings of chord cadences as very incorrect and unpleasant. These results suggested that, independently of musical expertise, dissonance is considered highly inappropriate as a cadence closure (at least, worse than a tonal-syntactic violation), which is possibly based on universal psychoacoustic constraints (Loui & Wessel, 2007). However, we must consider the possibility that behavioural evaluations of dissonant endings are influenced by a floor effect or a central tendency bias. Most importantly, when evaluating Neapolitan endings, musicians and non-musicians behaved different. Musicians rated Neapolitan endings as significantly more correct and pleasant than non-musicians did, and almost as appropriate as they rated the tonic ending. One explanation to that fact is that in modern and jazz tradition, Neapolitan chords (or bIIMaj7 in modern notation) can be interpreted as delayed resolution. Likely, the exposure of musicians to more complex harmonies might translate into a higher tolerance to tonal modulations and more complex harmonic rhythm structures (Popescu et al., 2019; Virtala & Tervaniemi, 2017).

When comparing between positions, behavioural results showed that musicians rated Neapolitan chords worse as a cadence ending than at the third position, but in non-musicians that difference was non-significant. As suggested by previous research, Neapolitan sixths introduce are less common and introduce a stronger violation of context as a cadence closure than as a substitute of the subdominant (at the third position). Thus, we can argue that for non-musicians, an out-of-key chord simply introduces "a disruption" in the chord cadence. But musicians consistently performed corresponding to the rules of western music (Brattico et al., 2016; Virtala & Tervaniemi, 2017b).

Our results also suggest that dissonance is similarly unacceptable (both at the pleasurableness and correctness level) regardless of the position that it occupies within a cadence for both listeners. That result further argues in favour of the inappropriateness of dissonance within a musical context for western listeners, which is perceived as unpleasant whenever it appears (Loui & Wessel, 2007). Although the difference was non-significant, musicians tended to rate dissonant endings slightly worse than dissonant chords at the third position. That tendency was not observed in non-musicians. Therefore, although dissonance is disliked and considered incorrect, its evaluation by musicians may still interact with harmonic expectation (Brattico et al., 2016; Virtala & Tervaniemi, 2017b). So, to experienced listeners, a dissonant cadence closure may instil stronger disruption than dissonance at intermediate positions of cadences.

2.5 References

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear Mixed-Effects Models Using S4 Classes (R Package Version 0.999999-0). http://cran.r-project.org/web/packages/lme4/index.html

Bidelman, G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology*, *4*, 1–13. https://doi.org/10.3389/fpsyg.2013.00264

Bones, O., Hopkins, K., Krishnan, A., & Plack, C. J. (2014). Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia*, *58*(1), 23–32. https://doi.org/10.1016/j.neuropsychologia.2014.03.011

Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2016). It's sad but i like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Frontiers in Human Neuroscience*, *9*(JAN2016), 676. https://doi.org/10.3389/FNHUM.2015.00676/BIBTEX

Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 48–61. https://doi.org/10.1037/a0031624

Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Genz et al. (2021). mvtnorm: Multivariate Normal and t Distributions. R package version 1.1-3. URL http://CRAN.R-project.org/package=mvtnorm

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*(5), 657–663. https://doi.org/10.1017/S0048577202010508

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception & Psychophysics 2007 69:7*, *69*(7), 1084–1092. https://doi.org/10.3758/BF03193946

Makowski, D., Ben-Shachar, M.S., Patil, I. & Lüdecke, D. (2020). Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. CRAN.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*(11), 1035–1041. https://doi.org/10.1016/j.cub.2010.04.019

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Therneau T (2022). A Package for Survival Analysis in R_. R package version 3.3-1, <https://CRAN.R-project.org/package=survival>.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Makowski, D., Ben-Shachar, M.S., Patil, I. & Lüdecke, D. (2020). Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. CRAN.

Virtala, P., & Tervaniemi, M. (2017b). Neurocognition of Major-Minor and Consonance-Dissonance. *Music Perception: An Interdisciplinary Journal*, *34*(4), 387–404. https://doi.org/10.1525/mp.2017.34.4.387

3. EXPERIMENTAL SECTION 2: ELECTRO-PHYSIOLOGICAL STUDIES

In this section, we ran a series of 3 electrophysiological studies, focusing on the more common event-related potentials (ERPs) related to musical perception and cognition of harmonic unexpectedness. Importantly, we focused on investigating some of the factors that can modulate these well-known neural responses, again with the aim to assess to what extent music processing is universal and imperturbable. Our studies were based on the work by Koelsch et al. (2000) and aimed to prove the effect of musicianship in processing musical unexpected endings. In Study 1, we investigated whether trained and naïve listeners would respond differently to unexpected musical events deviating in different dimensions (tonality, syntax and sensory dissonance). In this study, participants were presented with tonal-syntactic and dissonant irregularities while performing a secondary auditory detection task. Following the results and paradigm of Study 1, in Study 2 we investigated to what extent strong deviations in the dimension of consonance are processed proportionally to accumulated expectation and whether that depends on musicianship. To assess this issue, participants listened to musical irregularities placed at different positions of a musical sequence where expectation progressively increases. Finally, in Study 3 we explored whether plasticity of neural responses can be attained with short-term exposure to irregular events and whether such habituation differently takes place for trained and untrained listeners. To do so, listeners heard chord cadences ending on a deviation from consonance very frequently.

3.1 Manuscript 1

Pagès-Portabella C, Toro JM. Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians. Psychophysiology. 2019 Aug. 27(1). 1-11 DOI: https://doi.org/10.1111/psyp.13476

Received: 10 December 2018 | Revised: 30 July 2019 | Accepted: 8 August 2019
DOI: 10.1111/psyp.13476

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians

Carlota Pagès-Portabella[1] | **Juan M. Toro**[1,2]

[1]Language & Comparative Cognition Group, Center for Brain & Cognition, Universitat Pompeu Fabra, Barcelona, Spain

[2]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Correspondence
Carlota Pagès-Portabella, Language & Comparative Cognition Group, Center for Brain & Cognition, Universitat Pompeu Fabra, C. Ramon Trias Fargas, 25-27, 24.326., Barcelona 08005, Spain.
Email: carlota.pages@upf.edu

Funding information
H2020 European Research Council, Grant/Award Number: 312519; Fundação Bial, Grant/Award Number: 13/18

Abstract

In major-minor tonal music, the hierarchical relationships and patterns of tension/release are essential for its composition and experience. For most listeners, tension leads to an expectation of resolution. Thus, when musical expectations are broken, they are usually perceived as erroneous and elicit specific neural responses such as the early right anterior negativity (ERAN). In the present study, we explored if different degrees of musical violations are processed differently after long-term musical training in comparison to day-to-day exposure. We registered the ERPs elicited by listening to unexpected chords in both musicians and nonmusicians. More specifically, we compared the responses of strong violations by unexpected dissonant endings and mild violations by unexpected but consonant endings (Neapolitan chords). Our results show that, irrespective of training, irregular endings elicited the ERAN. However, the ERAN for dissonant endings was larger in musicians than in nonmusicians. More importantly, we observed a modulation of the neural responses by the degree of violation only in musicians. In this group, the amplitude of the ERAN was larger for strong than for mild violations. These results suggest an early sensitivity of musicians to dissonance, which is processed as less expected than tonal irregularities. We also found that irregular endings elicited a P3 only in musicians. Our study suggests that, even though violations of harmonic expectancies are detected by all listeners, musical training modulates how different violations of the musical context are processed.

KEYWORDS
dissonance, ERAN, ERPs, harmonic expectations, musical training

1 | INTRODUCTION

A central part of tonal Western music composition and experience is the hierarchical relationships and patterns of tension/release that are created by the combination of certain chords (Meyer, 1956). When chords are combined into sequences, they can create tensions (e.g., the movement from the supertonic to the dominant) and lead the listener to expect a resolution of these (Bharucha, 1984). In the present study, we aim to explore how the brain of both musicians and nonmusicians reacts when these expectations are violated in different ways.

For any listener, the understanding of music requires the integration of new upcoming information, and for that to happen, a context has to be built. Within a musical context, a central stable pitch attracts the others, constituting the tonal center or tonic. The combination of the notes surrounding the tonic forms the tonality or key. It is precisely the tonality that provides the appropriate context for us to understand music (Meyer, 1956). Within this context, a hierarchy of stability is established (Bharucha & Krumhsansl, 1983). In the first level of the hierarchy, some tones are perceived as more stable than others, depending on the relation to the tonic (e.g., nonscale tones are the least stable because they are the most remotely related to the tonic; Bharucha, 1984). The second level of the harmonic hierarchy appears when multiple notes are played at the same time, forming chords. Each tone establishes hierarchical relationships with the other tones in the chord (e.g., in the tonic triad of C [*c-e-g*], the note *c* is perceived as the more stable tone). Finally, the third level of the hierarchy is based on the relation between the chords and the distance to the tonal center. Each chord performs a function inside a tonality, where a tonic triad (*c-e-g*) is more stable than a mediant (*e-g-b*), and this in turn is more stable than a dominant (*g-b-d*), even if they share notes. Thus, when chords are combined into sequences, this hierarchy of stability provides the appropriate context to create harmonic expectations.

Previous research on music perception has focused on the violation of musical expectations. Moving away from the tonal center is perceived as tensioning, and going back to it is perceived as tension releasing (Koelsch, Gunter, & Friederici, 2000). In general, when chords break the expectations built by the previous context, they are perceived as erroneous (Berent & Perfetti, 1993) and can create negative affective reactions (Brattico, Jacobsen, De Baene, Glerean, & Tervaniemi, 2010). Different ERPs have been linked to the processing of expectancy violations in music. Unexpected endings of chord sequences elicit an early right anterior negativity (ERAN) peaking around 200 ms after the onset of the ending chord (Koelsch et al., 2000). This component has been taken to reflect the fast and automatic processing of music-syntactic irregularities. The ERAN can be followed by a late frontal bilateral negativity with the onset around 500 ms (the N5). Koelsch and collaborators claimed that the N5 may reflect the amount of semantic integration because it is functionally and morphologically reminiscent of the N400 component (Kutas & Dale, 1997), the former being related to music and the latter to language. Their amplitude is enhanced when the integration of an event into the previous context is difficult, and thus unexpected ending chords elicit a larger N5 than expected chords. Musical irregularities can also elicit the frontal P3a, which appears for unlikely and unexpected sounds that attract the attention of the listener (Koelsch, 2000). Thus, broken expectations elicit distinct neural responses that include the ERAN, the N5, and the P3a, among others. However, an open issue is how the degree of violation of musical context can affect these neural responses to expectation violations.

Ending chords can deviate from given context in different ways. On the one hand, unexpected within-key tones are a violation at the syntactic level and out-of-key tones represent a violation at the tonality level. On the other hand, dissonant chords violate the harmony rules that establish the distance between simultaneous tones to form chords. Most chords in music are triads (with a tonic, a third 3 or 4 semitones apart, and a fifth 3 semitones apart), while clusters are not and hence are nonharmonic. There are thus two different levels of expectation deviation: mild and strong or (to put it differently) syntactic and nonsyntactic. Different levels of context violation can be reflected in ERPs. For instance, the amplitude of N5 is larger for ending clusters than for syntactically unexpected endings (like the Neapolitan sixth). Clusters at the third position of a chord sequence elicit larger ERAN and N5 than Neapolitans (Koelsch et al., 2000). This result supports that clusters represent a stronger violation than syntactic violations even when the tonal context is not well defined, and, consequently, the expectation for resolution is weaker. Moreover, Leino, Brattico, Tervaniemi, and Vuust (2007) showed that syntactic violations elicit the ERAN while nonsyntactic violations (mistuned chords with the fifth altered by 50 cts) elicit the mismatch negativity (or MMN) because they introduce frequencies deviating from the tempered scale. While clusters do not represent a frequency deviant, the matter whether nonsyntactic violations of context elicit the ERAN or a MMN is under current debate.

Koelsch (2009) first argued that the ERAN is a kind of "abstract feature" MMN (Saarinen, Paavilainen, Schöger, Tervaniemi, & Näätänen, 1992), elicited by the abstract feature "in key/out of key." Yet, there are functional differences between them: while the MMN is based on sound relationships extracted online, the ERAN depends on the culturally biased mental schema of regularities stored in the long-term memory (Garza-Villarreal, Brattico, Leino, Ostergaard, & Vuust, 2011). In oddball paradigms, dissonant chords elicit the MMN (Crespo-Bojorque, Toro, & Monte-Ordoño, 2018; Virtala et al., 2011), because they are infrequently occurring tones that do not recruit higher-order processes. But when placed in a chord sequence, they violate the stored template of tonal music and elicit an ERAN (as shown in Koelsch et al., 2000). Importantly, Koelsch and collaborators demonstrated that dissonant clusters (as well as Neapolitans) at the fifth position of a chord sequence elicit a bigger ERAN than at the third position. Meanwhile, the amplitude of a MMN should not change (Koelsch, 2000; Koelsch, Schmidt, & Kansok, 2002; Koelsch et al., 2001). Such results show that the processing of clusters is music rule based and interacts with the establishment of tonal context. Therefore, we should expect to find the ERAN for both mild and strong violations. However, different musical experiences may modulate its features.

Regarding the effect of long-term musical training, the ERAN has been shown to be larger in musicians than in musically naive listeners (for Neapolitans; Koelsch et al., 2002). The more specific and robust musical expectancies of musicians could cause stronger reactions to violations. MMN can be enhanced by musical training (with slightly deviant pitches, see Koelsch, Schröger, & Tervaniemi, 1999; with dissonant intervals see Crespo-Bojorque et al., 2018), but not necessarily (it is indistinguishable for timbre deviants; Koelsch et al., 2002). Since clusters in a chord sequence elicit the ERAN and musicians show an enhanced ERAN (but not necessarily enhanced MMN), we believe that musicians could show a larger ERAN for dissonant clusters. If so, it would indicate that musicians react to dissonance based on an elaborated music rule processing rather than on an enhanced general auditory sensitivity (Koelsch et al., 2002).

In summary, the aim of the present study is to test how the neural responses to different degrees of musical unexpectedness are influenced by long-term musical training in comparison to day-to-day musical exposure. To tackle this issue, we will compare the responses elicited by three different types of endings of chord progressions: (a) expected tonic, (b) mild violations by out-of-key chords equivalent to Neapolitan sixths, and (c) strong violations by dissonant chords. We will compare the responses to these different levels of violation between musicians and nonmusicians. Our hypothesis is that musicians will be more sensitive to different degrees of violations than nonmusicians. According to the literature, long-term formal musical training leads to a more fine-tuned auditory processing together with a robust knowledge of music rules. We thus expect that musicians will show larger neural responses to irregular endings. Moreover, our prediction is that musicians' ERAN will be larger for dissonant endings than for mild violations while the responses of nonmusicians to these different endings will not differ.

2 | METHOD

2.1 | Participants

Twenty-eight volunteers participated in the study. Almost half of the participants ($N = 13$) were professional musicians (5 female; mean age 29.9 years ± 7.97). Musicians started musical education at 8.1 years of age (± 3.84) and had a mean of 21.1 years (± 7.99) of formal musical training in Western harmony. They all specialized in an instrument (mostly guitar, violin, and piano) and were musically active at the time of the experiment. The rest of the participants ($N = 15$) were nonmusicians (13 female; mean age 24.3 years ± 6.03). Nonmusicians did not have musical training besides normal school education and had not learned to play an instrument. Twenty-five out of 28 participants reported to have right-hand preference, two were left-handed, and one was ambidextrous (according to the Edinburgh Handedness Inventory; Oldfield, 1971). All of them reported normal (or corrected-to-normal) vision and normal hearing. They signed a written consent and received a monetary compensation for their participation in the study. The data from four participants were discarded because of extreme drifts and artifacts. Only the data from 12 musicians and 12 nonmusicians were analyzed.

2.2 | Stimuli

The stimuli were 324 unique five-chord sequences. Each chord was composed of four notes (in respect to the tonic: fundamental, third, fifth, and seventh). The chord at the first position in each sequence was always the tonic of the following cadence. The second chord could be the submediant (sixth grade minor seventh or VIm7) or the subdominant (fourth grade major seventh or IVMaj7). The third chord could be the supertonic (second grade minor seventh or IIm7) or the subdominant (fourth grade major seventh or IVMaj7). The fourth chord was always the dominant seventh (V7) to create the maximum expectation for resolution. The ending fifth chord could be of three different types: expected, ambiguous, or dissonant. The expected ending was the tonic. The ambiguous ending was a major seventh chord built on the second flat grade of the tonality (bIIMaj7), equivalent to the Neapolitan sixth. Importantly, even though Neapolitan chords are unexpected chords, they are considered to be ambiguous because they can be appealing to listeners by creating a suspended sensation and introducing an aesthetically pleasurable degree of novelty (Berlyne, 1971; Brattico et al., 2010). The dissonant ending was a nonharmonic cluster, which kept the fundamental of the tonic but had three notes separated by a semitone (in respect to the tonic, the notes were the third, the fourth, and the augmented fourth). In order to add more variability to the stimuli, each sequence was modified so that the first chord was either in root position, first inversion (third in the lower voice), or second inversion (fifth in the lower voice). The top voice always moved by conjunct melodic motion, never jumping more than a tone or a semitone. All the sequences were transposed to the 12 different keys.

Fifty per cent of the sequences (108 unique sequences, each played twice for a total of 216 sequences) had expected endings, 25% of the sequences (108 sequences) had ambiguous endings, and 25% of the sequences (108 sequences) had dissonant endings. Within each sequence, Chords 1 to 4 lasted for 600 ms each (Koelsch et al., 2000, 2002) and Chord 5 lasted for 1,200 ms. Each sequence lasted for 3,600 ms. All sequences were recorded with Cubase LE AI Elements 9.5 using the piano instrument. In order to create a secondary task (see procedure below), 10% of the sequences (four randomly selected sequences of each tonality) included a chord in the second or third position played by a deviant instrument other than a piano (violin, guitar, or vibraphone). The 36 sequences belonging to the same key were presented together in a block. Thus, there were 12 blocks (one per key). The order of presentation of blocks was randomized.

2.3 | Procedure

The experiment was run in an acoustically and electrically shielded room. Participants sat comfortably in an armchair and listened to the sequences, presented via headphones using MATLAB. Participants were told not to blink or move their body or their eyes while listening. They were instructed to look at a fixation cross at the center of the screen. We presented the sequences with a silent period of 500 ms between them, during which three stars appeared on the screen. That break between sequences performed a double function: first, it allowed participants to blink between sequences, which prevented having artifacts overwritten in the signal of interest. Second, a pause prevented the responses to the last and longer chord from overlapping with those elicited by the first chord of the next sequence. In order to keep the participants engaged while listening to the sequences, we used a secondary task. The participants were asked to detect and count how many times they heard an instrument different than a piano. They were not informed about the presence of irregular ending chords. Twice in each block (every 18 sequences), the participants were asked to provide a response by a text showed on the screen. They had to respond with the right arrow key on the keyboard if they heard more than one sequence produced by a different instrument. They were asked to respond with the left key otherwise. At the end of each block, there was a 10-s countdown as a break. There was also a longer break at the middle of the experiment so participants could move, after which participants pressed the space bar to continue with the experiment. The duration of the entire experimental session was approximately 35 min.

2.4 | Recording

EEG measurements were recorded with a 32-channel ActiCAP with Ag/AgCl electrodes. Twenty-eight scalp locations following the 10-20 system were recorded (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, Oz). Two electrodes were placed on the left and right mastoid (MSDL, MSDR) and two more on the outer side and below the right eye in order to monitor the ocular movements and blinking. Measurements were referenced online to the tip of the nose and offline to the average of the mastoids. The sampling rate was 500 Hz, and impedances were kept below 10 kΩ. Data were recorded using BrainVision Recorder, and the triggers were sent simultaneously with MATLAB.

2.5 | Analyses

Data were preprocessed offline with Fieldtrip Toolbox: filtered from 0.4 to 40 Hz and electrode corrected via neighbor interpolation. Elimination of artifacts caused by eye, heart, and muscular movement was done via independent component analysis (ICA) and manual removal of the components reflecting the artifacts. On average, 2.2% of trials were rejected from further analysis because of residual excessive noise. Epochs of 1,200 ms were extracted for all ending chords. A relative baseline correction was applied from −200 ms to the onset of the chord by calculating the baseline amplitude separately for each condition. A visual inspection of the ERPs identified stronger responses in anterior-frontal areas (especially in F8 electrode; see also Koelsch, 2000, Koelsch, Jentschke, Sammler, & Mietchen, 2007; Koelsch et al., 2002). Based on the difference waves (dissonant minus correct and dissonant minus ambiguous) of group averages, we chose the time windows of interest: 100–150 ms for N1, 150–250 ms for ERAN, and 325–500 ms for P3. We did not observe a negativity consistent with the N5; therefore, no data from the time window of N5 were included in the analyses. Additionally, we ran a hypothesis-driven permutation test (Crespo-Bojorque et al., 2018; Kalashnikova, Varghese, Di Liberto, Lalor, & Burnham, 2018; Maris & Oostenveld, 2007) separately for each group (musicians, nonmusicians) and each irregular ending (ambiguous, dissonant) with the data of F8 electrode as we expected to observe an ERAN and a P3a in frontal electrodes. For musicians and nonmusicians, we calculated the mean difference between regular and each irregular ending (dissonant-correct and ambiguous-correct) for each time point. Then, we mixed the data of the regular with each of the irregular conditions and randomly assigned them to two parts, with the labels correct versus ambiguous or correct versus dissonant assigned to each part. One thousand reassignments were run, and we obtained a p value that reflected the probability of obtaining similar differences. The permutation tests revealed significant responses at the different time windows that support the manually selected time windows of interest (see Table 1).

To statistically compare the ERPs across conditions, we used the local peaks measurement. First, for each time bin within the time windows of interest, we averaged across electrodes of the regions of interest. Then, for the individual ERP waveform of each chord, a local peak was considered the value larger than the two neighboring samples. If there were many local peaks, the biggest was chosen. To compare the local peak amplitudes of ERAN and P3 across conditions, we performed mixed-design repeated measures analysis of variance (ANOVA) with the factors region of interest (ROI; right-frontal electrodes, F4, F8, FC2, FC4; left-frontal electrodes, F3, F7, FC1, FC3) and chord types (correct, ambiguous, dissonant) as within-subject factors, and group (musicians, nonmusicians) as a between-subjects factor. Additionally, to explore a possible N1 response, we performed two mixed-design ANOVAs with the factors Chord Type × Group over

TABLE 1 Significant time windows and peak amplitudes of the difference waves of each comparison

	Time window (ms)		Peak amplitude (μV)	
	ERAN	P3	ERAN	P3
Musicians				
Correct-dissonant	164–254	324–468	−3.23 (0.30)	2.50 (0.26)
Correct-ambiguous	176–222	340–412	−2.70 (0.37)	2.57 (0.26)
Nonmusicians				
Correct-dissonant	188–248		−2.64 (0.19)	2.35 (0.51)
Correct-ambiguous	164–226		−2.47 (0.36)	1.38 (0.35)

Note: Time windows were extracted from permutation tests on F8. Peak amplitudes and standard deviation of the mean (*SEM*) of difference waves (dissonant-correct, ambiguous-correct) for right-anterior electrodes and time windows 150–250 and 325–500 ms are shown.

the mean amplitude observed between 100 and 150 ms: one with the three ending chords and one including all the chords of the sequence (context Chords 1, 2, 3, 4 and the three different endings). All results were corrected with Greenhouse-Geisser when the sphericity was violated. In all analyses, Bonferroni correction was applied on multiple comparisons.

3 | RESULTS

3.1 | Early components

3.1.1 | N1

For the time window 100–150 ms, the Chord Type × Group ANOVA over the local peak amplitudes revealed no differences across groups or ending chords. However, the Chord Type × Group ANOVA over the mean amplitudes including all the chords revealed a main effect of chord type, $F(6, 22) = 27.024$, $p < .001$. Post hoc tests revealed that the largest early negativity was observed after Chord 1 ($p < .001$). Chords 2–4 elicited a bigger negativity than correct endings ($p < .001$). Both ambiguous and dissonant endings elicited a larger N1 than correct endings ($p < .001$) but undistinguishable from Chords 3 and 4. When analyzing only the ending chords, the ANOVA revealed a smaller N1 for correct than for ambiguous and dissonant endings, $F(2, 44) = 8.574$, $p = .001$. No significant difference was found between musicians and nonmusicians (see Figures 1 and 2).

3.1.2 | ERAN

For the time window 150–250 ms, the ROI × Chord Type × Group ANOVA revealed a main effect of chord type, $F(2, 44) = 30.093$, $p < .001$, and an effect of group, $F(1, 22) = 5.811$, $p = .025$. We also observed an interaction between chord type and group, $F(2, 44) = 5.619$, $p = .007$, and a marginal effect between chord type and ROI, $F(2, 44) = 3.004$, $p = .060$. The significant effect of chord type indicates that both dissonant and ambiguous endings elicit a negativity in comparison to correct endings ($p < .001$). These results are consistent with the appearance of an early right anterior negativity, the ERAN in both groups of participants (see Figures 1 and 2). The main effect of group indicated that the amplitude of the ERP in this time window was more negative in musicians than in nonmusicians for all chord types (see Figure 3). Moreover, the interaction between group and chord type reflected that dissonant endings elicited a larger negativity in musicians than in nonmusicians ($p = .003$). Even more, the difference between the negativities elicited by ambiguous and dissonant endings is significant only in musicians ($p = .016$), suggesting a discrimination of the degree of violation that is not found in nonmusicians. The almost significant interaction between chord type and ROI could suggest that the early negativity tended to be right lateralized, as can be clearly seen in Figure 2. Although in musicians the ERAN seemed to be more widely distributed (see Figure 1), the triple interaction of Chord Type × Group × ROI was nonsignificant, suggesting that there were no differences between musicians and nonmusicians in the response lateralization.

3.2 | Late components

For the time window 325–500 ms, the ROI × Chord Type × Group ANOVA revealed an effect of chord type, $F(2, 44) = 8.734$, $p = .001$, showing that irregular endings (ambiguous and dissonant) elicit a positivity in comparison to regular endings ($p = .005$) that bears resemblance to the P3. There was an interaction between chord type and group, $F(2, 44) = 3.365$, $p = .044$. Post hoc effects showed that the bigger positivity for irregular endings is only present in musicians ($p < .001$; see Figure 1). Although a small positivity was observed in nonmusicians, too, it was not significant (see Figure 2). Also, it showed that the positivity elicited by ambiguous endings has bigger amplitude in musicians than in nonmusicians ($p = .030$). The main effect of ROI, $F(1, 22) = 9.843$, $p = .005$, suggested that the positivity is right localized.

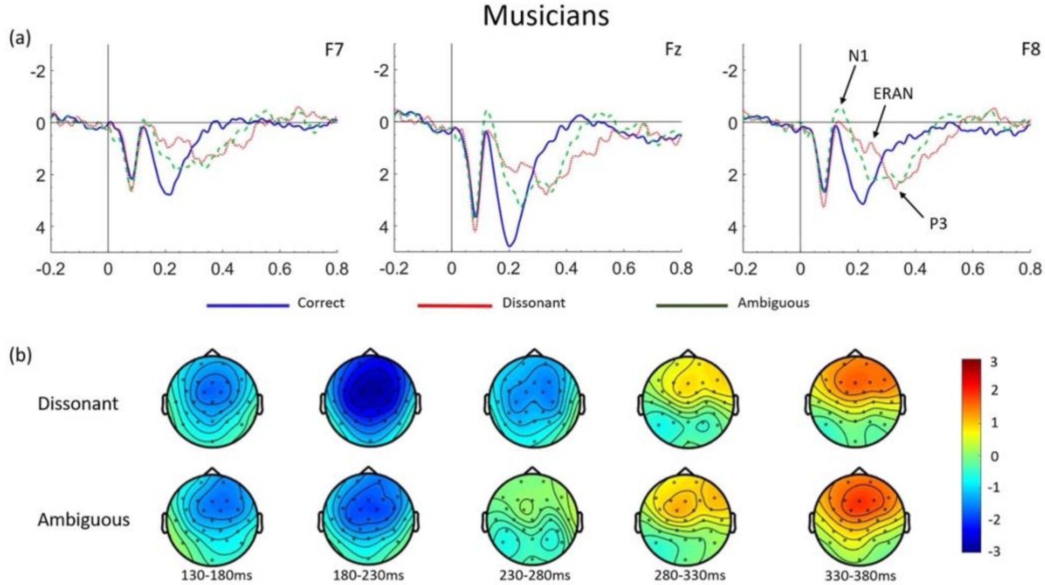


FIGURE 1 ERPs and scalp topographies elicited by correct, dissonant, and ambiguous endings in musicians. ERPs are baseline corrected relative to the average at each electrode. (a) ERPs of musicians at frontal electrodes representing left, central, and right positions. We can observe N1, ERAN, and P3, especially clear at F8. (b) Scalp topographies of the difference waves (dissonant-correct and ambiguous-correct). At around 200 ms, there is a clear anterior negativity (slightly bilateral for dissonant endings) and around 300 ms, a clear right-preponderant frontal positivity

3.3 | Behavioral data

Based on self-reported comments of participants, we suggest that the task was easy for both groups. Most participants reported that the task was easy enough to pay attention to deviant instruments while involuntarily detecting irregular chords (and some musicians even expected them), but none reported getting distracted from the main task. Timbre deviants (second or third) never overlapped in time with irregular endings (fifth position); hence, it is unlikely that irregular endings elicited a distraction (Escera & Corral, 2007). However, the main task stored binary data of only "1 or more than 1 deviant instruments"; therefore, we could not directly assess task difficulty from performance results.

4 | DISCUSSION

In the present study, we explored whether musical training shapes how the brain reacts to different degrees of violations of harmonic expectancies. We thus presented musicians and nonmusicians with common chord progressions that generated high expectancies of resolution to the tonic chord. We registered the ERPs triggered by different possible endings that broke these expectations. There were two degrees of violation: going to an unexpected but consonant ambiguous ending (mild violation created by a Neapolitan chord) or to an unexpected and highly dissonant ending (strong violation by dissonant clusters). We observed that irregular endings elicited an anterior negativity peaking around 200 ms with a tendency to be right lateralized, consistent with the ERAN (Koelsch 2011; Koelsch et al., 2000, 2002, 2007). In our study, the early negativity was weakly lateralized. However, it has been shown that the lateralization of the ERAN can be even absent (Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005, or termed as EAN; Steinbeis, Koelsch, & Sloboda, 2006), differ between genders (Koelsch, Maess, Grossmann, & Friederici, 2003), or be weaker for less salient musical irregularities (Koeslch et al., 2007). In musicians, the ERAN was larger for dissonant endings than for ambiguous endings. Also, dissonant endings elicited a larger ERAN in musicians than in nonmusicians. Moreover, both irregular endings elicited a frontal positivity at around 350 ms significant only in musicians. These differential neural responses may reflect processing advantages for harmonic chord progressions deriving from musical training.

The ERAN has been described as a response to music-syntactic irregularities based on the relationship of the target chord with the preceding context (Koelsch et al., 2007; Leino et al., 2007). In our study, both types of irregular endings (Neapolitans and dissonant clusters) elicited the ERAN in all participants (musicians and nonmusicians). Thus, our study provides support to the idea that the human brain automatically

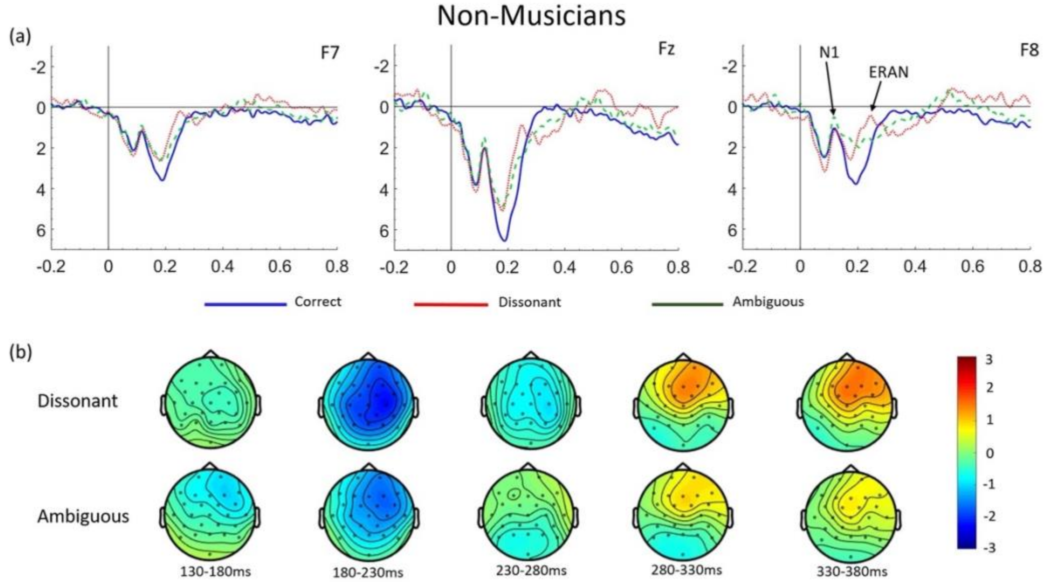


FIGURE 2 ERPs and scalp topographies elicited by correct, dissonant, and ambiguous endings in nonmusicians. ERPs are baseline corrected relative to the average at each electrode. (a) ERPs of musicians at frontal electrodes representing left, central, and right positions. We can observe N1, ERAN, but not a clear P3, although there is a tendency (especially for dissonant endings at F8). Note that the scale of the vertical axis of the ERPs is different from Figure 1. (b) Scalp topographies of the difference waves (dissonant-correct and ambiguous-correct). At around 200 ms, there is a clear right-anterior negativity at F8 and around 300 ms, a tendency to a frontal positivity

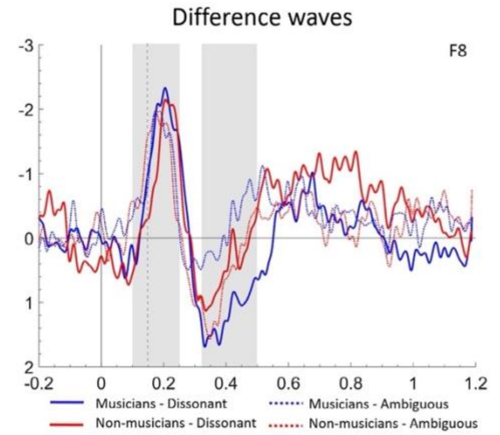


FIGURE 3 Difference waves (irregular subtracted from regular endings) of the ERPs for musicians and nonmusicians over the electrode site F8. Shaded areas represent time windows of interest: 100–150, 150–250, 325–500 ms

responds to musical violations (e.g., Koelsch et al., 2000, 2002). Importantly, we also found that, only in musicians, the amplitude of the ERAN was larger for dissonant than for ambiguous endings. On the contrary, we observed no significant differences across endings in nonmusicians. Thus, our study suggests that only musicians activate a neural mechanism able to distinguish between different kinds of musical irregularities.

Koelsch et al. (2000, 2007) showed that the ERAN can be elicited both in the presence of physical deviance (with Neapolitan sixths and dissonant clusters) or without it (with a dominant-supertonic V-IIm cadence). In the present study, irregular chords differently matched the auditory memory traces of the previous context. For instance, Neapolitans repeated two notes of the context (in C major, *f* and *c*), while clusters repeated three (*c, f,* and *g*). Of these repeated notes (and in comparison with the syntactically expected tonic: *c, e, g,* and *b*), Neapolitans contained only one tone of the tonic (*c*), whereas clusters contained two (*c* and *e*). Moreover, Neapolitans introduced two out-of-key notes (*d flat* and *a flat)* while clusters introduced only one (*f sharp)*, matching better with the tonality. Thus, clusters could better match the context than Neapolitans (in terms of pitch repetition) but the relation of high dissonance between their tones negatively affects their perceptual similarity (in terms of consonance). Dissonant clusters include two successive semitone intervals, which have been rated as the most dissonant intervals (Bidelman, 2013) and have the highest sensory dissonance based on roughness

(Proverbio, Orlandi, & Pisanu, 2016). Consistently, Tillmann, Janata, Birk, and Bharucha (2008) showed that pitch repetition is not the only factor that determines congruency with the context. In their priming paradigm, pitch repetition facilitated the processing of the expected tonic but could not explain the responses expected by the tonal hierarchy. Therefore, perceptual similarity may not be fully described by pitch commonality (perhaps other factors, such as relaxation seeking, should be taken into account). Because of their high degree of dissonance, clusters become perceptually unrelated and, therefore, an inappropriate ending for the context. If clusters were processed in relation to the chord sequence, they would represent a strong violation of context and could have triggered a bigger ERAN. Moreover, while the perception of dissonance is known to be fairly universal (Fritz et al., 2009; Virtala & Tervaniemi, 2017), a dissonant chord could be more salient for trained musicians than for naïve listeners. Musicians have better chord perception abilities than nonmusicians, as seen in Bowling, Purves, and Gill (2017), where they distinguish more and more accurately between chords with very subtle differences of consonance.

4.1 | The degree of violation and musical training

Different unexpected elements may create multiple degrees of violation of context. In our study, Neapolitans violated tonality (with out-of-key notes) and syntax, because they were presented after a dominant (which has a tendency to resolve to the tonic; Leino et al., 2007). On the contrary, dissonant clusters after a dominant chord violated both syntax and tonality and harmony rules. Thus, dissonant clusters would represent a stronger violation of context than Neapolitans. In fact, dissonant chords elicit a much larger N5 than Neapolitans in nonmusicians because strong violations are harder to integrate in the musical context (Koelsch et al., 2000). In contrast, we did not observe the N5. Koelsch and collaborators reported that clusters at the third position of the sequence elicit bigger ERAN and N5 than Neapolitans, suggesting that clusters are so dissonant that they elicit considerable brain responses at a position where the context is not very well defined. However, they also stated that amplitude of ERAN at the fifth position was already maximal with mild violations, and suggested that, when the tonal context is well defined, ERAN is not affected by the degree of violation. Nonetheless, dissonant clusters were not tested with musicians, leaving the door open to the possibility that differences in sensitivity to the degree of violation arise from musical expertise.

Previous findings suggest that formal musical training enhances the neural responses to unexpected musical stimuli. Koelsch et al. (2002) showed that the ERAN is larger in musicians than in nonmusicians, likely because of more specific and robust musical expectancies based on explicit tonal knowledge. Consistently, our results suggest that musicians have not only stronger but also more sensitive reactions to different degrees of musical violations than nonmusicians. As the ERAN reflects the degree of violation of context, we suggest that the difference in amplitude between clusters and ambiguous endings only in musicians demonstrate that they reacted to dissonance as a stronger violation. In contrast, the lack of differences across endings in nonmusicians suggests that they were able to detect the tonal-syntactic irregularity, but not to encode dissonance.

Musicians' sensitivity to dissonant violations could derive from their enhanced auditory ability (bottom-up sensory processes of perceptual organization) or from their explicit musical knowledge (top-down cognitive processes; Regnault, Bigand, & Besson, 2001). Koelsch et al. (2000, 2002) showed that timbre-deviant instruments elicited a MMN that did not differ between musicians and nonmusicians, while the ERAN for Neapolitans was bigger for musicians. This suggests that the ERAN is not due to enhanced general auditory sensitivity but rather to more elaborated music-syntactic processing. However, while Neapolitans and clusters are similar in tonal and syntactic terms, clusters are dissonant and nonharmonic. Thus, the exclusive application of explicit music-syntactic knowledge may not be enough to explain a bigger ERAN in musicians. Musicians possess an early representation of consonance/dissonance (Itoh, Suwazono, & Nakada, 2010) and a good perceptual differentiation of dissonant dyads and tetrads that is reflected in their ERPs (e.g., Bidelman, 2013; Crespo-Bojorque et al., 2018; McDermott, Lehr, & Oxenham, 2010; Minati et al., 2009; Schön, Regnault, Ystad, & Besson, 2005). Dissonance is hard to distinguish (Bidelman & Krishnan, 2009; McDermott & Oxenham, 2008); hence, while musicians could detect dissonance due to their enhanced auditory ability, nonmusicians could have difficulties detecting such violations. However, the detection of clusters due to enhanced auditory abilities in musicians could trigger a MMN response. To disentangle the MMN from the rule-based ERAN, dissonant clusters could be tested at different positions of the sequence. In fact, it has been observed in nonmusicians that clusters at the fifth position elicit a bigger ERAN than at the third position (Koelsch et al., 2000, 2002). Meanwhile, the frequency-MMN does not change across positions (Koeslch et al., 2001). Such results suggest that context buildup influences the music rule-based ERAN. Interestingly, similar experiments have not been done with musicians. We would expect musicians to show an increasing amplitude of ERAN toward the end of the sequence while being always larger than the ERAN in nonmusicians in both third and fifth position. Such possibility opens the door for future studies to confirm if after musical training the enhanced processing of dissonance is music rule based.

4.2 | Effects of attention

Our study did not require to explicitly pay attention to musical violations. However, we observed that both irregular endings elicited a significant right-frontal positivity around 350 ms in musicians that is reminiscent of the P3a. This component has been proposed to reflect attention allocation after syntactically unexpected endings in all listeners (Regnault et al., 2001), although it can be larger and with an earlier onset in musicians than in nonmusicians (Besson & Faïta, 1995). The ERAN represents sound violation, and it is known that unexpected sounds attract more attention. Thus, as P3a is related to attention, the presence of a P3a after irregular endings supports the presence of the ERAN. Moreover, the P3a indicates that sensitivity to musical violations emerges without the need of explicit attention. As we observed it only in musicians, it suggested an enhanced sensitivity to irregular endings. Importantly, Koeslch et al. (2000) found that, in nonmusicians, Neapolitans elicit a frontal P3a around 380 ms, while dissonant clusters elicit an N2b-P3a-P3b complex. The right-frontal N2b and frontal P3a would reflect attentional processes, while the parietal P3b decisional processes (participants were tempted to respond to clusters). Therefore, we must consider the possibility of a presence of a N2b-P3a-P3b complex in our results. Considering that the main task in our study was relatively easy, participants could have been performing a dual task by attending the deviant instruments and consciously evaluating the musical irregularities. Notably, such conscious evaluation of dissonant endings could have allowed musicians to extract more information (like patterns of beating and relations with the other notes of the chord). Therefore, they could have characterized clusters as stronger violations of context, thus eliciting a stronger ERAN. Meanwhile, nonmusicians would have responded to clusters only in terms of "not matching the context" (Koelsch et al., 1999), thus eliciting a similar ERAN between clusters and Neapolitans.

As musicians have an explicit knowledge about music, they may tend to focus more on the chord sequences (even not intentionally), which could trigger a selective enhancement of the N1. In our study, however, both musicians and nonmusicians showed a bigger N1 after dissonant endings. Therefore, our results do not suggest that musicians attended musical irregularities more than nonmusicians. To the best of our knowledge, previous studies on this issue have not explicitly controlled for possible attentional differences across groups. The enhancement of early sensory components in musicians due to attentional modulation is a topic of ongoing debate. Studies have shown that intentionally directing the attention toward the stimuli does not affect the N1 (Baumann, Meyer, & Jäncke, 2008), rather, that an increased N1 might be due to a general enhancement in the encoding of some acoustic features of complex stimuli (Kaganovich et al., 2013) and not to differences in attention allocation. Moreover, although our main task did not recollect information about the ERPs of deviant instruments, Koelsch et al. (2000, 2002) already showed that they elicit a timbre-sensitive MMN and a N2b-P3a-P3b complex. N2b-P3a-P3b is known to reflect attentional and decisional processes, and it was equal for musicians and nonmusicians. Thus, previous research, together with the fact that we did not find a significant difference in N1 between groups, suggests that the reported differences in the ERAN between musicians and nonmusicians are not confounded by attention.

4.3 | Conclusion

The present study suggests that violations of harmonic sequences can be detected irrespective of musical training while providing further evidence for a processing enhancement in musicians. The stronger reactions (taken as an ERAN) that we observed in musicians may represent the ability of musicians to automatically process the degree of sensory dissonance of chords and use it as a tool to establish the degree of fitting with the previous context. We suggest that the specific enhancement of the ERAN by strong as compared to mild violations reflects that musicians are able to encode unexpected dissonance as a strong violation of musical context. Future studies will disentangle if this effect emerges from enhanced auditory capacities or from the application of explicit musical knowledge. Our study also shows that, despite the fact that irregular endings were task irrelevant and not directly attended, they elicited a P3a in musicians. Therefore, irregular musical events can be salient enough to attract the attention of the listener.

ACKNOWLEDGMENTS

This research was supported by the European Research Council (ERC) Starting Grant agreement n.312519 and by a grant from the BIAL Foundation (reference 13/18). We would like to thank Sofia Russo for her help in the experimental sessions of the study.

ORCID

Carlota Pagès-Portabella https://orcid.org/0000-0001-5151-8844

REFERENCES

Baumann, S., Meyer, M., & Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *Journal of Cognitive Neuroscience*, *20*, 2238–2249. https://doi.org/10.1162/jocn.2008.20157

Berent, I., & Perfetti, C. A. (1993). An on-line method in studying music parsing. *Cognition*, *46*(3), 203–222. https://doi.org/10.1016/0010-0277(93)90010-S

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with non-musicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1278–1296. https://doi.org/10.1037/0096-1523.21.6.1278

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*(4), 485–518. https://doi.org/10.1016/0010-0285(84)90018-5

Bharucha, J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, *13*(1), 63–102. https://doi.org/10.1016/0010-0277(83)90003-3

Bidelman, G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology*, *4*, 1–13. https://doi.org/10.3389/fpsyg.2013.00264

Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, *29*(42), 13165–13171. https://doi.org/10.1523/JNEUROSCI.3900-09.2009

Bowling, D., Purves, D., & Gill, K. Z. (2017). Vocal similarity predicts the relative attraction of musical chords. *PNAS*, *115*(1), 216–221. https://doi.org/10.1073/pnas.1713206115

Brattico, E., Jacobsen, T., De Baene, W., Glerean, E., & Tervaniemi, M. (2010). Cognitive vs. affective listening modes and judgments of music—An ERP study. *Biological Psychology*, *85*(3), 393–409. https://doi.org/10.1016/j.biopsycho.2010.08.014

Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, *117*, 188–198. https://doi.org/10.1016/j.neuropsychologia.2018.06.005

Escera, C., & Corral, M. J. (2007). Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology*, *21*(3–4), 251–264. https://doi.org/10.1027/0269-8803.21.34.251

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., … Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, *19*, 573–576. https://doi.org/10.1016/j.cub.2009.02

Garza-Villarreal, E. A., Brattico, E., Leino, S., Østergaard, L., & Vuust, P. (2011). Distinct neural responses to chord violations: A multiple source analysis study. *Brain Research*, *1389*, 103–114. https://doi.org/10.1016/j.brainres.2011.02.089

Itoh, K., Suwazono, S., & Nakada, T. (2010). Central auditory processing of noncontextual consonance in music: An evoked potential study. *The Journal of the Acoustical Society of America*, *128*(6), 3781–3787. https://doi.org/10.1121/1.3500685

Kaganovich, N., Kim, J., Herring, C., Schumaker, J., Macpherson, M., & Weber-Fox, C. (2013). Musicians show general enhancement of complex sound encoding and better inhibition of irrelevant auditory change in music: An ERP study. *European Journal of Neuroscience*, *37*, 1295–1307. https://doi.org/10.1111/ejn.12110

Kalashnikova, M., Varghese, P., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, *8*, 13745. https://doi.org/10.1038/s41598-018-32150-6

Koelsch, S. (2000). Brain and music: A contribution to the investigation of central auditory processing with a new electrophysiological approach. *MPI Series in Cognitive Neuroscience*, *11*, 1–247.

Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, *46*(1), 179–190. https://doi.org/10.1111/j.1469-8986.2008.00752.x

Koelsch, S. (2011). Toward a neural basis of music perception – a review and updated model. *Frontiers in Psychology*, *2*(110), 1–20. https://doi.org/10.3389/fpsyg.2011.00110

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541. https://doi.org/10.1162/089892900562183

Koelsch, S., Gunter, T., Schröger, E., Tervaniemi, M., Sammler, D., & Friedrici, A. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*, *12*(7), 1385–1389. https://doi.org/10.1097/00001756-200105250-00019

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*(3), 476–490. https://doi.org/10.1111/j.1469-8986.2007.00517.x

Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport*, *14*(5), 709–713. https://doi.org/10.1097/00001756-200304150-00010

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*(5), 657–663. https://doi.org/10.1017/S0048577202010508

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Cognitive NeuroScience*, *10*(6), 1309–1313. https://doi.org/10.1097/00001756-199904260-00029

Kutas, M., & Dale, A. M.. (1997). Electrical and magnetic readings of mental functions. *Cognitive Neuroscience*, *53*(4), 197–242.

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, *1142*(1), 169–177. https://doi.org/10.1016/j.brainres.2007.01.049

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research*, *25*(3), 678–687. https://doi.org/10.1016/j.cogbrainres.2005.08.019

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *162*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*(11), 1035–1041. https://doi.org/10.1016/j.cub.2010.04.019

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch and the auditory system. *Current opinion in Neurobiology*, *18*(4), 452–463. https://doi.org/10.1016/j.conb.2008.09.005

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., … Bruzzone, M. G. (2009). Functional MRI/Event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *NeuroReport*, *20*(1), 87–92. https://doi.org/10.1097/WNR.0b013e32831af235

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Proverbio, A. M., Orlandi, A., & Pisanu, F. (2016). Brain processing of consonance/dissonance in musicians and controls: A hemispheric asymmetry revisited. *European Journal of Neuroscience*, *44*(6), 2340–2356. https://doi.org/10.1111/ejn.13330

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255. https://doi.org/10.1162/089892901564298

Saarinen, J., Paavilainen, P., Schöger, E., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport*, *3*, 1149–1151. https://doi.org/10.1097/00001756-199212000-00030

Schön, D., Regnault, P., Ystad, S., & Besson, M. (2005). Sensory consonance. *Music Perception*, *23*(2), 105–118. https://doi.org/10.1525/mp.2005.23.2.105

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2008). Tonal centers and expectancy: Facilitation or inhibition of chords at the top of the harmonic hierarchy? *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 1031–1043. https://doi.org/10.1037/0096-1523.34.4.1031

Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., & Tervaniemi, M. (2011). The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. *Neuroscience Letters*, *487*(3), 406–10. https://doi.org/10.1016/j.neulet.2010.10.066

Virtala, P., & Tervaniemi, M. (2017). Neurocognition of major-minor and consonance-dissonance. *Music Perception*, *34*(4), 387–404. https://doi.org/10.1525/mp.2017.34.4.387

How to cite this article: Pagès-Portabella C, Toro JM. Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians. *Psychophysiology*. 2019;00:e13476. https://doi.org/10.1111/psyp.13476

3.2 Manuscript 2

Pagès-Portabella C, Bertolo M, Toro JM. Neural correlates of acoustic dissonance in music: The role of musicianship, schematic and veridical expectations. PLoS ONE. 2021 Dec. 16 (12). DOI: 10.1371/JOURNAL.PONE.0260728

PLOS ONE

RESEARCH ARTICLE

Neural correlates of acoustic dissonance in music: The role of musicianship, schematic and veridical expectations

Carlota Pagès-Portabella[1]*, **Mila Bertolo**[1,2], **Juan M. Toro**[1,3]

1 Center for Brain & Cognition, Universitat Pompeu Fabra, Barcelona, Catalunya, Spain, **2** Department of Psychology, Harvard University, Cambridge, Massachusetts, United States of America, **3** Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalunya, Spain

* carlota.pages@upf.edu

OPEN ACCESS

Citation: Pagès-Portabella C, Bertolo M, Toro JM (2021) Neural correlates of acoustic dissonance in music: The role of musicianship, schematic and veridical expectations. PLoS ONE 16(12): e0260728. https://doi.org/10.1371/journal.pone.0260728

Editor: Alice Mado Proverbio, University of Milano–Bicocca: Universita degli Studi di Milano-Bicocca, ITALY

Received: August 11, 2021

Accepted: November 15, 2021

Published: December 1, 2021

Data Availability Statement: All relevant data are within the manuscript and its Supporting information files.

Funding: This research was supported by a grant from the Bial Foundation (reference 13/18, https://www.bial.com/pt/fundacao-bial/) to J.M.T and an F. I. grant from Agència de Gestió d'Ajuts Universitaris i de Recerca (ref. BDNS 362582, FI_B00768, https://agaur.gencat.cat/ca/inici) to C. P.P. The funders had no role in study design, data

Abstract

In western music, harmonic expectations can be fulfilled or broken by unexpected chords. Musical irregularities in the absence of auditory deviance elicit well-studied neural responses (e.g. ERAN, P3, N5). These responses are sensitive to schematic expectations (induced by syntactic rules of chord succession) and veridical expectations about predictability (induced by experimental regularities). However, the cognitive and sensory contributions to these responses and their plasticity as a result of musical training remains under debate. In the present study, we explored whether the neural processing of pure acoustic violations is affected by schematic and veridical expectations. Moreover, we investigated whether these two factors interact with long-term musical training. In Experiment 1, we registered the ERPs elicited by dissonant clusters placed either at the middle or the ending position of chord cadences. In Experiment 2, we presented to the listeners with a high proportion of cadences ending in a dissonant chord. In both experiments, we compared the ERPs of musicians and non-musicians. Dissonant clusters elicited distinctive neural responses (an early negativity, the P3 and the N5). While the EN was not affected by syntactic rules, the P3a and P3b were larger for dissonant closures than for middle dissonant chords. Interestingly, these components were larger in musicians than in non-musicians, while the N5 was the opposite. Finally, the predictability of dissonant closures in our experiment did not modulate any of the ERPs. Our study suggests that, at early time windows, dissonance is processed based on acoustic deviance independently of syntactic rules. However, at longer latencies, listeners may be able to engage integration mechanisms and further processes of attentional and structural analysis dependent on musical hierarchies, which are enhanced in musicians.

Introduction

The hierarchical organization of listeners' perceptions of tensions and relaxations through time forms musical syntax [1]. Musical regularities (e.g. how musical events are combined into

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

sequences) creates predictions about upcoming musical events or *schematic expectations* [2], by which unstable chords should lead to more stable ones. Our brain also makes predictions based on the probability of appearance of unexpected musical events, which generate *veridical expectations* (e.g., through the repetition of violations; [3,4]). One of the key questions of music cognition research is how the creation and violation of these musical expectations are encoded in the brain. In other words, how expectation and prediction interact with occurrence, as music perception unfolds over time. To address these issues, a large body of empirical research [5–8] studied the set of neural responses that are elicited by unexpected musical events. However, the music-specificity and the contribution of cognitive/sensory processes behind these responses are still under debate [9–11]. In the present study, we contribute to this line of research by exploring whether the processing of tonal expectations broken by strong dissonance is modulated by schematic and veridical expectations and musical expertise.

In musical terms, irregular chords that violate schematic expectations (for instance, by disrupting the expectation for resolution after a dominant chord) elicit an *early right anterior negativity* (or ERAN, which is a type of MMN), the P3 and the N5 [5,9,12]. Traditionally, the ERAN is suggested to represent the cognitive processing of a violation of music-syntactic rules. In more general terms, the brain makes predictions about upcoming events and aims to reduce the uncertainty of these predictions by comparing prior models to sensory inputs (which is known as Bayesian surprise; [13–15]). Unpredictable or surprising events are informative because they contribute to reduce the uncertainty of the brain's predictions and, therefore, elicit neural responses, such as the MMN [16,17]). Interestingly, when irregularities become predictably surprising through repetition, the subsequent neural responses (such as the ERAN) diminish, as listeners extract veridical expectations about their appearance and become able to anticipate them [5,18]. It thus seems that, in music, schematic expectations are modulated to some extent by veridical expectations [3,19]. Likely, the reduction of the neural responses after exposure to unexpected events is related to the sensory cortex having evolved to stop responding to those events that match its predictions and are, therefore, uninformative [20].

Syntactic processes in music are intimately entwined with sensory-driven processes (in contrast to language). Western tonal syntax is deeply rooted in the acoustic properties of sounds, their psychoacoustic effects and their storage in auditory memory. For example, the syntactically most important events have also strong overlap in harmonic spectra, which is an acoustic feature. Thus, tonal hierarchies and their subsequent neural responses might be (at least partly) an emergent property of auditory short-term memory (ASTM), which is the overlap of the auditory image of any tone or chord with the auditory image created by the previous events accumulated in auditory memory [11,21].

A critical issue in music cognition is to determine the respective weights that acoustic information stored in ASTM and learned syntactic representations have in musical syntax processing [21]. To address that issue, most research focused on minimizing acoustic deviance and introducing only musical deviance [9,10]. Some studies explored chords that introduce pure acoustical deviance independently of their relationship with the musical context. For example, auditory deviants like mistuned chords [7] and frequency violations [22]. Koelsch and colleagues [5] compared the ERPs triggered by dissonant and syntactic violations of tonal context. In that study, highly dissonant semitone clusters introduced acoustic deviance without being syntactically correct. At the same time, Neapolitan chords introduced contextual acoustic dissimilarity (with their out-of-key tones) and were rendered either syntactically correct or incorrect depending on their position within the cadence. The critical finding of that study (that was later supported by other studies [7,9] was that the amplitude of the ERAN was larger for Neapolitans presented at the end of the cadence than for Neapolitans at the third position. At the third position, Neapolitan chords substituted a subdominant (*syntactically correct)* while at the end of the cadence sabotaged the resolution to the tonic (*syntactically incorrect*). The authors thus suggested that the ERAN reflects syntactic-like processes because irregular chords violate sound expectancy to a higher degree at the end of a cadence [5,7]. However, the authors also found a difference between positions for dissonant clusters, that were not syntactically correct at any of the positions, as if it was difficult for the brain to distinguish syntactic from acoustic violations. Moreover, more recent research demonstrated that ASTM can also account for the responses reported in most electrophysiological studies [11]. Thus, there is a current debate around the sensory and cognitive mechanisms behind music processing.

In the present study, we advance this line of research by using dissonant chords that introduced a strong acoustical violation but that, in terms of ASTM, [11,21], were undistinguishable between positions. Dissonant chords include intervals with complex frequency ratios, roughness, which creates the overlapping of amplitude envelopes in the basilar membrane [23] and a lack of harmonicity [24]. The convention of consonance/dissonance is an always-changing continuum [25]. For instance, the perception of seventh chords changed from dissonant in the 18th century to commonplace in classical music [26] or mildly dissonant and even preferred in jazz tradition [27,28]. In the present study we thus use chords deliberately containing highly dissonant intervals (such as the minor second and the tritone) that are very rare in most western genres [25,29] and do not act as culturally specific chord prototypes [30]. Thus, their interpretation should depend less on the musical experience of the listener while they represent a good model for investigating the role of prolonged experience on the ability to distinguish sensory from syntactic violations.

Musical training enhances the neural responses triggered by most musical violations. Western musicians show a larger responses to irregular chords because their expectations for tonal resolutions are more finely tuned [12]. Musicians are also more likely exposed to intentionally dissonant or mistuned chords, which are more frequent in genres like avant-garde or free jazz and contemporary music [27]. Thus, they display a flexible auditory system with a greater ability to quickly categorize dissonant chords [28,30–33] and more rapid perceptual learning [34]. The comparison of highly trained musicians with naïve listeners might provide valuable information about the discrimination between syntactic and acoustic violations.

Thus, in the present study, we investigate whether the neural responses to dissonance change as a function of the position that the chords occupy in a sequence (Experiment 1). Second, we study whether these responses are modulated by the predictability of dissonant endings (Experiment 2). To account for a possible role of training in how these conditions modulate the neural responses of the listener, in both experiments we compared highly trained musicians against naïve listeners.

Experiment 1

Introduction

In previous studies, syntactic irregularities (in the absence of acoustic deviance) were found to elicit stronger ERAN as closure than at middle positions of chord cadences [9,10], because at the end of the cadence they violate sound expectancy to a higher degree. The first chords of a chord sequence do not clearly establish a key (e.g., a C followed by Am can lead to 6 different possible keys). But the key is unequivocally established after four chords. Moreover, unstable chords (as the dominant at the penultimate position) increase the demand of resolution [35]. In terms of Bayesian surprise, precision estimates accumulate over time based on the probability distribution of each event [36]. Therefore, the auditory system collects more evidence in favour of in-key chords after 4 chords than after 2 chords.

In our first experiment, we investigated the neural responses elicited by acoustic violations in a tonal context, introduced by dissonant chords at different positions of a chord cadence. In contrast with previous research, we studied whether these responses are modulated by musical expertise. We presented trained musicians and naïve listeners with dissonant clusters placed at the third (middle) and at the fifth (ending) position of a chord cadence. Besides being highly dissonant, they both introduced an out-of-key tone but were indistinguishable between positions at the auditory level. Therefore, if we observe stronger responses at the ending position, they would be based on stronger tonality establishment and expectations of resolution (which, in turn, might be based on the accumulation of sensory evidence in favour of the tonic). On the contrary, if the responses to dissonance were not different between positions, it would suggest that dissonance is processed as an acoustic anomaly not affected by hierarchical relationships. We hypothesized that musicians might show a different pattern of neural responses than non-musicians.

Methods

Participants

24 volunteers participated in the experiment, most of them were undergrade students. 12 were musicians and 12, non-musicians [8,31,37]. Musicians (3 identified as cisgender women, mean age 23.66 ± 6.53), finished or were studying Advanced Studies in Western music, starting at 6.9 (± 2.2) years of age and had been musically active for 14.6 (±2.8) years on average. They all specialized in an instrument (mostly piano), 5 out of 12 studied jazz and modern music and all were musically active at the time of the experiment. Non-musicians (8 of which identified as cisgender women, mean age 24.3 ± 5.4) never received formal musical training besides the compulsory program at school. All participants were right-handed and reported normal (or corrected to normal) vision and no diagnosed hearing problems. The study protocol was approved by the Etical Committee for Drugs Research Parc de Salut Mar (Comité de Ética de la Investigación con Medicamientos Parc de Salut Mar, CEIm, under the reference number 2018/7888/I). We obtained written informed consent from the participants who received a monetary compensation for their participation of 5€ each 30min. The data was analyzed anonymously.

Stimuli

We created different 5-chord sequences. Each chord was composed of four notes (with respect to the tonic: fundamental, third, fifth and seventh). The chord at the first position in each sequence was always the tonic seventh chord. The second chord could be the submediant (VIm7) or the subdominant (IVMaj7). The third chord could be the supertonic (IIm7), the subdominant (IVMaj7) or a dissonant cluster. The fourth chord was always the dominant seventh (V7) to create the maximum expectation for resolution. These combinations established three types of context sequences: I-VIm-IV-V, I-IV-IIm-V and I-IV-IIm-V. The sequences could end in either the tonic or a dissonant cluster (Fig 1). Dissonant chords kept the fundamental and the third of the tonic but included a cluster of semitones (with respect to the tonic: the third, the fourth and the augmented fourth). Thus, they were a dissonant version of the tonic chord, that is syntactically correct at both the third and the fifth position. Preliminary analyses with the ASTM model proposed by Leman [11,21] were performed to test the goodness of fit of dissonant clusters with the musical context. The ASTM model calculates the acoustic congruency of a chord with the auditory sensory memory traces established by the previous chords, whose auditory information decays but is kept in the echoic memory for a certain time. Thus, it calculates the correlation between the pitch image of a chord and the

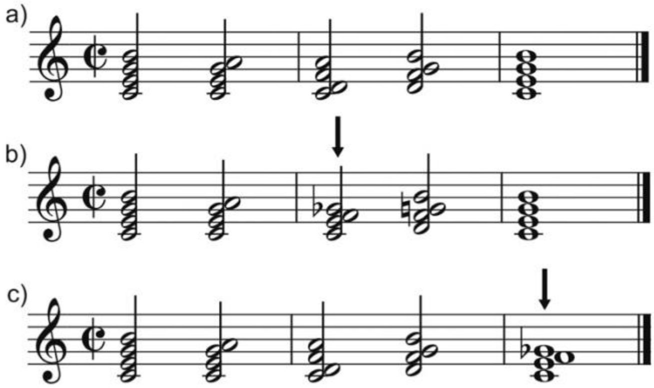


Fig 1. Examples of chord sequences in root position. Dissonant clusters are indicated by black arrows. a) Cadences consisting of in-key chords that resolve to the tonic. b) Cadences containing a dissonant cluster in the third position. c) Cadences containing a dissonant cluster in the fifth position.

https://doi.org/10.1371/journal.pone.0260728.g001

pitch image of the previous chords stored in the echoic memory [10]. According to the model, clusters in our study were acoustically more similar to the tonic than a Neapolitan violation, despite being highly dissonant, because they shared more tones with the tonic. More importantly, they were acoustically undistinguishable between positions. Therefore, any differences in the ERPs between positions could not be explained by differences in auditory deviance.

In order to add more variability to the stimuli, the three basic context sequences were modified so that the first and last chord were either in root position, first inversion or second inversion. The lead voice always moved by conjunct melodic motion. The resulting sequences were moreover transposed to the 12 different keys, resulting in 108 context sequences (without taking into account the sequences that had the dissonant ending at the third position). We manipulated the proportion of appearance of each type of target chord. Fifty percent of the total sequences (216) ended as expected (in the tonic). 108 sequences presented a dissonant cluster at the third position, and another 108 sequences presented a dissonant cluster at the fifth position (Fig 1). Thus, both kinds of sequences containing dissonant clusters appeared with a probability of 25% each. There was a total of 432 sequences played. Within each sequence, chords 1 to 4 lasted for 600ms each [5,12] and chord 5 lasted for 1200ms. Each sequence lasted for 3600ms. All sequences were recorded with Cubase LE AI Elements 9.5 using the piano instrument. In order to create a secondary task (see procedure below), 10% of the sequences (four randomly selected sequences of each tonality) included a chord in the second or third position played by a deviant instrument other than a piano (violin, guitar or vibraphone). When the sequence had a dissonant chord at the third position, it was never played by a deviant instrument. However, deviant instruments at the second position could be followed by a dissonant chord at the third position. The 36 sequences belonging to the same key were presented together in a block. Thus, there were 12 blocks (one per key). The order of presentation of blocks was randomized across participants.

Procedure

The experiment was run in an acoustically and electrically shielded room. Participants sat comfortably in an armchair and listened to the sequences, presented via headphones using

Matlab. Participants were told not to blink, move their body or their eyes while listening. They were instructed to look at a fixation cross at the centre of the screen. We presented the sequences with a silent period of 500ms between them, during which three stars appeared on the screen. That break between sequences performed a double function: first, it allowed participants to blink between sequences, which prevented having artefacts overwritten in the signal of interest. Second, a pause prevented that the responses to the last and longer chord overlapped with those elicited by the first chord of the next sequence. In order to keep the participants engaged while listening to the sequences, we used a secondary task. The participants were asked to detect and count how many times they heard an instrument different than a piano. They were not informed about the presence of dissonant ending chords to avoid overlapping attentional and decisional effects in the responses of interest. Twice in each block (every 18 sequences) the participants were asked by a text shown on the screen to provide a numerical answer with the keypad. At the end of each block there was a 10s countdown as a break. There was also a longer break in the middle of the experiment so participants could move, after which they had to press the SPACE bar to continue with the experiment. The duration of the entire experimental session was approximately 40 min.

Recording and analyses

EEG measurements were recorded with a 32 channels actiCAP Slim (Brain Products) with Ag/AgCl electrodes. 28 scalp locations following the 10–20 system were recorded (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, TP9, CP5, CP1, CP2, CP6, TP10, P7, P3, Pz, P4, P8, Oz). Two electrodes were placed in the left and right mastoid (MSDL, MSDR) and two more in the outer side (HEOG) and below (VEOG) the right eye to monitor the ocular movements and blinking. Measurements were referenced online to the tip of the nose, and offline to the average of the mastoids. The sampling rate was 500Hz and impedances were kept below 10kΩ. Data was recorded using BrainVision Recorder and the triggers sent simultaneously with Matlab 2019b.

Data was pre-processed offline with Fieldtrip Toolbox: filtered from 0.4 to 40Hz and bad electrodes corrected via neighbour interpolation. Independent Component Analysis (ICA) was applied to the whole trial epochs (3.6s) to identify the variance caused by eye movements, heart and muscular movement. The components reflecting the artefacts were manually removed individually on each participant. On average, 2% of trials were rejected from further analysis because of residual excessive noise. After artefact rejection, the data was divided in epochs of 600ms for chords at the third position in Experiment 1 and of 1200ms for chords at the ending position (nevertheless, the responses of interest were always within the first 600ms). We applied a baseline correction from 200ms to the onset of the chord performed in two steps: first trial by trial and, after calculating average ERPs, a condition-specific baseline for all participants. Data for statistical analyses were selected from the time windows in which the responses of interest are usually found in the literature: 150-250ms for the early negativity, 300-450ms for the frontal positivity and 500-600ms for the late negativity [5,12,22]. We later confirmed these time windows by visually inspecting the grand-average difference waves (dissonant minus in-key chords). That visual inspection also revealed a parietal subcomponent of a positivity that was further analysed within the time window 350-550ms. Note that we were not interested in the effect of Chord itself, but rather in the interactions. To avoid priming effects on dissonant ending chords, we excluded from the analyses regarding the effect of dissonant chords those sequences containing deviant instruments. We moreover verified that the statistical results were not affected by the exclusion of these trials.

We calculated the ERPs by averaging across the trials belonging to each condition (in-key 3rd position: 324 trials, in-key 5th position: 216, dissonant 3rd position: 108, dissonant 5th position: 108). Then, for each time window of interest, we calculated the mean amplitudes for each participant in each condition. After, we averaged these mean amplitudes across the electrodes of our regions of interest (ROIs) (right-frontal electrodes [F4, F8, FC2, FC4], left-frontal electrodes [F3, F7, FC1, FC3]) based on previous literature regarding the responses of interest [5,12], which entered the statistical analysis. We performed mixed-design repeated measures ANOVA with the within-subject factors Region of Interest (ROI; left-anterior, right-anterior), Chord Type (in-key, dissonant), Position of the dissonant chord within the sequence (3rd, which corresponds to a lower expectation violation and 5th, a high expectation violation) and the between-subjects factor Group (musicians, non-musicians). To analyse the parietal response, parallel ANOVAs were run over the parietal ROIs (right-parietal [CP2, CP6, P4, P8], left-parietal [CP1, CP5, P3, P6]). We also analysed the responses to deviant instruments, by comparing the amplitudes of the windows 150–200 and 250-400ms with the factors Instrument (piano vs deviant), ROI and Group. In all the analyses, to correct for multiple comparisons (in all the tests over different time windows), we applied a Bonferroni correction and when sphericity was violated, results with Greenhouse-Geiser correction are reported.

Results

Dissonant chords

We observed a negative response for dissonant chords in comparison to in-key chords in both the third and the fifth positions around 200ms with a right-frontal distribution (Fig 2). We also observed a frontal positivity peaking around 350ms and a late parietal positivity arising around 400ms (Fig 2). Finally, we observed a bilateral late negativity arising around 500ms.

For the 150-250ms time window, the ANOVA yielded a main effect of Chord ($F(1, 22) = 42.46$, $p < .001$, $\eta 2 = .66$), ROI ($F(1, 22) = 6.50$, $p = .018$, $\eta 2 = .23$) and Position ($F(1, 22) = 8.90$, $p = .007$, $\eta 2 = .29$). It also showed interactions between Chord and Group ($F(1, 22) = 5.32$, $p = .031$, $\eta 2 = .20$), Chord and ROI ($F(1, 22) = 7.06$, $p = .014$, $\eta 2 = .24$) and Position and ROI ($F(1, 22) = 8.86$, $p = .007$, $\eta 2 = .29$). The effect of Chord suggested that dissonant chords elicit a negativity in comparison to in-key chords in both groups (mean amplitude difference between chord types ± SEM: -1,176 ± 0.18). The amplitude of the negativity appeared larger in musicians (-1.59 ± 0.26) than in non-musicians (-0.76 ± 0.26) and in the right ROI (-1.30 ± 0.21) than in the left ROI (-1.05 ± 0.21), which could suggest a right-lateralization of the response. Importantly, the lack of interaction between Chord and Position indicated that the magnitude of the effect did not depend on whether the dissonant chord was presented in the middle of the sequence (third position) or at its end (fifth position; Table 1).

For the 300-450ms time window, the ANOVA yielded a main effect of Chord ($F(1, 22) = 6.24$, $p = .021$, $\eta 2 = .22$), Position ($F(1, 22) = 23.74$, $p < .001$, $\eta 2 = .51$) and Group ($F(1, 22) = 20.82$, $p < .001$, $\eta 2 = .49$). Also, it revealed interactions between Chord and Position ($F(1, 22) = 12.81$, $p = .002$, $\eta 2 = .37$), Chord and Group ($F(1, 22) = 6.01$, $p = .023$, $\eta 2 = .21$), Chord and ROI ($F(1, 22) = 8.02$ $p = .010$, $\eta 2 = .27$), ROI and Group ($F(1, 22) = 18.40$, $p < .001$, $\eta 2 = .46$), Position and Group ($F(1, 22) = 8.84$, $p = .007$, $\eta 2 = .29$), ROI, Chord and Group ($F(1, 22) = 5.11$, $p = .034$, $\eta 2 = .19$) and between Chord, Position, ROI and Group ($F(1, 22) = 5.10$ $p = .034$, $\eta 2 = .19$). Together, these effects suggested that dissonant chords elicit a positivity in comparison to in-key chords in this time window (mean difference 0.45 ± 0.18). Post hoc tests for the interactions revealed that this effect reached statistical significance for chords at the 5th position in musicians (0.89 ± 0.26, $p < .001$, 95% CI = 1.11, 2.65) but not in non-musicians (0.01 ± 0.26) and was larger in the right ROI (1.89 ± 0.37 vs 1.25 ± 0.41).

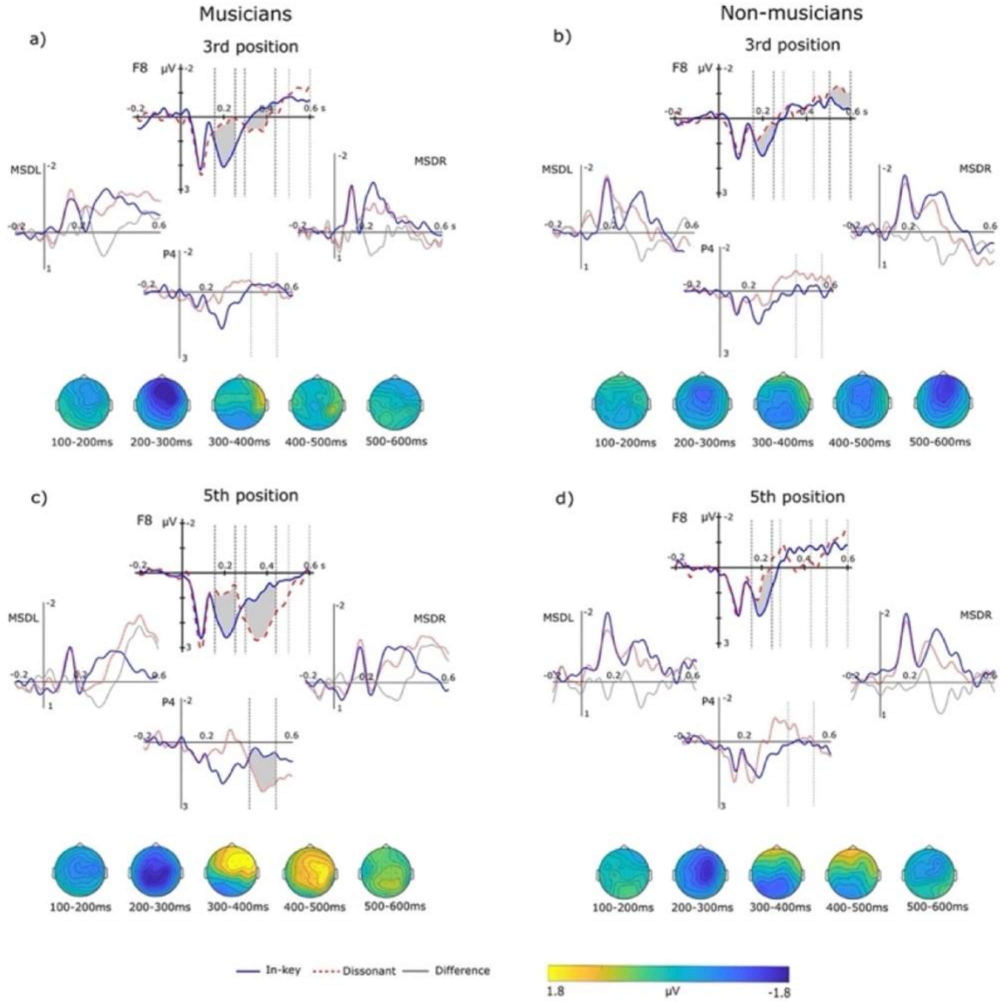


Fig 2. Experiment 1: Event-related potentials of musicians and non-musicians. Graphs include grand-average ERPs and the topographical distribution of the difference waves (dissonant minus in-key) of musicians (a, c) and non-musicians (b, d) at the third (a, b) and fifth position (c, d). Dotted vertical lines indicate the limits of the time windows of interest analysed. Shaded areas correspond to significant ERP effects, which are best seen in electrode F8. Left (MSDL) and right mastoidal (MSDR), and parietal (P4) are also shown. Topographical plots depict the grand-average difference wave to better observe the evolution of the topographical distribution of each ERP of interest.

https://doi.org/10.1371/journal.pone.0260728.g002

For the 350-550ms time window, the ANOVA yielded a main effect of Position ($F(1, 22) = 71.45$, $p < .001$, $\eta 2 = .77$), ROI ($F(1, 22) = 9.54$ $p = .005$, $\eta 2 = .30$), Group ($F(1, 22) = 13.64$, $p = .001$, $\eta 2 = .38$) and interactions between Chord and Group ($F(1, 22) = 18.54$, $p < .001$, $\eta 2 = .46$), Position and Group ($F(1, 22) = 22.81$, $p < .001$, $\eta 2 = .51$) and Position and Chord ($F(1, 22) = 6.40$, $p = .019$, $\eta 2 = .23$). Post hoc tests for interactions suggested that dissonant chords

Table 1. Mean amplitudes and SEM for the ERPs of interest. Mean amplitudes reported are the grand-average of the mean amplitude of the difference waves calculated within each time window of interest.

			Mean amplitude (μV) + SEM			
			EN	P3a	P3b	N5
Experiment 1	Musicians	3rd position	-1.65 (0.22)	0.21 (0.22)	0.08(0.22)	-0.17 (0.21)
		5th position	-1.52 (0.39)	1.56 (0.42)	0.96(0.30)	0.42 (0.36)
	Non-musicians	3rd position	-0.71 (0.30)	-0.34 (0.28)	-0.6(0.19)	-1.01 (0.24)
		5th position	-0.83 (0.23)	0.36 (0.33)	-0.27(0.18)	-0.37 (0.18)
Experiment 2	Musicians		-1.20 (0.33)	1.90 (0.56)	0.90(0.28)	0.25 (0.27)
	Non-musicians		-1.34 (0.34)	0.73 (0.33)	0.37(0.35)	-0.98 (0.23)

https://doi.org/10.1371/journal.pone.0260728.t001

elicited a positivity in comparison to in-key chords, which was larger in musicians (0.52 ± 0.16) than in non-musicians (-0.43 ± 0.16), as the effect of Chord only reached significance in the former group (p < .003, 95% CI = 0.20, 0.84). The positivity also appeared larger at the fifth than at the third position (0.35 ± 0.18 vs -0.25 ± 0.15).

For the 500-600ms time window, the ANOVA revealed a main effect of Chord (F(1, 22) = 4.99, p = .036, η2 = .19), Group (F(1, 22) = 5.34, p = .031, η2 = .20) and Position (F(1, 22) = 12.46, p = .002, η2 = .36). Moreover, there were interactions between Chord and Group (F (1,22) = 9.79, p = .005, η2 = .31), Chord and Position (F(1, 22) = 5.52, p = .028, η2 = .20), ROI and Group (F(1, 22) = 8.13, p = .009, η2 = .27) and between Position, ROI and Group (F(1, 22) = 4.58, p = .044, η2 = .17). The effect of Chord indicated that dissonant chords elicit a negativity in comparison to in-key chords (-0.31 ± 0.14). Post hoc tests for interactions suggested that the Chord effect was significant only at the third position (p = .001, 95% CI = -0.92, -0.26) and was larger in non-musicians (-0.74 ± 0.20) than in musicians (-0.12 ± 0.20) as it only reached significance in the former group (p = .001, 95%CI = -1.15, -0.34).

Deviant instruments

The presentation of deviant instruments elicited a central negativity that peaked around 180ms. The negativity was followed by a large frontal positivity peaking around 300ms and a late parietal positivity arising around 450ms (Fig 3).

For the 150-200ms time window, we found an effect of Instrument (F(1, 22) = 28.47, p < .001, η2 = .56, 95% CI = -5.48, -2.41). The magnitude of the negativity did not differ between groups. For the 250-400ms time window we found a main effect of Instrument (F(1, 22) = 109.30, p < .001, η2 = .83, 95% CI = 5.42, 8.10) and of Group (F(1, 22) = 10.88, p = .003, η2 = .33, 95% CI = 0.85, 3.71). As suggested by the lack of interactions with Group, deviant instruments elicited comparable responses between musicians and non-musicians.

In the behavioural task of counting the number of deviant instruments, we calculated the performance of participants as the number of correct counts of deviant instruments over the total of times that they were asked. Musicians had a performance of 92.36% ± 2.80 and non-musicians, of 88.89% ± 2.96. A good performance in both groups of participants indicated that they paid attention to the task. Moreover, the performance was not perfect in either group, indicating that the task was not too easy, not even for musicians.

Discussion

In the present experiment dissonant clusters elicited a right-frontal early negativity (which we will refer to as EN for simplicity) that reached its maximum around 200ms. The amplitude of the EN did not significantly differ between positions, but it was larger in musicians than in

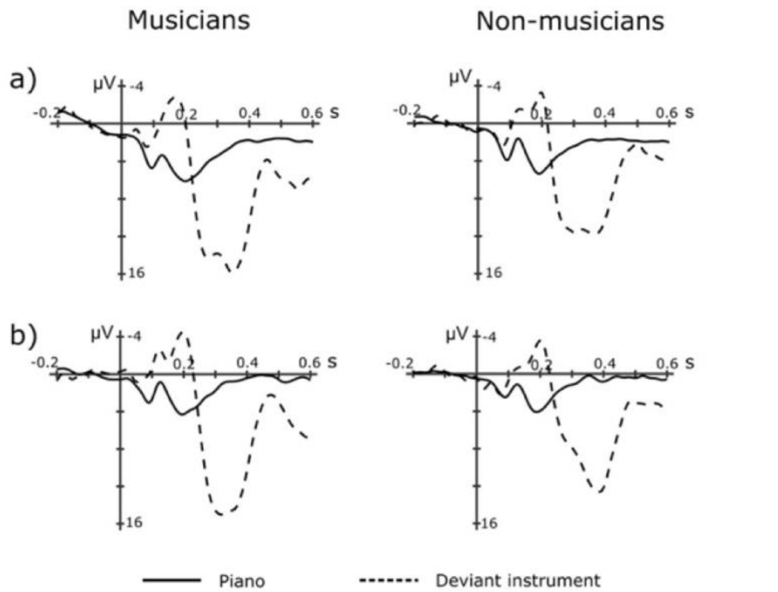


Fig 3. Event-related potentials of deviant instruments. Grand-average ERPs for deviant and standard (piano) instruments for musicians and non-musicians for Experiment 1 (a) and Experiment 2 (b). Potentials at Cz are being depicted.

https://doi.org/10.1371/journal.pone.0260728.g003

non-musicians. Moreover, we found a robust frontal positivity peaking around 300ms (taken as the P3a) that was larger in musicians for ending clusters and was followed by a late parietal positivity (taken as the P3b). Finally, clusters elicited a late bilateral negativity at 500ms (taken as the N5), which was most prominent in non-musicians for middle clusters.

Early negativity

Dissonant clusters elicited an EN in musicians and non-musicians. The amplitude of the EN was statistically undistinguishable between the middle and the ending positions in both groups. Syntactic irregularities well controlled for acoustic deviance have been shown to trigger larger negativities for ending positions [9] that cannot be simply accounted by ASTM [11]. In our experiment, clusters did not differ across positions in terms of ASTM. Because early responses are more related to basic acoustic deviations, the similarity in the EN across positions that we observed may reflect the detection of the acoustic deviance introduced by the distinction between consonant/dissonant (which did not differ across positions) and may be simply interpreted as a MMN [7,38]. However, such early response may not be sensible to higher-level processes related to schematic expectations or the degree of sensory surprise based on the evidence in favour of the tonic accumulated across the cadence. That is a striking result, especially in musicians, who are able to automatically categorize chords and should form stronger expectations.

Importantly, musicians showed a larger EN than non-musicians. Musicians have enhanced discrimination of dissonance at early time windows [30,39,40]. However, the contributions to this enhancement are under current debate. One possibility is that musicians possess more sophisticated overall auditory abilities for any kind of sound. This may provide them with greater sensitivity for the beats in the dissonant chord (as shown in the N2; [41]). However, an enhancement in sensory processing should also be reflected in the initial feature analysis in the auditory cortex, and signalled by the P1 or the N1, which is not the case [30]. A complementary view is that musical training provides listeners with a perceptual expertise that allows them to differentiate culture-specific musical categories, such as non-prototypical chords [30]. Indeed, musicians' MMN is equally sensitive to non-prototypical dissonant chords in a dissonant context than to stereotypical chords in a consonant context [40,42]. Moreover, different types of musical training might separately shape sensory and categorization abilities. For instance, jazz training includes ear learning together with explicit knowledge of complex chord changes and rich harmonies, which can be reflected in a general enhancement of the MMN in comparison to classical or pop musicians [43]. Our results are consistent with both explanations and might emerge from the combination of both the sophistication of auditory abilities and the perceptual expertise that derive from musical training.

P3 effects

Dissonant clusters elicited a frontal positivity around 300ms, which is consistent with the P3a (Fig 2) and a late parietal positivity, interpreted as the P3b. The P3a and P3b are considered subcomponents of the P300, where the P3a (novelty P3 in traditional terms) is frequently linked to the non-intentional allocation of attention resources to a stimulus that is novel or unexpected [44,45] and the P3b (or target P3) is associated with context updating (see below). In terms of sensory surprise, stimuli with a low likelihood given a distribution of expected or learned stimuli are more informative, more surprising [20]. Precisely because unexpected events are informative, they require further evaluation, for which is necessary the attention of the listener. If sensory changes were predictable, further evaluation would not be necessary [46]. Therefore, that task-relevant deviant instruments were surprising and attracted the attention of participants was expected because the participants were informed about their presence and instructed to detect them. But, importantly, participants were not informed about the presence of dissonant chords [5,12]. Therefore, that clusters elicited the P3a could suggest that dissonant chords were surprising, salient or informative enough for listeners and subsequently attract their attention, because we actively seek or attend to sensory cues that are surprising [47]. This is in line with the idea that, for the P3a to be elicited, it requires that the change exceeds a certain threshold to be salient [46].

Importantly, the P3a elicited by dissonant clusters was statistically significant in musicians but not in non-musicians. Musical training is known to enhance P3 effects [2,48]: musicians have better perceptual learning, and stronger and faster involuntary attention switching [49]. Thus, the lack of a significant in the P3a in non-musicians might suggest that the change that clusters introduced in the chord sequences might have not been as surprising or salient for them than for musicians. Furthermore, in musicians, the P3a was significantly larger for ending than for middle clusters (Fig 4, left panel). Thus, strong dissonance at the end of a cadence might be perceived as less expected, more surprising, than in the middle position. In musical terms, ending clusters are more surprising than middle clusters because they substitute the very expected tonic, while at the third position they substitute a subdominant. Similarly, within a framework of Bayesian surprise, ending clusters are more surprising because the sensory system has accumulated more evidence in favour of the tonic by the fifth position.

Ending clusters, but not middle clusters, elicited a later parietal P3b that was significant in musicians but not in non-musicians. The P3b is related to subsequent memory processing arising from the revision of the brain's current model of the musical structure induced by

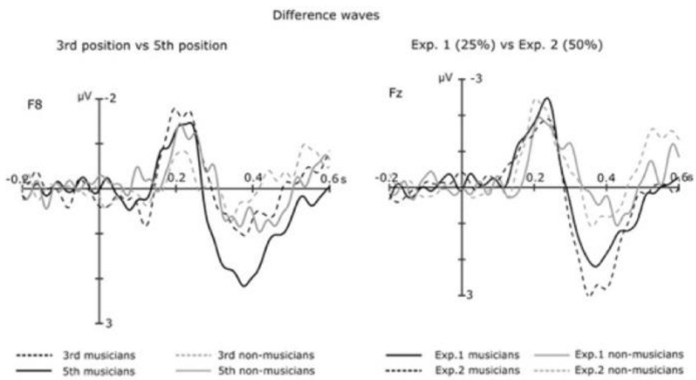


Fig 4. Difference waves the ERPs of musicians and non-musicians. Left panel: Difference wave comparing chords at the 3rd versus the 5th position. Right panel: Difference wave comparing Experiment 1 (where dissonant endings occurred with a probability of 25%) versus Experiment 2 (where dissonant endings occurred with a probability of 50%).

https://doi.org/10.1371/journal.pone.0260728.g004

incoming stimuli [50] or, in other words, context-updating [51]. The P3b is also related to processes of structural analysis in rule-governed sequences [50], which reflects proper conscious perception [45] and ultimately leads to coordination of behavioural responses [5]. Importantly, the amplitude of the P3b reflects the amount of information extracted from the unexpected stimulus [46]. This suggests that ending clusters might have recruited further processing resources than middle clusters because they introduced a larger disruption of the current sensory/musical model. At the same time, the P3b is thought to be elicited only by task-relevant stimuli [46], suggesting that such disruption was surprising enough to overcome task relevancy. The fact that ending clusters elicited a P3b that was significant only in musicians suggests that they might be more efficient at matching the target chord to the memory trace of the context [49] than non-musicians [5]. Indeed, in our study some musicians reported that they were strongly surprised by dissonant chords and were tempted to respond to them.

In sum, in contrast to early responses, the subcomponents of the P3 elicited by dissonant chords suggest that ending clusters became more surprising than middle clusters, especially for experienced listeners. Importantly, acoustic differences in terms of ASTM cannot account for such enhancement of the P3 subcomponents toward the end of the cadence, suggesting that these late responses are sensible enough to register higher-order schematic expectations.

N5 effects

Dissonant clusters elicited a negativity around 500ms consistent with the N5. However, the N5 was statistically significant for middle clusters in non-musicians, but not in musicians. The N5 has been reported when irregular chords are task-irrelevant [12,52]. That we observed the N5 at least in non-musicians suggests that they made an effort in integrating clusters into the previous context [5]. On the other hand, it is unlikely that the lack of a N5 in musicians is due to an ease in the integration of clusters, because clusters are very strong acoustic violations and musicians are very sensitive to dissonance [30,32,53]. A possible explanation is that the positive potentials of the large P3 effects compensated the negative potentials of the N5 [5,9]. This reduction effect would be most salient in musicians, who showed a clear P3a. Similar effects have been observed when participants were asked to respond to irregular chords, which increased the P3 amplitude for target deviants [5,12]. Similarly, the N5 that we observed in non-musicians was larger for middle than for ending clusters, likely because the increasing magnitude of the P3a toward the end of the cadence increasingly masked the N5. The fact that the P3a possibly masked the N5 leaves open the question of whether musicians did not invest neural resources in integrating clusters because they are so unacceptable that are processed independently of music rules (while non-musicians attempted to do so) or whether musicians did but that was not reflected in the N5 due to the overlapping attentional effects.

Thus, in Experiment 1, we investigated whether the processing of dissonance is modulated by the syntactic relations in a chord cadence. We inserted infrequent dissonant chords in either the third or the fifth position of the cadence (each in 25% of the sequences). This design might tap into the interaction between veridical and schematic expectations because the appearance of a dissonant chord in the third position predicts that a dissonant chord will not appear in the ending position. To specifically test whether listeners can extract online higher-order veridical expectations about the appearance of dissonant closures, in experiment 2 we manipulated the predictability of dissonant closures by increasing their frequency of appearance.

Experiment 2

Introduction

The neural correlates responsive to irregularities (such as the MMN/ERAN and P3) have been suggested to be reduced by veridical expectations [5,18,51]. The precision of the brain's predictions is contextual, because they accumulate over time, based on the statistics of the environment (e.g., the stimuli during an experiment). Listeners can accumulate evidence for the relative precision about both the local model (consonant/dissonant) and about the regularity of the pattern of stimuli itself (highly frequent dissonant chords), creating superordinate expectations. After exposure to a very predictable context for enough time, higher-order predictions are held with greater confidence, making them relatively impervious to disconfirmatory sensory evidence (which minimizes the subsequent responses) [36]. In other words, as initially unexpected events become predictable with repeated exposure, neural responses attenuate because they become uninformative for the brain's internal models [20].

In the present experiment, we tested whether an increased predictability of acoustic violations in a tonal context (introduced by dissonant endings) would elicit reduced neural responses. We exposed listeners to musical sequences that, with a high probability (50%), would end on a dissonant cluster instead of the expected tonic. In this setting, the representation of correctly ending sequences would be weaker and clusters should become easier to anticipate, in comparison to Experiment 1, where clusters equally appeared at the middle or ending position. However, given the evidence suggesting that dissonance is psychoacoustically hard to process [53] and disruptive at the auditory level, we hypothesized that the neural responses to dissonant clusters would hardly diminish, even if listeners can easily anticipate their appearance. Alternatively, if the neural responses to frequent dissonant endings are in fact reduced in comparison to infrequent dissonant endings, it would suggest that veridical observations about any acoustical irregularity that appears often enough could prevail over the learned rules of hierarchical expectations, as if the irregularity becomes the new rule. Even music with quartertones sounds unpleasant at first, but after some repetitions, listeners tend to like it more as a result of the "mere exposure" effect [26]. Importantly, long-term musical training might lead to an increased experience with dissonant chords, faster perceptual learning and an ease in creating higher-order predictions for musical stimuli. Thus, we sought to explore the role of musical training in processing highly predictable dissonant closures.

Musicians may render the anticipation of clusters easy enough to (at least partly) compensate the strong acoustic disruption that they introduce.

Methods

Participants

The participation criteria were identical to those in Experiment 1. 12 musicians (5 of which identified as cisgender women, mean age = 23 ± 4.3 years) and 12 non-musicians (3 identified as cisgender women, mean age = 23 ± 3.3 years) were recruited, none of whom had taken part in Experiment 1. Musicians started playing their instrument at a mean age of 8.1 ± 3.3 and had been playing for 14.1 ± 2.9 years on average. 5 out of 12 musicians studied modern music/jazz.

Stimuli

The stimuli were identical to those used in Experiment 1 except for the proportion of appearance of chords. In the present experiment, 50% of the sequences ended with the tonic and 50% ended in a dissonant cluster. Contrary to Experiment 1, there were no dissonant clusters at the third position of the sequence.

Procedure

The experimental procedure, including the secondary task, was identical to that of Experiment 1. Each experiment was ran separately.

Analyses

To compare the mean amplitudes of the ERPs between conditions, we performed mixed-design repeated measures ANOVA with the within-subject factors Region of Interest (ROI; left-anterior, right-anterior), Chord Type (in-key, dissonant) and the between-subject factor Group (musician, non-musician). Moreover, for assessing the effect of proportion of clusters as ending chords, we compared the mean amplitudes of Experiment 1 (where 25% of sequences ended in a dissonant cluster) and Experiment 2 (where 50% did) with a between subjects univariate ANOVA with the amplitudes of the difference waves and the fixed factors Experiment (25% vs 50%) x Group (musicians vs non-musicians).

Results

Dissonant chords

We observed a negative response for dissonant chords in comparison to in-key chords peaking around 200ms with a fronto-central distribution. We also observed a frontal positivity around 350ms followed by a late parietal positivity and a late frontal-bilateral negativity that was only evident in non-musicians (Fig 5).

For the 150-250ms time window, the ANOVA revealed an effect of Chord ($F(1, 22) = 29.23$, $p < .001$, $\eta 2 = .57$, 95% CI = -1.76, -0.78) which suggested that frequent dissonant chords in comparison to in-key chords elicit a negative response in both groups (-1,27 ± 0.24). Such effect was not lateralized, as indicated by the lack of interaction between Chord and ROI. Moreover, the lack of interaction between Chord and Group indicated that there was no difference between musicians and non-musicians. This contrasted with the responses that we observed in Experiment 1, where there was an enhancement in the EN in musicians.

For the 300-450ms time window, the ANOVA revealed an effect of Chord ($F(1, 22) = 16.44$, $p = .001$, $\eta 2 = .43$, 95% CI = 0.53, 1.68) as dissonant endings elicited a positivity in comparison

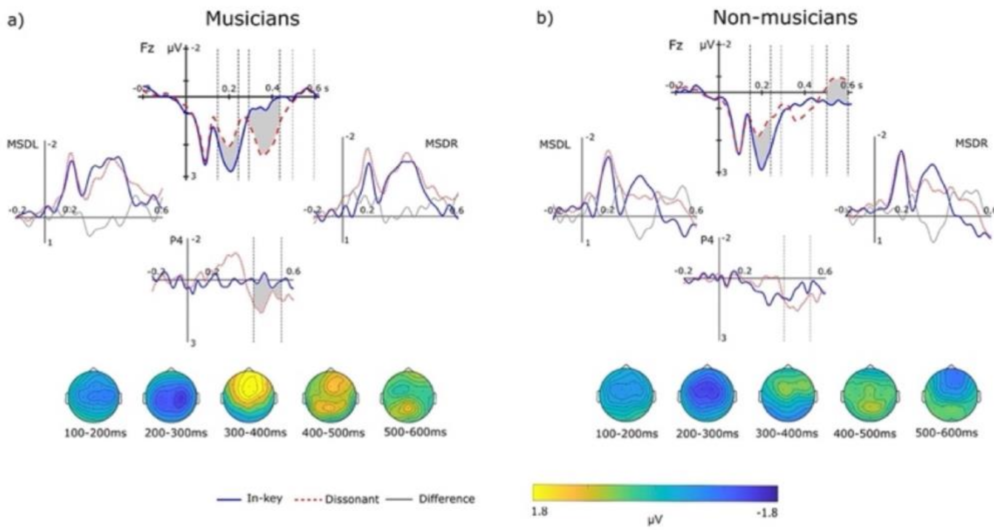


Fig 5. Experiment 2: Event-related potentials of musicians and non-musicians. Graphs include grand-average ERPs and topographical distribution of the difference waves (dissonant minus in-key) of musicians (a) and non-musicians (b) appearing in 50% of the sequences. Dotted vertical lines indicate the limits of the time windows of interest analysed. Shaded areas correspond to significant ERP effects, which are best seen in electrode F8. Left (MSDL) and right mastoidal (MSDR), and parietal (P4) are also shown. Topographical plots depict the grand-average difference wave to better observe the evolution of the topographical distribution of each ERP of interest.

https://doi.org/10.1371/journal.pone.0260728.g005

to in-key chords (1.31 ± 0.32). The interaction between Chord and Group ($F(1, 22) = 3.23$, $p = .085$, $\eta 2 = .13$) was only marginally significant, with musicians having a slightly larger response than non-musicians (1.90 ± 0.46 vs 0.73 ± 0.46), as supported by the post hoc test ($p < .001$, 95% CI = 0.948, 2.850 in musicians).

For the 350-550ms time window the ANOVA yielded an effect of Chord $F(1, 22) = 8.02$, $p = .010$, $\eta 2 = .27$), suggesting that dissonant chords elicited a P3b.

For the 500-600ms time window, the ANOVA yielded a main effect of Chord ($F(1, 22) = 4.36$, $p = .049$, $\eta 2 = .17$) and an interaction between Chord and Group ($F(1, 22) = 12.04$, $p = .002$, $\eta 2 = .35$). Thus, dissonant endings elicited a negativity in comparison to in-key chords (-0.37 ± 0.18). The negativity was larger in non-musicians than in musicians (0.98 ± 0.25 vs 0.24 ± 0.25), as supported by the post hoc test, that reached significance in non-musicians ($p = .001$, 95% CI = -1.50, -0.46).

Deviant instruments. Deviant instruments elicited a central negativity peaking around 200ms and a fronto-central positivity around 300ms (Fig 3). The ANOVA confirmed these responses for both the 150-200ms time window ($F(1, 22) = 45.12$, $p < .001$, $\eta 2 = .67$, 95% CI = -4.62, -2.44) and the 300-400ms window ($F(1, 22) = 212.01$, $p < .001$, $\eta 2 = .91$, 95% CI = 6.87, -9.15). We observed no difference between groups.

The performance in the behavioural task was 86.11% ± 5.55 in musicians and in non-musicians, 85.41% ± 5.34. As in Experiment 1, this suggests that participants were paying attention to the task in a consistent manner during the whole experiment.

Effect of proportion (Experiment 1—25% vs Experiment 2—50%). Both in Experiment 1 and Experiment 2, clusters elicited similar responses at the time windows of interest (best seen in difference waves, Fig 4, right panel). To explore whether the amplitude of these ERPs was in fact affected by the predictability of ending clusters we compared them between experiments.

For the 150-250ms time window, the ANOVA showed that the main effects of Experiment and Group were non-significant.

For the time windows of the subcomponents of the P3b (300-450ms and 350-550ms), the effect of Experiment was non-significant. However, the main effect of Group was significant for both time windows ($F(1,44) = 15,45$, $p < .001$, 95%CI = 0.51, 1.60 and $F(1, 44) = 7.821$, $p = .007$, 95%CI = 0.33, 2.04, respectively), which could indicate that the P3a and P3b was larger in musicians in both experiments.

For the 500-600ms time window, the main effect of Experiment was non-significant but the effect of Group was significant ($F(1,44) = 9.44$, $p = .004$, 95%CI = -1.45, -0.30), which might indicate that the N5 was larger in non-musicians in both settings.

These results suggest that the amplitude of the ERPs of interest were not significantly modulated by the proportion of appearance of dissonant chords as cadence closures.

Discussion

In the present experiment, dissonant clusters presented frequently elicited an EN with a fronto-central distribution in both musicians and non-musicians. Clusters also elicited a P3a and a P3b that were significant in musicians and an N5 that was significant in non-musicians. The amplitude of these responses did not significantly change when compared with those observed in Experiment 1.

Early negativity

The comparison between Experiment 1 and Experiment 2 showed that the amplitude of the EN is independent of the predictability of dissonant chords as cadence closures, at least for the present manipulation.

The most likely explanation could be that the syntactic hierarchy of the cadences (or even the organization of chords into cadences) might not be accessible at early stages of processing clusters. That could be linked to the psychoacoustic difficulty in processing roughness. Clusters represented 10% of all the chords presented. Thus, the alternation of consonant and dissonant chords may have simply elicited a MMN as in an oddball paradigm. Moreover, in both experiments, the same number of sequences (50%) contained a cluster (either always at the end or in two positions), causing that the proportion of sequences containing *a* cluster was the same across experiments. However, previous studies with similar paradigms showed that the responses to syntactic irregularities (such as a Neapolitan sixth or a supertonic as a cadence closure) decrease after enough repetitions [5,18]. Thus, future research should address whether the lack of effect of predictability is due to the psychoacoustical difficulties for processing roughness [54] or whether is more related to a difficulty in extracting the experimental regularities.

P3 effects

Frequent ending clusters elicited a P3a and a P3b, which were significant in musicians (Fig 5), but not in non-musicians. When compared with the results from Experiment 1, we observed that the amplitude of the P3 did not change as a function of the proportion of appearance of dissonant endings. That is in contrast with previous research, that shows that the P3a and P3b decline after the exposure to repeated syntactic irregularities [19,55], as listeners become able to predict the appearance of irregular events [56]. Moreover, musicians show habituation of the P3a during passive exposure to pitch deviants while non-musicians show an enhancement of that neural signature [49]. That the P3 did not decrease in our study after sustained repetition of clusters could be due to the fact that dissonant clusters are a very disruptive violation of sound expectancy based on the consonance dimension. Moreover, a dissonant closure of a chord cadence can hardly become less surprising, even if it was becoming easier to anticipate.

N5 effects

Frequent dissonant clusters elicited a N5 that was significant in non-musicians. Previous studies argue that, when listeners familiarize to frequently occurring violations (like Neapolitans that can act as subdominants), their integration is facilitated [5]. However, clusters are very unlikely to be perceived as a function of the key, independently of how often they appear. Thus, the presence of the N5 suggests that frequently appearing ending clusters are still difficult to integrate in the cadence, at least for non-musicians. Similarly, to Experiment 1, the strong attentional effects (P3a) elicited by dissonant clusters might have masked the N5 in musicians.

General discussion

In the present study, we investigated to what extent the neural responses to acoustic violations introduced by dissonant chords are modulated in a similar manner as the responses to syntactic irregularities are [5,9,10], with the aim to assess to what extent they share neural resources. The factors that we took into account were the musical training of the listener, and the musical and experimental context. First, we aimed to clarify whether the neural responses elicited by dissonant chords depend on the syntactic hierarchies created by chord cadences. Second, we investigated to what extent an increased predictability of dissonant closures of cadences modulates the subsequent neural responses. We found that unexpected dissonance elicited responses related to the detection of the unexpected event and subsequent responses related to the attraction of attention of the listener, structural analyses based on internal model updates and the integration into the previous context. We also found that musical training shapes how each of these neural responses differently interact with syntactic and experimental regularities.

The present study shows that both musicians and non-musicians readily detect dissonant clusters as an irregularity, as reflected by the EN that we observed in Experiments 1 and 2. The EN that we observed was not only a N2b because it inverted polarity at mastoid leads (Figs 2 and 5) which the N2b does not [9,57]. Even if more evidence is accumulated in favour of in-key chords and expectation for resolution are stronger at the end of the cadence, the amplitude of the EN did not significantly differ depending on the position of dissonant chords. This is in line with previous research showing that the acoustic deviance of dissonant chords is processed independently of their harmonic expectedness [32]. Moreover, even if the predictability of dissonance as a cadence closure was high, the amplitude of the EN did not decrease. Thus, early responses might be sensitive to the acoustic deviance introduced by dissonance (which did not differ across positions in terms of ASTM), but not to the higher-order schematic expectations (induced by syntactic regularities) and veridical expectations (induced by experimental regularities about frequency of appearance). Dissonance engages psychoacoustic constraints [54] because the roughness that it introduces is relevant at the auditory level, even to listeners with amusia [58,59]. Besides introducing an out-of-key pitch (*f#*; [9,10,21] clusters in the present study engaged a process of pure sensory analysis based on the roughness and beating of the chord [30,60]. Thus, the auditory system may have simply reacted to clusters as infrequent dissonant auditory deviants among frequent consonant chords [7,22]. By contrast, the early responses to syntactic deviance in the absence of acoustic deviance is in fact modulated by syntactic [9] and experimental regularities [18]. Here, we provide evidence that syntactic and acoustic deviance might engage different processes at early latencies. This pattern of results in music resembles language in the sense that syntax and changes in the acoustic information (phonetics) can be processed separately. Our results might be thus relevant to the current debate about whether music and language share neural resources [61,62 but see 11].

Previous studies had already reported increased MMN responses for dissonant chords [30,39] in comparison to consonant chords. In contrast, in our study, dissonant chords were presented within a musical context, which recruited subsequent higher-order processes of analysis reflected in the late neural responses (both the P3 and the N5). Once the sensory deviance of dissonance has been assessed [19], if a change in any stimulus attribute is detected, an attention-driven mechanism (reflected in the P3a) leads to further processing for updating the mental model of previous events in working memory (which is reflected in the P3b) [20,45,51]. Importantly, the larger P3 (P3a and P3b) that we observed after ending clusters suggests that they were more surprising than middle clusters, at least in trained listeners. In line with this, mistuned chords are rated worse as the tonality gets more established, although that is not reflected in an increase in the early neural responses (such as the MMN; [63]. Moreover, the lack of reduction in the P3 components suggests that, even if clusters became highly predictable through repetition, that did not render them less surprising, likely because they are strongly disruptive as a cadence closure.

Effects of musical training

In a previous study, musicians displayed an increased EN for dissonant endings in comparison to Neapolitan endings [33]. However, it remained unclear whether such enhancement reflected the reaction to the greater sensory dissonance or the cognitive processing of clusters as stronger violation of musical context than out-of-key chords. In the present study, we replicated these results and clarified that musical-syntactic hierarchies do not influence such early processing of dissonance, as reflected by the lack of difference in the EN across positions of the cadence. Thus, the increase of the EN that we observed in musicians might reflect their more efficient processing of dissonant chords [30,32,40,53], and maybe their classification as non-prototypical chords [30]. Meanwhile, non-musicians might have more difficulty in processing the roughness of dissonance at early latencies [54] and only distinguish more conventional distinctions, such as mistuning [30].

Our results also show that musical expertise is linked to an enhancement of the P3 response toward dissonant clusters. Although we lack direct behavioral measurements of attention allocation to dissonant clusters, previous studies suggest that musicians possess faster involuntary attention allocation than non-musicians (as reflected by the P3a [2,49]). Moreover, musicians in our study showed a larger P3b, which is linked to processes of structural analysis in rule-governed sequences arising from the revision of the stimulus into the current model of the musical structure [49,50] together with the coordination of behavioral responses. That might suggest that musicians possess a more efficient updating of the sensory model [49]. The observed P3 subcomponents interacted with syntactic hierarchies suggesting that even if the early automatic detection of is independent of musical rules, musicians are still able to efficiently sort its unexpectedness according to the musical context. In short, dissonant chords might represent a stronger surprise to musicians, especially when it is placed as a cadence closure. Meanwhile, we observed an EN and a N5 in non-musicians, but neither a P3a nor a P3b. That suggests that non-musicians may have detected clusters as dissonant deviants and even attempted to integrate them, but that these clusters did not represent such a strong surprise as they did in musicians.

Conclusion

Our study focused on the influence that syntactic (schematic) and experimental (veridical) regularities have on how the brain reacts to a purely acoustic violation such as strong dissonance. Our results show that, at early latencies, acoustic deviance may be processed independently of schematic and veridical expectations. They also show that musical training modulates the neural responses to unexpected dissonance by enhancing and refining them. Musicians showed larger responses related to the detection of dissonance (EN) and the attraction of attention (P3a), suggesting that dissonance in a musical context are more surprising for musicians. At late latencies, the neural responses of musicians reflective of attention allocation (P3a) and model-updating (P3b) were indeed influenced by syntactic (but not experimental) regularities. Moreover, while it is likely that the P3 effects overrode the N5 in musicians, in non-musicians the presence of the N5 suggested that dissonance engaged mechanisms of integration. Thus, our study advances the understanding of the processing of music by exploring the factors that modulate the neural responses to acoustic violations in a musical context and contributes to untangling the neural mechanisms behind the processing of syntactic irregularities and acoustic disruptions.

Supporting information

S1 Data.
(XLS)

S2 Data.
(XLS)

Acknowledgments

We would like to thank David Ricardo Quiroga-Martínez for performing the ASTM modelling and for the fruitful discussions; and Evangelia Andreopolou for her help in the experimental sessions of the study.

Author Contributions

Conceptualization: Carlota Pagès-Portabella, Mila Bertolo, Juan M. Toro.

Data curation: Carlota Pagès-Portabella, Mila Bertolo.

Formal analysis: Carlota Pagès-Portabella, Mila Bertolo.

Funding acquisition: Juan M. Toro.

Investigation: Carlota Pagès-Portabella.

Methodology: Carlota Pagès-Portabella, Mila Bertolo.

Project administration: Juan M. Toro.

Resources: Juan M. Toro.

Software: Carlota Pagès-Portabella.

Supervision: Juan M. Toro.

Validation: Carlota Pagès-Portabella, Juan M. Toro.

Visualization: Carlota Pagès-Portabella.

Writing – original draft: Carlota Pagès-Portabella.

Writing – review & editing: Carlota Pagès-Portabella, Mila Bertolo, Juan M. Toro.

References

1. Jackendoff R. Parallels and nonparallels between language and music. Music Percept. 2009; 26:195–204. https://doi.org/10.1525/mp.2009.26.3.195
2. Guo S, Koelsch S. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. Sci. Rep. 2016; 6(1): 1–11.
3. Tillmann B, Bigand E. Musical structure processing after repeated listening: Schematic expectations resist veridical expectations. MUSICÆ SCIENTIÆ. 2010; 33–47. https://doi.org/10.1177/10298649100140S204
4. Justus TC, Bharucha JJ. Modularity in Musical Processing: The Automaticity of Harmonic Priming. J. Exp. Psychol. Hum. Percept. Perform 2001; 27(4): 1000–1011. https://doi.org/10.1037//0096-1523.27.4.1000 PMID: 11518142
5. Koelsch S, Gunter T, Friederici AD, Schröger E. Brain Indices of Music Processing: "Nonmusicians" are Musical. J. Cog. Neurosci. 2000; 12(3):520–541. https://doi.org/10.1162/089892900562183 PMID: 10931776
6. Janata P. ERP measures assay the degree of expectancy violation of harmonic contexts in music. J. Cog. Neurosci. 1995; 7(2):153–164. https://doi.org/10.1162/jocn.1995.7.2.153 PMID: 23961821
7. Leino S, Brattico E, Tervaniemi M, Vuust. Representation of harmony rules in the human brain: Further evidence from event-related potentials. Brain Res. 2007; 1142(1): 169–177. https://doi.org/10.1016/j.brainres.2007.01.049 PMID: 17300763
8. Loui P, Grent-'t-Jong T, Torpey D, Woldorff M. Effects of attention on the neural processing of harmonic syntax in Western music. Cogn. Brain Res. 2005; 25(3): 678–687. https://doi.org/10.1016/j.cogbrainres.2005.08.019 PMID: 16257518
9. Koelsch S, Jentschke S, Sammler D, Mietchen D. Untangling syntactic and sensory processing: An ERP study of music perception. Psychophysiology. 2007; 44(3): 476–490. https://doi.org/10.1111/j.1469-8986.2007.00517.x PMID: 17433099
10. Koelsch S, Sammler D. Cognitive components of regularity processing in the auditory domain. PLoS ONE, 2008; 3(7). https://doi.org/10.1371/journal.pone.0002650 PMID: 18612421
11. Bigand E, Delbé C, Poulin-Charronnat B, Leman M, Tillmann B. Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. Front. Syst. Neurosci. 2014; 8. https://doi.org/10.3389/fnsys.2014.00094 PMID: 24936174
12. Koelsch S, Schmidt BH, Kansok J. Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. Psychophysiology. 2002; 39(5): 657–663. PMID: 12236333
13. de Ridder D, Vanneste S, Freeman W. The Bayesian brain: Phantom percepts resolve sensory uncertainty. Neurosci Biobehav Rev. 2014; 44: 4–15. https://doi.org/10.1016/j.neubiorev.2012.04.001 PMID: 22516669
14. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 2004; 27 (12). https://doi.org/10.1016/j.tins.2004.10.007 PMID: 15541511
15. Friston K. The free-energy principle: A unified brain theory? Nat Rev Neurosci. 2010; 10: 127–38. https://doi.org/10.1038/nrn2787 PMID: 20068583
16. Lieder F, Daunizeau J, Garrido MI, Friston KJ, Stephan KE. Modelling Trial-by-Trial Changes in the Mismatch Negativity. PLoS Comput Biol. 2013; 9(2):1002911. https://doi.org/10.1371/journal.pcbi.1002911 PMID: 23436989
17. Quiroga-Martinez DR, Hansen NC, Højlund A, Pearce MT, Brattico E, Vuust P. Reduced prediction error responses in high-as compared to low-uncertainty musical contexts. Cortex. 2019; 120: 181–200. https://doi.org/10.1016/j.cortex.2019.06.010 PMID: 31323458
18. Koelsch S., Jentschke S. Short-term effects of processing musical syntax: An ERP study. Brain Res. 2008; 1212: 55–62. https://doi.org/10.1016/j.brainres.2007.10.078 PMID: 18439987
19. Guo S, Koelsch S. The effects of supervised learning on event-related potential correlates of music-syntactic processing. Brain Res. 2015; 1626: 232–246. https://doi.org/10.1016/j.brainres.2015.01.046 PMID: 25660849
20. Itti L, Baldi P. Bayesian surprise attracts human attention. Vision Res. 2009; 49(10): 1295–306. https://doi.org/10.1016/j.visres.2008.09.007 PMID: 18834898

21. Leman M. An auditory model of the role of short-term memory in probe-tone ratings. Music Percept. 2000; 17(4): 481–509. https://doi.org/10.2307/40285830

22. Koelsch S, Gunter TC, Schröger E, Tervaniemi M, Sammler D, Friederici AD. Differentiating ERAN and MMN: An ERP study. NeuroReport. 2001; 12(7): 1385–1389. https://doi.org/10.1097/00001756-200105250-00019 PMID: 11388416

23. Plomp R, Levelt WJM. Tonal Consonance and Critical Bandwidth. J. Acoust. Soc. Am. 1965; 38(548). https://doi.org/10.1121/1.1909741 PMID: 5831012

24. McDermott JH, Lehr AJ, Oxenham AJ. Individual differences reveal the basis of consonance. Curr. Biol., 2010; 20(11): 1035–1041. https://doi.org/10.1016/j.cub.2010.04.019 PMID: 20493704

25. Virtala P, Tervaniemi M. Neurocognition of major-minor and consonance-dissonance. Music Percept. 2017; 34(4):387–404. https://doi.org/10.1525/MP.2017.34.4.387

26. Johnson-Laird P, Kang O, & Chang Leong Y. On Musical Dissonance. Music Percept. 2012; 30(1): 1533–8312. https://doi.org/10.1525/mp.2012.30.1.19

27. Lahdelma I, Eerola T. Mild Dissonance Preferred Over Consonance in Single Chord Perception. I-Percept. 2016; 7(3). https://doi.org/10.1177/2041669516655812 PMID: 27433333

28. Schön D, Regnault P, Ystad S, Besson M. Sensory Consonance. Music Percept. 2005; 23(2): 105–118. https://doi.org/10.1525/mp.2005.23.2.105

29. Virtala P, Berg V, Kivioja M, Purhonen J, Salmenkivi M, Paavilainen P, et al. The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. Neurosci. Lett. 2011; 487(3): 406–410. https://doi.org/10.1016/j.neulet.2010.10.066 PMID: 21055444

30. Brattico E, Pallesen KJ, Varyagina O, Bailey C, Anourova I, Järvenpää M, et al. Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. J. Cogn. Neurosci. 2009; 21(11): 2230–2244. https://doi.org/10.1162/jocn.2008.21144 PMID: 18855547

31. Minati L, Rosazza C., D'Incerti L, Pietrocini E, Valentini L, Scaioli V, et al. Functional MRI/Event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. NeuroReport. 2009; 20(1): 87–92. https://doi.org/10.1097/WNR.0b013e32831af235 PMID: 19033878

32. Regnault P, Bigand E, Besson M. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. J. Cogn. Neurosci. 2001; 13(2): 241–255. https://doi.org/10.1162/089892901564298 PMID: 11244549

33. Pagès-Portabella C, Toro JM. Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians. Psychophysiology, 2019, 27(1): 1–11. https://doi.org/10.1111/psyp.13476 PMID: 31512751

34. Seppänen M, Hämäläinen J, Pesonen AK, Tervaniemi M. Passive sound exposure induces rapid perceptual learning in musicians: Event-related potential evidence. Biol. Psychol. 2013; 94(2): 341–353. https://doi.org/10.1016/j.biopsycho.2013.07.004 PMID: 23886959

35. Bharucha J, Kruhmansl CL. The representation of harmonic structure in music: Hierarchies of stability as a function of context. Cognition. 1983; 13(1): 63–102. https://doi.org/10.1016/0010-0277(83)90003-3 PMID: 6681743

36. Todd J, Petherbridge A, Speirs B, Provost A, Paton B. Time as context: The influence of hierarchical patterning on sensory inference. Schizof. Res. 2018; 191: 123–131. https://doi.org/10.1016/j.schres.2017.03.033 PMID: 28343741

37. Loui P, Wu EH, Wessel DL, Knight RT. A generalized mechanism for perception of pitch patterns. J. Neurosci. 2009; 29(2): 454–459. https://doi.org/10.1523/JNEUROSCI.4503-08.2009 PMID: 19144845

38. Koeslch S. Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. Pyschophysiology. 2009; 46(1):179–190. https://doi.org/10.1111/j.1469-8986.2008.00752.x PMID: 19055508

39. Bigand E, Parncutt R, Lerdahl F. Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Percept. Psychophys. 1996; 58(1): 125–141. https://doi.org/10.3758/BF03205482

40. Crespo-Bojorque P, Monte-Ordoño J, Toro JM. Early neural responses underlie advantages for consonance over dissonance. Neuropsychologia. 2018; 9(3): 183–186. https://doi.org/10.1016/j.neuropsychologia.2018.06.005 PMID: 29885961

41. Proverbio AM, Orlandi A, Pisanu F. Brain processing of consonance/dissonance in musicians and controls: a hemispheric asymmetry revisited. Eur. J. Neurosci. 2016; 44(6): 2340–2356. https://doi.org/10.1111/ejn.13330 PMID: 27421883

42. Neuloh G, Curio G. Does familiarity facilitate the cortical processing of music sounds? NeuroReport. 2004; 15(16): 2471–2475. https://doi.org/10.1097/00001756-200411150-00008 PMID: 15538177

43. Vuust P, Brattico E, Seppänen M, Näätänen R, Tervaniemi M. Practiced musical style shapes auditory skills. Ann. N. Y. Acad. Sci. 2012; 1252(1): 139–146. https://doi.org/10.1111/j.1749-6632.2011.06409.x PMID: 22524351

44. Escera C, Alho K, Schröger E, Winkler I. Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol. Neurootol. 2000; 5(3–4): 151–166. https://doi.org/10.1159/000013877 PMID: 10859410

45. Quiroga-Martinez DR, Hansen NC, Højlund A, Pearce M, Brattico E, Vuust P. Decomposing neural responses to melodic surprise in musicians and non-musicians: Evidence for a hierarchy of predictions in the auditory system. NeuroImage. 2020; 215:116816. https://doi.org/10.1016/j.neuroimage.2020.116816 PMID: 32276064

46. Sussman E, Winkler I, Schröger E. Top-down control over involuntary attention switching in the auditory modality. Psychon. Bull. Rev. 2003; 10(3): 630–637. https://doi.org/10.3758/bf03196525 PMID: 14620357

47. Koelsch S, Vuust P, Friston K. Predictive Processes and the Peculiar Case of Music. Trends Cog. Sci. 2019; 23(1): 63–77. https://doi.org/10.1016/j.tics.2018.10.006 PMID: 30471869

48. Steinbeis N, Koelsch S, Sloboda JA. The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. J. Cogn. Neurosci. 2006; 18(8): 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380 PMID: 16859422

49. Seppänen M, Pesonen AK, Tervaniemi M. Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. Atten. Percept. Psychophys. 2012; 74: 600–612. https://doi.org/10.3758/s13414-011-0257-9 PMID: 22222306

50. Koelsch S, Siebel WA. Towards a neural basis of music perception. Trends Cogn. Sci. 2005; 9(12): 578–584. https://doi.org/10.1016/j.tics.2005.10.001 PMID: 16271503

51. Polich J. Updating P300: An integrative theory of P3a and P3b. 2007; 118(10): 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019 PMID: 17573239

52. Carrión RE, Bly BM. The effects of learning on event-related potential correlates of musical expectancy. Psychophysiology. 2008; 45(5): 759–775. https://doi.org/10.1111/j.1469-8986.2008.00687.x PMID: 18665861

53. Itoh K, Suwazono S, Nakada T. Central auditory processing of noncontextual consonance in music: An evoked potential study. J. Acoust. Soc. Am. 2010; 128(6): 3781–3787. https://doi.org/10.1121/1.3500685 PMID: 21218909

54. Bidelman GM, Krishnan A. Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem. J. Neurosci. 2009; 29(42): 13165–13171. https://doi.org/10.1523/JNEUROSCI.3900-09.2009 PMID: 19846704

55. Gunter TC, Stowe LA, Mulder G. When syntax meets semantics. Psychophysiology. 1997; 34(6): 660–676. https://doi.org/10.1111/j.1469-8986.1997.tb02142.x PMID: 9401421

56. Sommer W, Leuthold H, Matt J. The expectancies that govern the P300 amplitude are mostly automatic and unconscious. Behav. Brain Sci. 1998; 21(1): 149–150. https://doi.org/10.1017/s0140525x98210958

57. Schröger E. Measurement and interpretation of the mismatch negativity. Behav. Res. Methods Instrum. Comput. 1998; 30(1): 131–145. https://doi.org/10.3758/BF03209423

58. Cousineau M, McDermott JH, Peretz I. The basis of musical consonance as revealed by congenital amusia. PNAS USA. 2012; 109(48): 19858–19863. https://doi.org/10.1073/pnas.1207989109 PMID: 23150582

59. Brattico E, Tervaniemi M, Valimaki V, Zuigen T, Peretz I. Cortical correlates of acquired deafness to dissonance. Ann. N. Y. Acad. Sci. 2003; 999:158–160. https://doi.org/10.1196/annals.1284.019 PMID: 14681131

60. Brattico E, Tervaniemi M, Näätänen R, Peretz I. Musical scale properties are automatically processed in the human auditory cortex. Brain Res. 2006; 1117(1): 162–174. https://doi.org/10.1016/j.brainres.2006.08.023 PMID: 16963000

61. Koelsch S, Gunter TC, Wittfoth M, Sammler D. Interaction between syntax processing in language and in music: An ERP study. J. Cogn. Neurosci. 2005; 17(10):1565–77. https://doi.org/10.1162/089892905774597290 PMID: 16269097

62. Steinbeis N, Koelsch S. Shared Neural Resources between Music and Language Indicate Semantic Processing of Musical Tension-Resolution Patterns. Cerebral Cortex. 2008; 18(5): 1169–78. https://doi.org/10.1093/cercor/bhm149 PMID: 17720685

63. Garza Villarreal EA, Brattico E, Leino S, Østergaard L, Vuust P. Distinct neural responses to chord violations: A multiple source analysis study. Brain Res. 2011; 1389: 103–114. https://doi.org/10.1016/j.brainres.2011.02.089 PMID: 21382359

4. STIMULI SIMULATION WITH A COMPUTATIONAL MODEL OF AUDITORY SHORT-TERM MEMORY (ASTM)

A study by Bigand and collaborators (2014) challenged the traditional cognitive music-syntactic account by simulating the stimulation paradigms of previous literature with an auditory short-term memory (ASTM) model (Leman, 2000). They showed that most of the "music-specific cognitive" ERPs may be more parsimoniously explained by the perceptual contrasts that physical features create with those of previous events accumulated in ASTM. Thus, Bigand and collaborators proposed that simulating stimuli based on an ASTM model may be a more reliable control for sensory contributions than, for instance, counting the musical tones shared by the target and the preceding context.

One of the main issues that needed clarification in our electrophysiological data was the extent of the influence of auditory sensory mechanisms to the registered neural responses. We thus presented the stimuli from our electrophysiological studies to the auditory short-term memory model proposed by Leman (2000). The model is freely available as a Matlab toolbox (Leman et al., 2001, see www.ipem.ugent.be/Toolbox). The ASTM model is useful in defining and quantifying sensory expectations, because it compares the current pitch (or chord) representation with the accumulated previous context and provides an estimate of the sensory surprise that a pitch (or chord in a specific position) generates at every time point. In other words, it gives an estimate of the accumulation of evidence towards a specific tonality.

Listening to a chord progression leads to an integration of auditory images over a time-window, and the resulting integrated auditory image can be conceived as occupying a particular region of the tonal space. This spot or activation region thus represents a time-dependent context with which target chords can be compared. If the target chord falls outside this region, then the similarity will be low and vice-versa (Leman, 1995). The auditory short-term memory model has four components: in the first component, audio signals are transformed into neural firing probabilities by a model of the "peripheral auditory system" (Van Immerseel and Martens, 1992). The audio signal is low and band pass filtered (to simulate the filtering of the outer, middle and inner ear) and then half-wave rectified and compressed to simulate the conversion into neural rate-code patterns by the hair-cells, resulting in an auditory nerve image (ANI). The second component performs a summed "pitch periodicity analysis" of the ANI for all auditory channels, based on a (windowed) autocorrelation function, which results in a summary pitch image (PI). In the third component, the "echoic memory" model performs a leaky integration (low pass filtering over time), so the PI becomes smeared over time. In this component, the duration that incoming pitch images kept in short-term memory can be manipulated. This component of the model is key: with shorter echo, a local pitch image (LPI) is obtained, and it represents the pattern for the immediate pitch percept. With a longer echo (taking in consideration more of the preceding contextual information or pitch context), a global pitch image (GPI) is obtained. Finally, the "tonal contextuality" (TC) index calculates the tension of a LPI with respect to the GPI over time and we obtain a continuous stream of TC values, one per each time point. Crucially, higher TC values reflect higher similarity between the two images, while lower values represent tension.

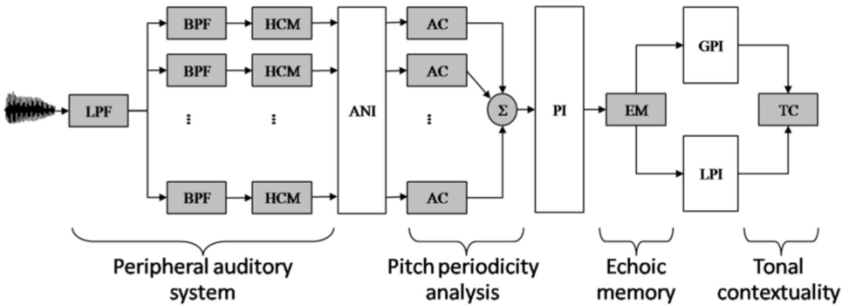


Figure 12. Schematic diagram of the auditory model of Leman (2000). Adapted from Bigand et al. (2014).

For simplicity, from now on, we will refer to the study reported in Manuscript 1 as Study 1, and the studies reported in Manuscript 2, as Study 2 and Study 3.

4. 1 Simulation of Study 1

First, the mean TC differences "tonic *minus* Neapolitan" chords (Figure 13, left panel) and "tonic *minus* cluster" chords (Figure 13, central panel) were positive. That suggests that both kinds of deviant chords elicited a stronger dissimilarity in ASTM (which is reflected in lower TC values) than did the in-key tonic chords (which reflected in higher TC values). That effect was more localized in the region of short local integration windows (which fit the latency of early negativities) and longer global integration windows. This result is coherent with the elicitation of a sensory early negativity as a response to the deviation in ASTM and would account for the ERPs. Note that we report TC computed for the time window of 0-600ms post-onset for simplicity, but simulations for the time window of 200ms ± 25ms post-onset (which would correspond specifically to the latency of early negativity) provided essentially the same results.

The comparison between Neapolitans and clusters (Figure 13, right panel) revealed that the mean TC difference values were negative (for the same parameter space as described above). That means that dissonant clusters have higher TC values than Neapolitans, suggesting that clusters are more contextual (i.e. less violating). Likely, this is due to clusters sharing two notes with the tonic chord whereas Neapolitans share only one. This simulation would suggest that ASTM would not account for the ERPs because, in musicians, we found larger early negativities for dissonant than for Neapolitan chords. In non-musicians, we observed no difference between types of irregular chords. Likely, the early negativities that we observed were responses to the tonal irregularity instilled by Neapolitan chords (which is reflected in the ASTM model) and by the roughness instilled by dissonant clusters (which is not reflected in the ASTM model). In the case of musicians, who showed larger responses to dissonant clusters, acoustic dissonance (consisting of roughness) may be perceptually more relevant at early time windows than tonal deviation and than to non-musicians.

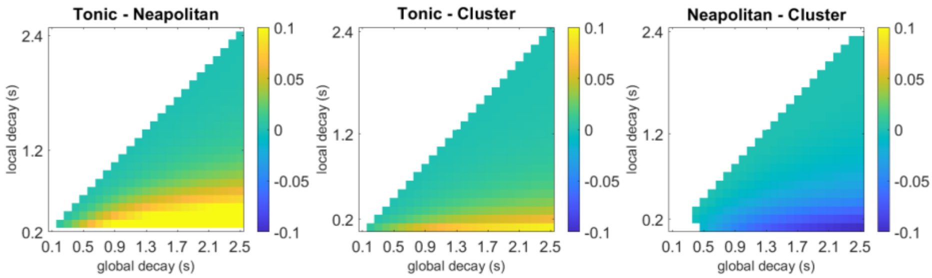


Figure 13. Simulation data obtained with the stimuli of Study 1. Mean differences between the TC values as a function of local and global context integration windows between tonic, Neapolitan and dissonant chords for the time window 0-600ms. Positive, negative and nonsignificant differences are represented by hot, cold and white colors, respectively (two-paired t test, p<.05).

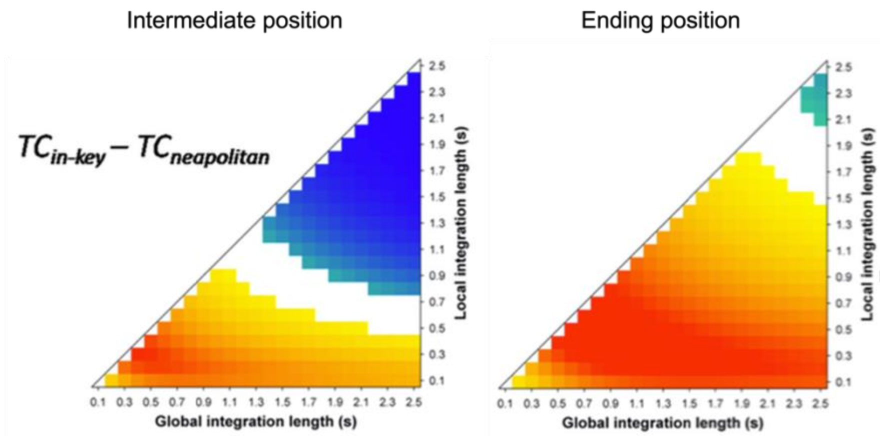


Figure 14. Simulation data obtained with the stimuli of Leino et al. (2007) where they compared tonic and Neapolitan chord for the 200±25ms for intermediate and ending positions. Adapted from Bigand et al. (2014).

4. 2 Simulation of Study 2

In principle, ASTM alone should be able to explain position effects in the ERPs because memory of auditory images decays over time. For instance, with the succession of in-key chords in a row, the ASTM model grows more and more confident that the next chord will be in-key and therefore out-of-key chords would be more and more surprising as the sequence progresses (Figure 14). Note that the same phenomenon can be explained from different theoretical frameworks: listeners create a mental model of chord sequences termed "evidence" under a Bayesian framework (Friston, 2010) / "predictive model" under a predictive processing framework (Koelsch, Vuust & Friston, 2019) / "syntax" from a music-language framework (Koelsch et al., 2000) / "memory" from an ASTM framework (Bigand et al., 2014). They all reflect the same idea: as chord sequences progresses in time, more evidence accumulates/predictions get stronger/there is a build-up of the syntactic phrase/there is a stronger memory trace, that affects how much surprised we are by new chords. Although we lack a computational approach to syntactic expectations of our stimuli, we believe that it would predict similar responses to clusters at both positions: because the tonic is equally appropriate at both positions in terms of syntax, a dissonant version of the tonic triad should be similarly incorrect at all positions.

The simulation of the stimuli from Study 2 revealed that the TC difference "tonic at the fifth position minus cluster at the fifth position" (Figure 15, left panel) was significant and positive (yellow area) for short integration windows (the latency of the early negativity) and larger global integration windows. That suggests that clusters at the fifth position elicited a larger dissimilarity than tonic chords and that ASTM could account for early time windows. For the rest of the parameter space the TC difference was negative (blue area), suggesting that ASTM could not account for differences in the ERPs. That the effect was larger in both comparisons for longer global decays (also referred to as sensory buffer) might suggest that the more information about the tonality is accumulated by hearing more context chords, the larger is the difference between context and target chords. Similar results were found for the TC difference between "tonic at the third position *minus* cluster at the third position" (Figure 15, central panel), it was positive for short local decays and only for longer global decays. Note that TC differences of Study 2 were smaller in comparison to Study 1, which may suggest that having dissonant chords at both at intermediate and ending positions may reduce the dissimilarity because their pitch images globally linger on ASTM longer.

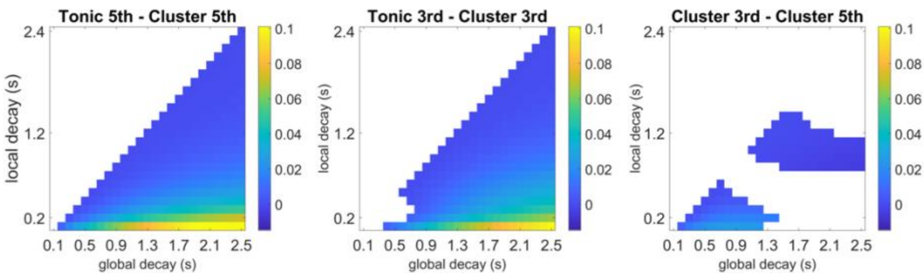


Figure 15. Simulation data obtained with the stimuli of Study 2. Mean differences between the TC values for the time window 0-600ms as a function of local and global context integration windows between tonic and dissonant chords at the third or fifth position of the cadence. Positive, negative ana nonsignificant differences are represented by hot, cold and white colors, respectively (two-paired t test, p<.05).

Crucially, the TC difference "clusters at the third position *minus* clusters at the fifth position" was non-significant for most of the parameter space (Figure 15, right panel). It was only positive (but very small) for short global decays. That could suggest that ending clusters are less contextual than intermediate clusters only when taking into account the immediately previous context (which may include the last 1 or 2 chords). Moreover, global decay values of clusters at the third position may also be integrating ending clusters from the previous sequence. It is also possible that the lack of differences between positions is due to accumulated memory traces throughout the block. In general, ASTM could not fully account for differences in the ERPs, if there were any (see *General discussion* for further information).

4. 3 Simulation of Study 3

The simulation of Study 3 revealed that the TC difference "tonic chord 50% *minus* cluster 50%" (Figure 16, left panel) was positive for short local decays and longer global decays. Note that, in comparison to Study 2, the magnitude of the effect of this comparison was smaller. This would suggest that, when clusters are presented frequently, the auditory system may hold clusters longer in memory so that when they appear again, they introduce a smaller auditory deviation than when presented seldom.

The difference "tonic chord 75% (Study 2) *minus* tonic chord 50% (Study 3)" was negative for long local decays and nonsignificant for the rest of the parameter space (Figure 16, central panel). If anything, this result would imply that the dissimilarity between tonic chords and the context was larger when they were presented less frequently. That would make sense because the representation of correctly ending sequences would become weaker than if presented often.

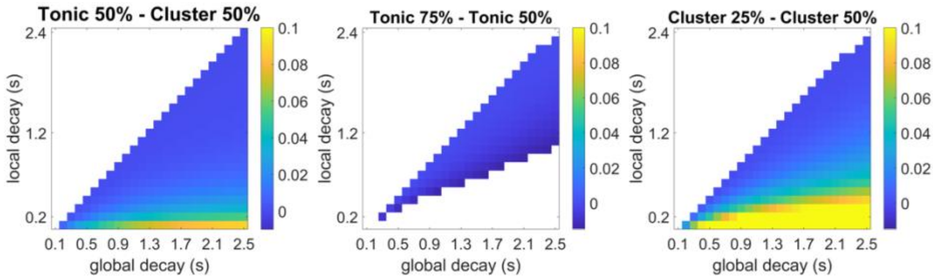


Figure 16. Simulation data obtained from the comparison between stimuli of Study 2 and Study 3. Mean differences between the TC values for the time window 0-600ms as a function of local and global context integration windows between tonic and dissonant chords presented 50% of the cases (left panel), tonic chords presented 75% or 25% of the cases (central panel) and dissonant chords presented 25% or 50% of the cases (right panel). Positive, negative and nonsignificant differences are represented by hot, cold and white colors, respectively (two-paired t test, p<.05; colorscale and mask is the same as in Study 2).

Finally, the difference "cluster 25% *minus* cluster 50%" was positive for all global integration windows but for short local integration windows (Figure 16, right panel). As suggested by the fact that the TC difference "tonic 50% *minus* cluster 50%" was smaller than in Study 2, this result indicates that clusters presented in 50% of the cases introduce a smaller dissimilarity with the context when compared to clusters presented 25% of the cases. As suggested above, that reflects that the auditory system may hold clusters longer in memory (maybe due to accumulated memory traces throughout the block), so that they become easier to anticipate. Thus, although ASTM would account for a (slight) reduction of the ERPs, we registered no difference between Studies 2 and 3 (at least when directly comparing between experiments) maybe because roughness is not considered in the simulation (see *5. General Discussion* for further details).

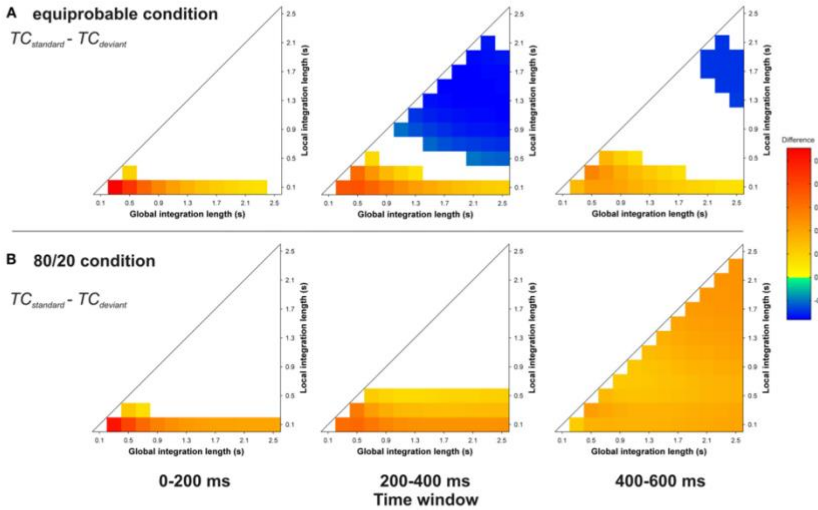


Figure 17. Simulation of the data of Loui et al, 2009. Mean TC differences between standard and deviant target chords built over a previously unheard musical scale (the Bohlen-Pierce scale; Mathews at al., 1988) with (A) equiprobable appearance or (B) 80% standards/20%deviants proportion for different time windows. (Adapted from Bigand et al., 2014).

5. GENERAL DISCUSSION

5.1 Aims

Throughout all our studies we were interested in investigating to what extent the neural responses to musical irregularities are different between musicians with long-term musical training and naïve listeners. For the sake of comparing the behavioural and electrophysiological results of our studies, from now on we will report them together as different sections of Study 1 and Study 2. With this comparison, we aimed to assess whether the neural responses registered in the electrophysiological studies were qualitatively good predictors of the pleasurableness and correctness judgements. Besides, each individual study addressed one specific question of interest in the current literature and how different relevant factors interact with musical training.

Study 1: are tonal-syntactic and strongly dissonant violations perceived and processed differently? Are any putative differences modulated by musicianship? In the electrophysiological study, we compared the ERPs triggered by Neapolitan chords with those of dissonant versions of the tonic. Neapolitan chords are syntactically incorrect as a cadence ending (although they are more common as a substitute of the subdominant), introduce a tonal violation (because they include out-of-key tones), but they can be enjoyable. Dissonant versions of the tonic would be also syntactically incorrect as cadence closures but, in terms of ASTM, introduce less deviance than Neapolitan chords and are considered unacceptable and unpleasant in most western music traditions. Even more, subjective evaluations of musical events tend to differ as a function of musical expertise. We aimed to investigate whether these differences are reflected in the conscious judgements and brain reactions of musicians and non-musicians.

Study 2: to what extent changes in the acoustic information (such as dissonance) of music are perceived and processed independently of schematic (syntactic) expectations? Is this dependent of musicianship? As a follow-up of Study 1, in Study 2 we investigated whether musicians and non-musicians recruit music-syntactic mechanisms when processing dissonance within a tonal context of chord cadences. To do so, we examined whether the neural responses to dissonant clusters change when placed at different positions of a chord cadence, where they violate different amounts of expectation/instil less or more tonal-syntactic deviance, respectively. We also evaluated whether syntactic expectations affected their conscious affective and cognitive judgements.

Study 3: how does the processing of dissonance in a structured musical context interact with veridical expectations (in the form of high probability of occurrence)? Do musicians and non-musicians react differently to veridical expectations? As a follow-up of Study 2, we aimed to investigate whether the neural responses to dissonant endings of chord cadences are modulated when they become predictable (even if they still introduce a violation), in order to compare them with tonal-syntactic violations. To do so, we examined how the neural responses to unexpected dissonant endings of chord cadences change when their probability of occurrence is high.

5.2 Summary of results

5.2.1 Study 1

5.2.1.1 Behavioural results (active listening task)

Behavioural results showed that the tonic chord as a cadence ending is highly preferred by both musicians and non-musicians in both the pleasurableness and correctness scales. However, musicians rated the tonic significantly higher than non-musicians. Moreover, musicians and non-musicians similarly evaluated dissonant endings of chord cadences as very incorrect. When evaluating Neapolitan endings, musicians and non-musicians behaved differently. Musicians rated Neapolitan endings as significantly more correct and pleasant than non-musicians did, and almost as appropriate as they rated the tonic ending.

5.2.1.2 ERP results (passive listening task)

We analysed the differences in the ERPs observed at the time windows of interest 100-150ms (N1), 150-250ms (ERAN/MMN), 325-500ms (P3) and 500-600 (N5). We found that ambiguous and dissonant endings elicit a larger N1 than tonic endings. However, the N1 was not different between musicians and non-musicians, suggesting that any differences in successive responses were not simply due to enhanced attention to musical stimuli in musicians. Dissonant chords elicited an early negativity with a slight right tendency that was larger in musicians than in non-musicians. More importantly, we also found that the difference between ambiguous and dissonant endings was larger in musicians. Both types of irregular endings (ambiguous and dissonant) elicited a frontal P3 (presumably the P3a) with a right lateralization, that was much larger in musicians than in non-musicians. Surprisingly, we did not find any effect at the time window of the N5, as we would have expected.

5.2.2 Study 2

5.2.2.1 Behavioral results (active listening task)

Behavioural assessments showed that both types of listeners rated dissonant clusters very low in the scales of pleasurableness and correctness, and that such evaluation was similar between positions. Moreover, both groups rated Neapolitans higher than the dissonant chords at corresponding positions. Importantly, musicians (but not non-musicians) rated Neapolitan chords at the fifth position significantly lower than Neapolitans at the third position.

5.2.2.2 ERPs (passive listening task) to dissonant chords

We analysed the time-windows of interest 150-250ms (ERAN/MMN), 300-450ms (P3a), 350-550ms (P3b) and 500-600ms (N5). We found that dissonant chords elicited an early negativity (EN) with a right-frontal distribution. The EN was larger in musicians than in non-musicians but we found no difference between positions in the amplitude of EN. Moreover, dissonant chords elicited a frontal right-lateralized P3a, that was modulated by the interaction between position and musicianship together, as it was larger fifth than at the third position in musicians, but this difference did not reach statistical significance in non-musicians. We also found a parietal P3b that was modulated, on one hand, by musicianship (as it was larger in musicians than not in non-musicians) and, on the other hand, by position (because it was larger for ending chords). Finally, dissonant chords elicited a frontal bilateral N5, that was modulated by position (as it was larger at the third position), and by musicianship (as it was larger in non-musicians).

5.2.2.3 ERPs to deviant instruments

Deviant instruments elicited an N2b-P3a-P3b complex that was similar between groups. Musicians and non-musicians had a similar performance in detecting and counting deviant instruments above chance, suggesting that the difficulty of the task was acceptable for both groups.

In sum, the results of Study 2 showed that early responses were insensitive to the degree of expectation accumulated along the cadence (which could be predicted both by syntactic and ASTM explanations). However, later responses were proportional to these stronger expectations, a modulation that was especially evident in musicians.

5.2.3 Study 3

5.2.3.1 ERPs to dissonant chords

We found that frequently-appearing dissonant endings elicited an EN with a fronto-central distribution. Importantly, the EN was undistinguishable between musicians and non-musicians. Dissonant endings also elicited a P3a (that was slightly larger in musicians than in non-musicians) and a P3b (that showed no difference between groups). Finally, frequent dissonant endings elicited an N5, that was larger in non-musicians than in musicians.

5.2.3.2 ERPs to deviant instruments

Deviant instruments elicited an N2b-P3a-P3b complex that was similar between groups. Like in Study 2, both groups showed a similar performance in the secondary task of counting deviant instruments that was above chance (although it was slightly higher in musicians).

5.2.4 Comparison between Studies 2 and 3

We compared the amplitudes of the responses to ending dissonant chords appearing in 25% of the sequences (Study 2) and 50% of the sequences (Study 3). Neither the EN, the P3a, the P3b nor the N5 were modulated by the proportion of appearance of dissonant endings. However, these analyses also revealed that the P3 was larger in musicians and the N5 was larger in non-musicians consistently in both studies.

In summary, our electrophysiological studies showed that unexpected tonal-syntactic and dissonant irregularities elicited responses related to the detection of acoustic deviance (such as those introduced by out-of-key tones in a tonal context and dissonance within a consonant context). Such irregularities also triggered responses related to the automatic switch of attention of the listeners, to structural context-updating and integration to the previous context. Musical training shaped how each neural response interacts with syntactic regularities. The responses of musicians were enhanced and more finely tuned to the detection of irregularities. But also, their responses reflected further processing of musical properties not evident in naïve listeners.

5.3 Relevance of our studies and implications for music cognition research

Our behavioural studies support that western listeners western listeners share the preference for consonant endings of chord cadences and an aversion toward strong and rough dissonance (Arthurs & Timmers, 2013), but that these are heightened in musicians (McDermott et al., 2010). Many studies evaluated the congruity of different types of chords with previous contexts (Brattico et al., 2010) very few studies directly compared consonance/preference evaluations of chords performing different functions within the context of a musical cadence (Arthurs & Timmers, 2013). These studies also showed that pleasurableness and correctness evaluations depended on, not only the type of chord, but the amount of expectation violated (or, in other words, the tonal function that such chord performed) and musicianship. The fact musicians (but not non-musicians) are able to judge Neapolitan chords at the third position as similarly correct as the tonic ending and also as more correct than Neapolitan endings supports a more robust stronger understanding of music theory, which allow them to discriminate between more subtle differences.

Our results on early responses to musical irregularities (such as the MMN/ERAN) extend our understanding of these components. For instance, to what extent they are plastic and modulated by musical experience and listening surroundings. A better knowledge of these neural responses has implications because it can facilitate cognitive treatments on individuals with amusia and other auditory-related diseases (Yu et al., 2015).

Previous electrophysiological research on music cognition and perception focused on the differential processing of various degrees of dissonance compared to consonance. However, even those that investigated consonance within a context, most of them introduced dissonance as a deviant in oddball paradigms (Brattico et al., 2009; Crespo-Bojorque et al., 2018). Moreover, some studies investigated the responses to different degrees of musical violations in structured, tonal context while other studies explored the effect of musicianship on the neural responses to tonal-syntactic violations. However, to the best of our knowledge, our studies are the first to directly address how acoustic dissonance is differently processed within a structured tonal context and to investigate the modulatory effect that musicianship might have.

More specifically, Study 1 replicated previous studies showing that the human brain of, at least, western listeners has the ability to automatically detect and categorize changes in a musical context such as deviations from tonality and consonance rules. Moreover, in Study 2, we challenged previous research in showing that, although dissonance is not categorized according to tonal hierarchy and expectations at early time windows (reflected in the EN/MMN), it is sorted according to these music-syntactic principles (or, at least, sensory surprise) at later stages of processing. This is reflected in the P3, that provides evidence for more conscious evaluation. Specifically, the P3b has been related to the conscious, deliberate detection of music irregularities that are task-relevant or involve decision (Koelsch et al., 2000; Regnault et al., 2001). To the best of our knowledge, ours are one of the first studies to show that the P3b can be elicited by irregular chords even when participants are neither aware about their appearance (at least at the beginning of the experiment) nor instructed to detect them. These results highlight that our brain automatically attend to and process disruptions of our internal environmental models if they are salient enough (which in turn renders dissonance very salient). Our studies also contribute to the extensive evidence indicating that the auditory system of musicians is more finely tuned than that of non-musicians to react to musical irregularities. For instance, we confirmed that their responses are significantly stronger in comparison to those of non-musicians. Not only that, in Study 2 we found that musician's show more efficient structural analysis of chord cadences even when facing disruptive violations such as dissonance. Previous studies using oddball paradigms showed an increased ability to categorize non-prototypical dissonant chords in musicians (Brattico et al., 2009). The present set of results advance this line by showing how dissonance is processed as a function of musical context.

One way to interpret the ERPs that we registered in the present set of studies is by acknowledging that they fit well within the framework of the adaptive orienting theory of error processing (Wessel & Aron, 2017). Unexpected perceptual events are followed by an automatic cascade of two sequential processes: an interruption of ongoing behaviour/cognitive processing, that frees up resources to identify the potentially important source of the unexpected and materializes in the attentional orienting (Näätänen et al., 2007). For unexpected events to be perceived as surprising and trigger an attention shift, they must exceed a certain threshold when compared with the established representation of the standard stimulus (which gives rise to the MMN). It is well known that the orienting of attention elicits the P3a, reflecting the transmission of the information to the frontal lobe (to make the event available to consciousness) and then, to temporal-parietal locations (where subsequent attentional resources promote memory operations) to elicit the P3b (Friedman et al., 2001; Wessel & Aron, 2017). The P3b reflects the mismatch with an internal model of the environment sustained consciously in working memory. In sum, incoming stimuli invoke top-down attention switching, and then bottom-up memory-driven operations to guide response organization and production (S. H. Patel & Azzam, 2005; Polich, 2007). Importantly, there is a dissociation between the MMN and the P3a, where the MMN can be elicited by both small and large deviants, but a significant P3a is elicited only by large deviants that exceed that threshold (Friedman et al., 2001; Nager et al., 2003). That we found a MMN in both groups but larger P3a and P3b in musicians suggests that Neapolitan and dissonant chords surpassed such threshold and engaged further processing, to a larger degree in musicians than in musicians (Nager et al., 2003).

5.4 Considerations on early responses: the debate on sensory and cognitive contributions

5.4.1 Study 1

The neural responses to tonal-harmonic violations have been extensively studied. However, few authors compared between different types of musical violations directly. For instance, acoustic violations elicit a larger EN than tonal violations (Koelsch et al., 2000), but whether tonal violations elicit larger responses than syntactic violations is unclear (Carrión & Bly, 2008; Koelsch et al., 2007).

Study 1 of the present dissertation is novel in directly comparing tonal with acoustic irregularities. In Manuscript 1, we claimed that the larger early negativity for dissonant clusters in comparison to Neapolitans represented that clusters introduced a larger violation of context, because they violate syntax, tonality and harmonicity, whereas Neapolitans "only" violate syntax and tonality. Previous studies showed that incongruous chords elicit an early negativity larger than ambiguous and congruous chords (Brattico et al., 2010). However, both ending chords of our study introduced out-of-key tones, which would confound acoustic and music-syntactic deviance. Various authors (Koelsch et al., 2007; Koelsch & Sammler, 2008) highlighted the difficulty in telling apart sensory from cognitive contributions because acoustic deviance usually co-occurs with music-syntactic deviance. However, most of them controlled for sensory contributions simply by counting the number of tones shared by target and context (Bigand & Pineau, 1997; A. D. Patel et al., 1998; Tekman & Bharucha, 1998), which does not define a compelling control according to Bigand et al. (2014). Bigand and colleagues (2014) suggested that a cognitive level of representation may not be indispensable for listeners to respond to some syntactic-like organizations in music. Precisely that study highlighted the potential of auditory short-term memory (ASTM) in explaining music tonal hierarchies more parsimoniously than cognitive accounts (Krumhansl & Kessler, 1982) and to further disentangle sensory from cognitive processes. To shed some light on this potential issue of our paradigms, we modelled our stimuli with the ASTM model (Leman, 2000; Bigand et al., 2014).

Regarding Neapolitan endings, the ASTM could account for the early responses registered, which would probably correspond to the elicitation of a MMN. That MMN would be simply reflective of the detection of an acoustic deviance from the memory of the tonality incited by the out-of-key tones of Neapolitan endings. In line with Bigand et al. (2014), the ERP results from Study 1 together with the ASTM model would contribute to the view that the early neural responses elicited by irregular chords with acoustic deviance can be parsimoniously explained by ASTM, while syntactic effects are not easily separable. There is agreement in that syntactic irregularities in the absence of physical deviance (such as ending in a supertonic that do not usually follow a dominant) can elicit neural responses studies (Koelsch et al., 2007; Koelsch & Sammler, 2008) that still cannot not be explained simply by ASTM. Likely, in parallel to an assessment of acoustic deviance, the brain may also evaluate the degree of syntactic deviance. Therefore, while ASTM is enough to explain the early negativity, we cannot rule out the possibility that received some contribution from cognitive processes of music-syntactic analysis.

Regarding dissonant endings, the ASTM account by itself could not explain the larger early negativity that they elicited. If the registered neural responses arose only from the deviance from the image of tonality stored in auditory memory, Neapolitans should have elicited a larger MMN than clusters, which was not the case. In fact, clusters introduced a smaller acoustic deviation than Neapolitans in relation to the auditory context built by the chord sequence (see Figure 13), because they had less out-of-key tones. In this case, ASTM failed to account for the larger amplitude of the early negativity. Although at first sight that could be interpreted as a cognitive account being plausible for our results (Bigand et al., 2014), we must keep in mind that roughness is not captured in the ASTM model. Moreover, from a cognitive point of view, clusters might have met the harmonic expectation of resolution better than Neapolitans because they were built over the tonic chord, which is the expected chord. Thus, the most likely explanation is that clusters may have met better the sensory and tonal expectations but because they were rough in a consonant context, they elicited a large phMMN. That idea is further supported by the fact that the MMN was larger in musicians than in non-musicians, as musicians are known to be very sensitive to roughness and to be able to quickly categorize dissonant chords at these early time windows (Brattico et al., 2009). Contrary to that view, previous studies suggested that dissonant chords in chord cadences are processed as strong violations *of the context* that elicit an ERAN and activation of the cortical network related to language processing (including the right frontal operculum, BA44). Importantly, these studies used dissonant clusters that did not share tones with the tonic and were built over the minor sixth. Thus, they likely confounded the irregularity introduced by harmonic relatedness with acoustic dissonance (Koelsch et al., 2000; Koelsch, Kasper, et al., 2004). Importantly, our results constitute a methodological novelty because by building dissonant chords over the tonic chord, we were able to disentangle acoustic dissonance from harmonic relatedness. Our results confirm that, even at early time windows, the brain may be able to assess these two features separately and that each process may elicit different (but sometimes overlapping) neural responses (Regnault et al., 2001).

5.4.2 Study 2

According to many studies, as a chord cadence advances, the degree of tonal-syntactic violation increases because the probability of each chord succession narrows, and the representation of tonality strengthens (Koelsch & Friederici, 2003). Tonality establishment may be ultimately rooted in ASTM, because the auditory memory for in-key chords grows more and more confident about the next chord being in-key and deviant chords become more surprising. We could also relate this effect with predictive coding, where our priors become stronger toward the end. In Study 2 tested to what extent the sensory processing of roughness interacts with syntactic appropriateness (that we operationalized as position within the cadence), with the aim to clarify whether the responses elicited by dissonance in Study 1 were influenced by hierarchic processing of musical context. The ASTM simulation of the experimental paradigm predicted only a slightly larger sensory surprise in the fifth position in comparison to the third position and only for short local decays. We found an EN that was not different between middle and ending positions. That result further supports that dissonant clusters in Study 1 recruited sensory processing of acoustic dissonance, that was not related to syntactic processing. That would have elicited a larger MMN in musicians because they are more sensitive to dissonance than non-musicians. Our results would imply that acoustic dissonance can be processed independently from syntactic deviance. As stated in Manuscript 2, this pattern of results in music resembles language in the sense that syntax and changes in the acoustic information (phonetic) can be processed separately. Our results might be relevant to the current debate about whether music and language share neural resources

(Bigand et al., 2014; Koelsch & Siebel, 2005; Steinbeis et al., 2006), because they would suggest that hierarchical structure is not unique to language, but rather a multidomain capacity of human cognition (Koelsch et al., 2013).

Another reason for the similarity of the early ERPs between positions could be that sequences were presented in blocks of 2 min of the same tonality and always started with the tonic. Likely, from the moment listeners confirmed that the first chord was the tonic they could build robust expectations for the resolution. Harmonic priming experiments show that hearing the tonic directly activates the "hierarchy of stability" (Bharucha & Krumhansl, 1983) and that listeners build accurate expectations of subsequent harmonies even when presented with a single chord (Bharucha & Stoeckig, 1986). Consequently, that the ASTM model did not show differences in across positions (only for very short decays) may be due to a longer-term accumulation of memory traces throughout each block. Thus, one way to look at our results is that tonality was similarly violated in both positions because listeners were habituated to the tonality. Moreover, musicians possess an enhanced representation of tonality and strong responses to musical violations even at initial positions of chord cadences (Guo & Koelsch, 2016; Koelsch, Schmidt, et al., 2002). The ERAN has been found to be enhanced in musicians irrespective of the position of irregularities in the cadence (Brattico et al., 2013). Therefore, the fact that the EN was indistinguishable between positions in Study 2 in musicians but tended to increase toward the end in non-musicians (see Figure 2 of our second paper) is consistent with a facilitated establishment of tonality, especially in trained listeners.

The findings of Study 2 contribute to the discussion around the functional significance of the ERAN. A key difference between Study 2 and previous literature (with similar paradigms presenting cadences in blocks of the same tonality) is that they found an increase of the EN toward the end for dissonant clusters (Koelsch et al., 2000). In that study, they named it ERAN and suggested that dissonant clusters introduced a stronger violation toward the end of the cadence because they violated the context more strongly. Importantly, these dissonant clusters were built either on the minor sixth or over the minor second of the tonality. Therefore, they were strongly non-syntactic wherever they appeared and likely were influenced by ASTM. In Study 2, we used clusters that were dissonant versions of the tonic triad, which is syntactically appropriate both at the third and ending position of a chord sequence (but because it is dissonant, we could say it is equally syntactically inappropriate). Likely, the brain reacted to roughness but, because it detected a chord that "mostly" matched the tonality expectation, it did not detect a strong violation of syntax and therefore did not elicit an ERAN. Thus, we provide evidence that tonality membership may weigh more than chord succession rules (syntax) in the processing of irregular chords and that the ERAN strongly relies on the introduction of tonal irregularities (which are grounded on ASTM) while the introduction of other acoustic irregularities (such as mistuning or roughness) do not have an additive effect (Koelsch et al., 2013).

5.4.3 Study 3

In Study 3 we aimed to further assess whether veridical knowledge extracted from experimental regularities about upcoming disruptive dissonance would impact the neural correlates of processing of such dissonance. Some ERP studies have shown that veridical knowledge in the form of familiarity (Besson & Faïta, 1995; Miranda & Ullman, 2007) and informed cueing modulate schematic expectations (Guo & Koelsch, 2015). Informed cueing modulates processes related to attention (and decreases the P3a and the P3b), but not partly automatic processes of sound analysis, such as those reflected in the ERAN or the MMN (Guo & Koelsch, 2015).

Only few neurophysiological studies directly assessed whether exposure to experimental regularities is enough to create veridical knowledge that can ultimately impact the processing of such an event (Guo & Koelsch, 2015). For instance, exposure to several hundreds of presentations of tonal-syntactic irregularities seem to reduce the amplitude of the ERAN (Koelsch et al., 2000; Koelsch & Sammler, 2008), maybe because the participants' initial representation of what was irregular was modified in an implicit way during the course of the experiment. Even in previously unheard musical scales, when presenting deviants equiprobably, the (Bigand et al., 2014). In other words, repeated stimulus exposure can lead to habituation as a result of a continuous comparison of the constructed representation to incoming information until they match sufficiently (Friedman et al., 2001). However, even under these conditions, the ERAN does not disappear, suggesting that it is remarkably stable (Bharucha & Stoeckig, 1986; Guo & Koelsch, 2015). Other studies in which participants' expectations to regular chords were boosted through training showed that tonal and syntactic violations elicit an enhancement of the P3b, but not the ERAN (Carrión & Bly, 2008). The predictive coding framework postulates that evoked responses correspond to prediction error that is explained away during perception and is suppressed during perceptual learning. With the aim to minimize free energy, the repetition of standards renders the suppression of prediction errors more efficient which leads to a reduction in ERPs and only un-learned stimuli elicit a mismatch response (Garrido et al., 2009). In that context, perceptual learning of irregularities (e.g. in the form of repetition) can progressively reduce prediction error, which is manifested in an attenuation of the MMN (Garrido et al., 2009; Koelsch et al., 2019). Together, previous research suggests that musical expectations may be modulated by veridical expectations to some degree and may even lead to a facilitated processing of irregular events, although they cannot be easily overridden. In line with this evidence, in Study 3 we showed that, in terms of ASTM, frequently appearing dissonant endings introduced a smaller violation of context than when they appeared rarely (like those on Study 2). However, they never become less violating than the tonic. We thus expected that an increased predictability of dissonant endings would elicit smaller neural responses. Nevertheless, we observed that the early negativity to frequently appearing ending clusters were not reduced in comparison to rarely appearing ending clusters. This result is surprising because, even if veridical expectations do not modulate schematic expectations, internal auditory models should be adjusted during the repetition of irregular events. Most likely, the registered early responses reflected the acoustic deviance introduced by dissonant chords alternated with consonant chords, with independence of schematic and veridical expectations. Both in Study 2 and Study 3, dissonant chords represented 10% of all the chords presented, which does not render dissonant chords frequent enough and mimics an oddball paradigm. In line with the outcome of Study 3, previous research showed that dissonant chords elicit a large MMN in comparison to consonant chords in oddball paradigms (Brattico et al., 2009; Crespo-Bojorque et al., 2018).

5.5 Considerations on late responses: the P3a and P3b

5.5.1 Study 1

In Study 1 we observed similar P3 components between Neapolitan and dissonant endings, which was an interesting result. For instance, non-identifiable sounds are known to elicit a larger P3a than identifiable sounds (Friedman et al., 2001) and clusters are known to be difficult to categorize within any known musical category. Also, tonal irregularities elicit a P3b after the ERAN while syntactic irregularities elicit an ERAN but not a P3b. This suggests that even non-musicians process syntactic violations as less regular in a harmonic sense (ERAN), but still as an accepted resolution at the end of a harmonic progression (P3b). In contrast, tonal violations are consistently processed as more unexpected (Carrión & Bly, 2008). Interestingly, according to Carrión & Bly (2008), tonal and dissonant deviations require the same degree of restructuring because they violate the *actual* elements within the musical vocabulary, as opposed to timbre deviations that violate the physical characteristics over which the vocabulary is built (Carrión & Bly, 2008). Therefore, that we found no difference in the P3 between types of chords would suggest that, despite the fact that they were highly deviant at the sensory level (at least for musicians), Neapolitan and dissonant chords required the same degree of restructuring and recruited attentional resources to a similar degree (with differences between groups, see section *5.10 Considerations on the effect of musicianship*).

5.5.2 Study 2

To the best of our knowledge, while the P3a can be found in auditory oddball tasks (Polich, 2007), the P3b is not (Carrión & Bly, 2008; Crespo-Bojorque et al., 2018). For instance, tonal violations, but not syntactic violations, elicit a P3b (Carrión & Bly, 2008). In our studies we introduced dissonant chords in structured musical contexts, which likely affected later cognitive processes originating from higher levels in the auditory processing hierarchy (Koelsch et al., 2019). That would suggest that participants did extract the regularities of the sequences, in addition to reacting to acoustic deviance from a consonant context (as seen in the MMN). We moreover observed a difference of amplitude of the P3a and P3b between positions. Even if there was some sort of long-term accumulation of memory traces and habituation to tonality (as the ASTM simulation could suggest) due to sequences always starting with the tonic and being presented in long blocks of the same tonality, our results indicate that later cognitive processes analysed the sound relationships sequence by sequence. Note that we found a difference when comparing across positions but not when comparing between Neapolitan and dissonant chords. While tonal and dissonant deviations in Study 1 would require the same degree of restructuring (Carrión & Bly, 2008), the difference between positions in the P3b would reflect the dissimilarity between the target chord and the established harmonic context and, ultimately, the higher degree of unexpectedness (Guo & Koelsch, 2015; Janata, 1995; Koelsch et al., 2000; A. D. Patel et al., 1998). This explanation fits with the "context-updating" model

(Donchin & Coles, 1988; Polich, 2007), according to which the amplitude of the P300 depends on the processing power required to resolve the structural discrepancies between deviant chords and the preceding musical context (Donchin & Coles, 1988). In the context of Study 2, dissonant chords would have elicited a large P3 because they required increased processing power to resolve the structural discrepancies. On top of that, the P3 modulation of the amplitude across positions would confirm that the more we advance in a chord sequence, the larger is the discrepancy that dissonance introduces and the larger is the power required to resolve that discrepancy.

Another consideration about the difference in the P3a and P3b amplitude between positions comes from the timing considerations while processing music. We believe that, after some repetition, listeners learned that sequences had five chords and would continue after dissonant chords at the third position. At that point, participants would be expecting the sequence to continue, and the brain would be prepared to process other chords that are going to occur after, consequently recruiting less attentional resources than for ending chords (which could be reflected in a smaller P3 at the third position).

5.5.3 Study 3

Study 2 showed that participants extracted regularities at the sequence level (referred to as schematic regularities, as shown in the difference between positions) but in Study 3 we found that the P3 was not reduced as a function of a higher frequency of appearance of clusters, as if participants did not extract predictions based on the experimental regularities. Previous research has shown that the repeated presentation of irregular targets reduces the P3 (Friedman et al., 2001; S. H. Patel & Azzam, 2005), to the point that, when errors become the more likely outcome, correct trials start eliciting a P3 (Carrión & Bly, 2008; Wessel & Aron, 2017). Evidence of music research show that, after entrainment to regular sequences, both syntactic and tonal violations elicit a P3b, but without it, only tonal violations do (Carrión and Bly, 2008). Accordingly, we could have expected a reduction of the P3b in our study. However, we did not find a significant reduction of neither the P3a and nor the P3b in Study 3. This could suggest that schematic expectations are not easily weakened (Bharucha & Stoeckig, 1987; Guo & Koelsch, 2016), at least not by familiarization through repeated processing of less-expected musical structures. Indeed, behavioural research showed that familiarization or habituation to unexpected chords does not facilitate their processing (Justus & Bharucha, 2001; Tillmann & Bigand, 2010). This might be especially true if such structures include dissonance, which introduces a psychoacoustical disruption that is difficult to habituate to. However, it remained unclear whether these results were actually due to dissonance being a very disruptive violation that can't become less surprising through exposure or to listeners having difficulty extracting higher-order experimental regularities.

Another interpretation would be that more instances of dissonant chords may be needed for the listeners in our studies to habituate. Previous studies have found systematic decline of the P3a and P3b through the course of the experiment with the repetition of harmonically irregular chords, especially in musicians (Guo & Koelsch, 2015; Seppänen et al., 2012). In these studies, the P3b likely decreased because the structural reanalysis of irregular events was facilitated by the repeated presentation of these events (Guo & Koelsch, 2015).

5.6 Other components to consider

According to the theory of error processing, perceptually novel events elicit the N2b/P3a complex over the medial prefrontal cortex. It is possible that dissonant chords in our studies elicited a novelty N2b/P3a complex, which likely overlapped with the MMN related to roughness. The N2b has central distribution (which matches the distribution of the early negativity registered in studies 2 and 3) and reflects the deviation from a template or mentally stored expectation, while the MMN reflects the departure from a collection of standard stimuli (S. H. Patel & Azzam, 2005). Taking into consideration the presence of the N2b/P3b complex is relevant for our studies because it is usually reflective of subjects selectively attending to deviations (Wessel & Aron, 2017) generating a response (Koelsch et al., 2000, 2001; Koelsch, Schoger, et al., 2002). Both the N2b and the P3b have been found for task-irrelevant deviants that are salient enough, which suggests that some stimuli's properties can lead to conscious registration of task-irrelevant deviance (Koelsch et al., 2007; Schröger & Wolff, 1998). Together, the presence of an N2b/P3a complex and the P3b in our studies could suggest that dissonant chords were salient enough to be registered consciously even if they were task irrelevant (Koelsch & Siebel, 2005; A. D. Patel et al., 1998). This is consistent with previous research by Koeslch et al. (2000) who already found that, in non-musicians, Neapolitans elicit a P3a while dissonant clusters elicit an N2b-P3a-P3b complex, reflecting attentional and decisional processes (as participants were tempted to respond to clusters). Considering that the main task in our studies was relatively easy, participants could have been performing a dual task by attending the deviant instruments while still consciously evaluating the musical irregularities (because they never co-occurred). Also, the presence of the N2b/P3 complex would argue in favour of dissonant chords being perceived as "novel" even if they are presented at different positions or presented very frequently as endings.

5.7 Considerations on late responses: the N5

Not many studies have investigated the N5 and even less have addressed how it behaves as a function of musicianship. First described in 2000 (Koelsch et al., 2000) the N5 is taken as reflective of harmonic integration or a modification of listeners' hierarchy of harmonic stability (Koelsch, Schmidt, et al., 2002). In other words, the N5 reflects the effort invested in integrating an irregularity into the previous context build up. The processes of harmonic integration resemble the processes of semantic integration during language perception (reflected in the N400). Therefore, the N5 is taken to reflect processing of musical meaning (at least partly) because it appears for irregular chord functions and deceptive cadences, which are prominent elements of tonal music that are used by composers as a means of expression (Koelsch, Kasper, et al., 2004; Koelsch & Siebel, 2005). The emergence of the N5 has been linked to Wernicke's area, which is known to process lexical-semantic aspects of linguistic stimuli. That Wernicke's area is activated for tonal modulations (that can only be differentiated by the application of implicit knowledge of the rules of harmony) and clusters reflects an attempt to find meaning in them (Koelsch, Wittfoth, et al., 2004).

However, the exact relation between N5 and processing of musical meaning remains to be specified (Koelsch et al., 2007). One of the aims of this dissertation was to assess whether musicians and non-musicians would show different approaches towards the integration of unexpected musically unacceptable events. For instance, because dissonant chords are musically unacceptable in terms of consonance, it would make sense to expect a musician to reflect show a larger response. However, we could also argue that precisely because dissonant chords cannot be categorized into any common musical category, a trained brain would likely not bother to invest any resource. The results of our studies did not answer that question easily; on the contrary, they raised further conundrums.

In our studies, there were two interactions of interest: musicianship and amount of expectation (position). Regarding the question of the effect musicianship on the N5, very few studies registered the N5 in musicians and non-musicians (Koelsch et al., 2007; Koelsch, Schmidt, et al., 2002). These studies consistently reported that musicians and non-musicians listening to Neapolitans, double dominants or supertonics show a similar N5. In our studies the N5 was always larger for non-musicians, very likely due to an overlapping P3. For that reason, we cannot draw a direct conclusion around the effect of musicianship on the N5. However, the interaction of the N5 with the amount of expectation provided more information. Previous studies showed that Neapolitan chords elicit larger N5 as endings than at the middle of the cadence (Koelsch et al., 2000). Also, the amplitude of the N5 is modulated by interacting global and local hierarchical harmonic relationships: chords that are harmonically irregular at the local level elicit an N5 only when related to the global context, while chords harmonically irregular at the global level elicit an N5 only when related to the local context (Zhang et al., 2018). However, in these studies, the stimuli were task irrelevant or even ignored (Koelsch, Schoger, et al., 2002). Because of the overlapping P3 in our studies, we do not have enough evidence to argue about the effect of harmonic relatedness, but it would fit well with our results: similarly to Koelsch et al. (2002), in Study 2 ending clusters elicited a larger P3 and the N5 was only slightly visible. Meanwhile, at the third position the P3 was smaller and therefore the N5 was significant. The overlap between the P3 and the N5 was further modulated by musicianship because that effect was present in non-musicians but not in musicians. In musicians the P3 was that it masked the N5 even at the third position. That the N5 that we registered was strong enough to reach statistical significance in spite of overlapping with the P3 argues in favour of the large effort that had to be invested by our listeners to process dissonant chords, because in previous studies if the P3 was present, it completely masked the N5.

We must highlight that, in Study 1, we did not observe an N5, maybe because irregular endings (Neapolitans and dissonant clusters) taken together appeared in 50% of the sequences. One speculative explanation could be that the integration of Neapolitan and dissonant endings became easier through exposure, as the N5 has been shown to decrease in previous studies where cadences end in an irregular ending very frequently (Koelsch et al., 2000). In line with that idea, in Study 3 maybe we registered an N5 although we presented 50% of irregular endings because they were all dissonant clusters, whose integration can hardly become easier. These results would suggest that an irregular event must also be rare to be processed as semantically irregular (unless it is strongly disruptive, like dissonant clusters), which draws an interesting parallel with language. Another explanation for the inconsistent results regarding the N5 across our studies could be the lack of statistical power, as the effect has been shown to be small or not reach significance for studies with few participants (Koelsch et al., 2007; Koelsch & Jentschke, 2009).

In sum, that we found the N5 in our studies provide further evidence that western listeners invest neural resources to integrate even strongly disruptive acoustic irregularities in order to extract harmonic meaning of music. Ours studies directly assess the effect of musicianship in the process of integrating dissonant chords in a musical context, although the effect on the N5 was likely masked by the P3. We moreover built on previous studies by showing that the N5 is reflective of the amount of expectation violated even for dissonant chords. Finally, we showed that, while the integration of syntactic irregularities can be facilitated through exposure, it is unlikely for strongly disruptive dissonant chords, similarly to the earlier ERP responses.

5.8 Considerations on the processing of consonance-dissonance

As stated early in this dissertation, the perception of dissonance has been long studied from the perspective of physics, ethnomusicology, psychoacoustics among other disciplines. The current state of literature agrees in that the perception of consonance and dissonance may be grounded in psychoacoustics but also mediated by culturally acquired preferences and further enhanced by musical training (Popescu et al., 2019). While all cultures hear roughness, it is associated with divergent aesthetic interpretations. At least for western listeners, the more accepted hypothesis for explaining the preference for consonance is harmonicity, or how harmonic are the aggregate frequency spectra of tone combinations (McDermott et al., 2010), although it is not applicable in all cultures or types of listeners, such as those with amusia (Cousineau et al., 2012). However, most studies focusing on studying dissonance focused on isolated chords, which is an over-simplification of music perception. For instance, it has not been studied whether harmonicity contributes to aesthetic responses to chord progressions or melodies, via integration of frequency information over time (McDermott et al., 2010).

Theoretical evidence highlights the importance of investigating dissonance within more realistic musical contexts. Currently, the quality of a chord is defined based on its role within a harmonic context, its acoustical/sensory consonance and the kind of melodic organization, instead of simply its fundamental frequency ratios. Critically, one same chord might be considered consonant or dissonant (and more or less pleasant) depending on the harmonic context in which it is embedded, independently of the presence of beating (Parncutt, 1989; Plantinga & Trehub, 2014; Steinbeis et al., 2006). A very dissonant chord can have a stable and important tonal function (McDermott & Oxenham, 2008) – e.g. the minor major seventh that ends some jazz pieces – while a consonant chord can have an unstable tonal function – e.g. a modulating dominant chord (Bigand et al., 1996). In the same line, chords are rated as more consonant when placed in a traditional harmonic progression than when in a non-traditional harmonic progression (Roberts, 1983). Moreover, major and minor triads are rated similarly when placed either in random sequences or chord cadences, while augmented and diminished triads are judged more consonant when presented alone (Arthurs & Timmers, 2013). It is precisely this interaction between musical context and perception of consonance and pleasurableness what inspired our research. As reviewed above, we have identified some of the key neural processes responsive to dissonance and observed their modulation by amount of expectation and musicianship. Our ERP results provided a further understanding of the interplay between musical expectations and acoustic dissonance observed in responses such as the P3a and P3b, suggesting that dissonance can be processed as more or less surprising as a function of context. This has not been observed in studies where dissonance is presented in isolation and has a nice fit with current musical practice. However, our ERP results alone do not provide enough information about the conscious perception of these musical irregularities.

5.9 Considerations on the relation between neural responses and behaviour

Our behavioural studies are coherent with previous studies regarding the preference of trained and untrained western listeners for consonant endings of chord cadences and an aversion toward strong and rough dissonance (Arthurs & Timmers, 2013). We further showed that both these preference and aversion are heightened in musicians (McDermott et al., 2010). The cognitive reactions to musical incongruities have been extensively measured by asking subjects to rate the congruity of chords with previous contexts (Brattico et al., 2010) and by evaluating different types of chords. For instance, evaluations of chord correctness can deviate from music theory and be independent from pleasurableness ratings, especially for ambiguous chords. However, very few studies directly compared consonance/preference evaluations of chords performing different functions within the context of a musical (Arthurs & Timmers, 2013). Most of these used rating in a binary forced choice, while we introduced the methodological novelty of rating chords' correctness/pleasurableness in a sliding scale, which is more sensitive to subtle changes in judgements.

Our studies showed that pleasurableness and correctness evaluations depended on, not only the type of chord, but the amount of expectation violated (or, in other words, the tonal function that such chord performed) and musicianship. These results are a good example of how one same chord might be considered more or less pleasant depending on the harmonic context and the listener (Parncutt, 1989; Plantinga & Trehub, 2014; Steinbeis et al., 2006). For instance, non-musicians have been shown to differentially rate the consonance and pleasantness of augmented and diminished triads performing different more or less common harmonic functions in chord sequences (Arthurs & Timmers, 2013). We showed that musicians (but not non-musicians) rated Neapolitan chords at the third position as similarly correct as the tonic ending and also as more correct (but not more pleasurable) than Neapolitan endings. Likely, a stronger understanding of music theory reflects in musicians' conscious perception and preference evaluations, which allow them to discriminate between more subtle differences (such as between Neapolitan chords at different positions). Another example of this would be that, while musicians rated dissonant chords worse than non-musicians in both conditions (which confirms that they have a stronger negative response towards dissonance; Brattico et al., 2010), they still rated them as more pleasurable at the third position than at the fifth.

Importantly, our behavioural results are consistent with evidence of affective and cognitive processing being distinct. That idea is further supported by electrophysiological and lesion studies showing that patients with disrupted perceptual music abilities keep intact music-induced emotion. There is even the hypothesis that affective judgement processes precede cognitive ones and require only primitive stimulus analysis (Brattico et al., 2010). While these two processes may be different at the neurophysiological level, that does not necessarily lead to divergent conscious evaluations, as for non-musicians, pleasantness and correctness are barely distinguishable at the attentive level. Our studies further contribute to research showing that musicians make a clear distinction between conscious evaluations of pleasantness and correctness (Lahdelma & Eerola, 2016; Plomp & Levelt, 1965; Popescu et al., 2019; Roberts, 1983). Again, musicians' stronger understanding of music theory and ability to extract information from dimensions such as structural cues and culturally or autobiographically relevant connotations (Lahdelma & Eerola, 2016), may affect their evaluations and preferences. In turn, these are further affected by musical style (and the aesthetic ideals associated with it) because each style employs different mechanisms to trade-off consonance/expectedness for pleasantness.

An interesting topic of discussion is the relationship between conscious evaluations and the neural responses for the same evaluated elements evaluations. For instance, previous research found distinct neural processes preceding cognitive and affective judgements, although they partially overlapped during initial sensory processing (Brattico et al., 2010). In our studies we do not have direct evidence of the relationship between cognitive/affective ratings and their neural responses. Nevertheless, we found various interesting dissociations between neural responses and behavioural evaluations of musical irregularities that are modulated by musicianship (note that we can only draw indirect relations because we obtained behavioural evidence in an active task [rating chords], while neural evidence was obtained in a passive listening task [detecting deviant instruments]). For instance, higher ratings of Neapolitan endings in comparison to intermediate Neapolitans are predicted by larger early ERPs (Koelsch et al., 2000; Koelsch, Schmidt, et al., 2002), a difference that is further increased in musicians. This indicates that such discrimination may be rooted to some extent in syntactic-acoustic attributes. Indeed, the application of musical knowledge based on experience can lead to different results. For instance, musicians and non-musicians' neural responses to Neapolitans were similar, despite musicians rated these closures higher than non-musicians. Therefore, musicians must have applied their explicit knowledge of musical rules. We found an opposite example in the fact that both groups rated dissonant endings as highly inappropriate although the neural responses of musicians were larger and modulated by the amount of expectation violation. This implies that, despite increased auditory and attentional responses toward dissonant chords in musicians, their ratings do not simply rely on the physical properties and music-syntactic relations of target chords (reflected in early neural responses) or on the degree of sensory surprise (reflected in late neural responses). Rather, these ratings seem to be based on top-down processed such as the application of a conscious culturally acquired negative association with dissonance. To explain the differences between untrained and trained listeners, it is possible that they rely on different features to elaborate their decisions (Kung et al., 2014). This may be further influenced by familiarity as sometimes, initial interest in a novel stimulus may be replaced by negative evaluations after longer exposure (Plantinga & Trehub, 2014).

5.10 Considerations on the effect of musicianship

Our studies advanced on previous research showing stronger neural reactions to musical irregularities (tonal or sensory dissonant). To summarize, the well-known sensitivity and ability to quickly categorize to dissonance of musicians were reflected in increased early negativities in comparison to non-musicians (Brattico et al., 2009). While increased responses of musicians have been reported in simple oddball paradigms, the effect of musicianship on neural responses within musical structured contexts has not been thoroughly addressed. Consistent with previous studies, we found an enhancement in musicians of attentional and structural analyses of chord cadences (even when facing disruptive violations such as dissonance), reflected in the increased P3a and P3b (Miranda & Ullman, 2007; Nager et al., 2003; Tervaniemi et al., 2005) and the modulation of these components by amount of expectation. The larger P3 component in musicians suggests that they are in possession of a more accurate and detailed representations (Proverbio et al., 2016). Interestingly, P300 potentials have been claimed to be influenced more strongly by musical expertise than the processes of music-syntactic analysis (as reflected in the ERAN; Guo & Koelsch, 2016). Regarding processes of harmonic integration, more work is needed to precisely understand how they were modulated by musicianship (because the N5 was masked by the large P3 due to the sensory surprise of dissonance). Taken together, our studies we reported robust evidence on the enhancement of neural responses by musicianship. However, there were some issues that deserve further discussion.

Regarding early responses, in our first paper (Study 1) we claimed that "the MMN was not different for timbre deviants between musicians and non-musicians, but it is for dissonant chords" and, based on that, we suggested that the larger EN was not due to enhanced general auditory sensitivity but rather to more elaborated music-syntactic processing. However, timbre deviance is a distinction that is salient to all listeners while dissonance is a little bit more subtle distinction. While detecting the sensory novelty of dissonance is a quite universal process to all western listeners, non-musicians may have trouble categorizing it into known musical categories. In contrast, dissonance may be, in fact, acoustically more salient to musicians. Another interesting observation is that we found that a larger EN for infrequent dissonant endings (Study 2) in musicians while frequent dissonant endings (Study 3) elicited similar responses across groups. While the interaction between Study and group was not significant (maybe due to individual peak latencies or small effect sizes), the lack of difference between levels of musical expertise could imply that musicians' responses decreased because they may easily become able to anticipate clusters that always happen in the fifth position. All in all, there might be an interaction between musicians being better in discriminating chords, the decreased predictability of standards (tonic endings) and the increased predictability of deviants that would be worth investigating thoroughly in the future.

One of our aims when studying the effect of musical training on the neural responses to music irregularities was to better understand the implications of long-term experience in our brain functioning. Importantly, music processing in musicians is influenced by the interplay between cultural exposure and the peculiarities of the training received. There is evidence of multi-musicality in musicians that depends on the degree of exposure to each musical system. For instance, folk musicians mix musical western and non-western cultures systems to various degrees (for instance, Finnish folk musicians). The resemblance of the neural responses of folk musicians with western classical musicians is proportional with the degree of exposure to western musical culture. Those less exposed to Neapolitans at the third position acting as subdominant (which is theoretically correct) show a weaker ERAN than folk musicians exposed to western music (Tervaniemi et al., 2012). Even within the same western culture there are differences depending on the style of training: selective attention capabilities are improved in all musicians, but conductors have an exquisite auditory spatial resolution and selectivity. While attended targets elicit an increase in the P3 in all musicians, unattended unexpected events only elicit a P3 in conductors, which could be related to decision confidence in detecting target stimuli (Nager et al., 2003).

Moreover, Neapolitans elicit larger ERAN and P3b in jazz musicians, suggesting that they have increased perceptual sensitivity and higher engagement. Meanwhile, the same chords elicit a persistent late positivity in classical musicians, because they may see Neapolitans as an error that requires subsequent further cognitive analysis. In classical training, musicians need to recognize errors so they can avoid the same mistake in subsequent performances. That jazz musicians recover quickly from these events suggests that they possess increased flexibility to switch to a different cognitive strategy immediately after unexpected chords, because improvisation occurs in real time (Przysinda et al., 2017). To the best of our knowledge, no study has yet investigated the neural responses to dissonance depending on the different types of musical training. Musicians participating in our studies were either trained in classical or jazz/modern music (to similar proportion) and we did not observe any obvious difference in their neural responses toward dissonance. A logical assumption would be that rough dissonance might represent an error regardless the type of training received. Because the experiments were not aimed to test for differences between types of trainings and we may lack statistical power, this should be addressed in future studies.

5.11 Future research

5.11.1 Methodological caveats and shortcomings

In Study 3, there was a matter that remained unresolved: whether the lack of effect of increased predictability of dissonant chords was due to psychoacoustical difficulties for processing roughness (which prevents listeners from habituating to it) or whether it was more related to a difficulty in extracting the experimental regularities. To investigate this issue an option could be to increase even more the probability of dissonant chords so that the alternation of consonant/dissonant chords do not resemble an oddball paradigm and then compare these results with a paradigm with the same proportions but with syntactic deviation (proven to facilitate habituation of neural responses). This latter paradigm is already being tested in our lab, although it was not included in the present dissertation. An additional manipulation could be to present dissonant chords frequently but randomly changing their position within the chord cadence. If we observed a similar response to Study 3, it would suggest that listeners did not apply expectations to dissonance at the level of cadence.

Moreover, in our studies, out-of-key notes were likely confounded with dissonance and syntactic appropriateness (Koelsch et al., 2007; Koelsch & Sammler, 2008) which left open issues about the extent to which each feature of the chords was responsible for the observed responses (although ASTM was very informative). We consider that we minimized syntactic deviation of dissonant clusters because they were versions of the tonic and were equally acceptable at both positions. However, to validate that claim, future research could separately manipulate the expectancy of target chords at middle and ending positions and their dissonance (similarly to (Regnault et al., 2001). Regnault and colleagues cleverly manipulated expectancy and dissonance separately but they did not test different positions. In such paradigm we could discern to what extent sensory dissonance is processed separately from syntactic expectations, building on the debate around the ability of the human brain to process structure independently from deviations of the auditory input. To control for acoustic deviance, future studies should make sure that target chords do not introduce tonal deviation (for instance, by using dissonant versions of syntactically irregular chords such as the supertonic) or, at least, make sure that any deviant tone already appeared in the chord sequence (Pearce, 2005, 2018; Virtala et al., 2011b). In fact, the best option would be to simulate both ASTM sensory expectations and syntactic expectations, for instance, with the model IDyOM, (Pearce, 2005, 2018) and then compare how they perform. Simulating syntactic expectations would be key to verify that experimental manipulations (such as changing the position of a chord within a cadence) actually affects syntactic expectations, rather than assuming it based on music theory.

5.11.2 Cross-cultural studies

In this line, scientists, ethnomusicologists and composers are still debating around the biological versus cultural origins of musical preferences. Studies using artificial musical contexts (Loui et al., 2009) show that western listeners can extract music regularities from any musical context. Such internalization may emerge from ASTM, a very fundamental mechanism of temporal integration. Thus, that ability should be transferable to any situation where listeners face a new musical idiom. Likely, the same processes (perceptual, cognitive, memory, expectations) will apply also to the processing of other musical systems (which define a further interesting challenge of cognitive accounts of music perception; Tillmann et al., 2014). Very few studies investigated the neural responses of expectancy violations in cross-cultural contexts and, to the best of our knowledge no study has done so with chord sequences or with dissonance in structured tonal contexts. For our line of research, cross-cultural ERP studies could be very enlightening. For instance, out-of-culture melodic incongruencies are usually rated as less congruous (Demorest & Osterhout, 2012) and elicit less robust P600 and P300 (Bischoff Renninger et al., 2006; Demorest & Osterhout, 2012). These evaluations (and probably the neural responses) of out-of-culture irregularities rely on the listener's own culture tonal knowledge to evaluate musical congruency and depend on the overlap between cultures (Bischoff Renninger et al., 2006; Curtis & Bharucha, 2009). It is unknown to what extent the accumulation of expectation through a musical cadence works the same way at the electrophysiological level between musical cultures, which could be investigated with our paradigms. We also wonder whether the sensory sensitivity toward dissonance can be modulated not only by explicit musical training but also by the musical context in which we grow up. For instance, people with no exposure to western music (McDermott et al., 2016) or from cultures where dissonance is used more freely (Maher, 1976) show no aversion toward inharmonicity, but still remains to be investigated whether that affects the underlying neural responses.

5.11.3 Effects of musical experience and musical training

Drawing conclusions about the origin of the differences between musicians and non-musicians is always tricky because of the caveats intrinsic to experimental settings, such as tasks with instructions including technical musical terms not understood by non-musicians (Bigand, 2006) but particularly because of the uncertainty as to whether any brain difference actually results from training or corresponds to genetic differences that predispose some individuals to become musicians (Bigand, 2006). There are methodological options to disentangle these confounding factors: directly studying the auditory nerve (Bidelman, 2013), performing longitudinal studies (Peretz & Zatorre, 2005) or studying musicians with different types of training, such as jazz and classical - who respond differently to tonal-syntactic violations (Przysinda et al., 2017). An interesting direction for future studies would be to study the effect of different types of musical training in the neural responses of processing strong dissonance, which is considered a universal musical deviant. Moreover, we have investigated the effect of musical training on the adaptation to strong dissonance but not to milder music-syntactic adaptations. The ERPs of naïve listeners habituate to a high proportion of Neapolitans (Koelsch et al., 2000) and one could explore whether musicians would adapt more easily than non-musicians or their stronger expectations would prevent them to do so. Given that dissonance may be hard to habituate to, frequent supertonics might be more well suited for investigating this issue without the confound of sensory processing of dissonance.

4.11.4 Music and language

Music and language have been argued to share the processing of structural relations, as supported by the evidence that syntactic violations interfere with the processing of sung phonemes, linguistic syntax and semantics (Fedorenko et al., 2009; Koelsch & Siebel, 2005; Regnault et al., 2001; Slevc et al., 2009; Steinbeis & Koelsch, 2008). The results from Study 2 could further support the *shared syntactic integration* *resource hypothesis* (Patel 2003, 2008) by showing that syntax and acoustic deviance are processed separately in our musical stimuli, like in language (Bigand et al., 2014; Koelsch & Siebel, 2005; Steinbeis et al., 2006). Future research should tackle the interfering effects between the processing of linguistic sentences and musical sequences. Note that strongly disruptive music violations may simply disturb attentional processes, which can eventually result in interference of linguistic processing but do not necessarily imply resource sharing. Thus, future studies investigating such interaction should control for acoustic disruption (Bigand et al., 2014)

5.12 Conclusion

In a set of studies, the present dissertation directly addressed the interaction between acoustic dissonance in musical structured tonal context, schematic (syntactic) and veridical (experimental) expectations in the form of repeated exposure and the modulatory effect of musicianship. We also assessed to that extent musicianship shapes conscious affective and cognitive judgements of musical expectedness and whether these dissociated. We provided evidence that day-to-day exposure to music is enough for western listeners to automatically detect changes in a musical context (such as deviations from tonality and consonance) at early stages of processing. We have shed more light on the interplay between musical expectations, syntactic appropriateness and acoustic dissonance and showed that the brain may be able to assess them independently, even at these early stages of processing. Also, tonal and (especially) dissonant deviations are disruptive enough for western listeners (even when they are naïve to their appearance) to attract their attention, to be registered consciously and to engage further context-updating processing. Importantly, at these later stages of processing, acoustic deviations may be sorted and processed according to the principles of tonal expectations (even though these might be rooted in sensory surprise and ASTM). Also, western listeners invest neural resources in integrating (even) strongly disruptive acoustic irregularities to try to extract harmonic meaning of music. In all the circumstances described, musicians showed stronger and more fine-tuned neural responses, both at the early sensory and later cognitive stages of processing. On the other hand, we also showed that the plasticity of schematic expectations may depend on the type of disruption, because highly disruptive acoustic deviance such as dissonance does not induce the habituation in listeners that one would expect after familiarization through repeated exposure. On a more methodological tone, auditory short-term memory simulations allowed us to better understand the sensory contributions to the complex interactions encapsulated in our studies and the theoretical implications of the subsequent results, such as to what extent musical syntax is grounded on peculiarities of the auditory functioning. Our behavioural results again confirmed the heightened preference and aversion present in musicians. More interestingly, the conscious evaluation of pleasurableness and correctness of unexpected musical events seemed to depend on the subsequent amount of expectation and, of course, of musicianship (for whom these two concepts can be independently assessed). At the same time, these evaluations sometimes dissociated from the underlying neural responses, revealing the application the conscious application of culturally acquired musical associations. Together, the set of studies of the present dissertation further demonstrate how plastic can music perception and cognition be.

6. REFERENCES

Arthurs, Y., & Timmers, R. (2013). Evaluating the Consonance and Pleasantness of Triads in Different Musical Contexts. *3rd International Conference on Music & Emotion (ICME3)*, *June*, 1–7.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear Mixed-Effects Models Using S4 Classes (R Package Version 0.999999-0). http://cran.r-project.org/web/packages/lme4/index.html

Berlyne, D. E. (1971). Aesthetics and Psychobiology. *Appleton-Century-Crofts,* 336p.

Besson, M., & Faïta, F. (1995). An Event-Related Potential (ERP) Study of Musical Expectancy: Comparison of Musicians With Nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1278–1296. https://doi.org/10.1037/0096-1523.21.6.1278

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*(4), 485–518. https://doi.org/10.1016/0010-0285(84)90018-5

Bharucha, J. J., & Stoeckig, K. (1986). Reaction Time and Musical Expectancy. Priming of Chords. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(4), 403–410. https://doi.org/10.1037/0096-1523.12.4.403

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception &*

Psychophysics 1987 41:6, *41*(6), 519–524. https://doi.org/10.3758/BF03210486

Bharucha, & Krumhansl. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, *13*, 53–102. https://doi.org/10.1016/0010-0277(83)90003-3

Bidelman, G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology*, *4*, 1–13. https://doi.org/10.3389/fpsyg.2013.00264

Bidelman, G. M., & Krishnan, A. (2009). Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem. *Journal of Neuroscience*, *29*(42), 13165–13171. https://doi.org/10.1523/JNEUROSCI.3900-09.2009

Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *The Journal of Neuroscience, 29*(42),131651-13171. https://doi.org/10.1097/WNR.0b013e328344a689

Bigand, E. (2003). More about the musical expertise of musically untrained listeners. *Annals of the New York Academy of Sciences.* 999(1), 304-312. https://doi.org/10.1196/annals.1284.041

Bigand, E. (2006). Are we Experienced Listerners. A review of the musical capacities that do not depend on formal musical training. *Cognition.* 100, 100–130. doi: 10.1016/j.cognition.2005.11.007

Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., & Tillmann, B. (2014). Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00094

Bigand, E., Madurell, E., Tillmann, B., & Pineau, M. (1999). Effect of global structural and temporal organization on chord processing. *Journal of Experimental Psychology. Human Perception and Performance. 25*(1). 184-197. https://doi.org/10.1037/0096-1523.25.1.184

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*(4), 237–254. https://doi.org/10.1007/S004260050053

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, *58*(1), 125–141. https://doi.org/10.3758/BF03205482

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics*, *59*(7), 1098–1107. https://doi.org/10.3758/BF03205524

Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. A. (2003). Sensory Versus Cognitive Components in Harmonic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 159–171. https://doi.org/10.1037/0096-1523.29.1.159

Bischoff Renninger, L., Wilson, M. P., & Donchin, E. (2006). The Processing of Pitch and Scale: An ERP Study of Musicians Trained Outside of the Western Musical System. *Empirical Musicology Review*, *1*(4), 185–197. https://doi.org/10.18061/1811/24117

Bones, O., Hopkins, K., Krishnan, A., & Plack, C. J. (2014). Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia*, *58*(1), 23–32. https://doi.org/10.1016/j.neuropsychologia.2014.03.011

Boso, M., Politi, P., & Barale, F. (2006). Neurophysiology and neurobiology of the musical experience. *Functional Neurology*. 21(4), 187-191.

Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(36), 11155–11160. https://doi.org/10.1073/PNAS.1505768112

Bowling, D. L., Sundararajan, J., Han, S., & Purves, D. (2012). Expression of Emotion in Eastern and Western Music Mirrors Vocalization. *PLOS ONE*, *7*(3), e31942. https://doi.org/10.1371/JOURNAL.PONE.0031942

Brattico, E. (2015). From pleasure to liking and back: Bottom-up and top-down neural routes to the aesthetic enjoyment of music. In *Art, Aesthetics, and the Brain* (pp. 303–318). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199670000.003.0015

Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2016). It's sad but i like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Frontiers in Human Neuroscience*, *9*(JAN2016), 676. https://doi.org/10.3389/FNHUM.2015.00676/BIBTEX

Brattico, E., Jacobsen, T., De Baene, W., Glerean, E., & Tervaniemi, M. (2010). Cognitive vs. affective listening modes and judgments of music - An ERP study. *Biological Psychology*, *85*(3), 393–409. https://doi.org/10.1016/j.biopsycho.2010.08.014

Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, M., Eerola, T., & Tervaniemi, M. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience*, *21*(11), 2230–2244. https://doi.org/10.1162/jocn.2008.21144

Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 48–61. https://doi.org/10.1037/a0031624

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, *1117*(1), 162–174. https://doi.org/10.1016/j.brainres.2006.08.023

Brattico, E., Tupala, T., Glerean, E., & Tervaniemi, M. (2013). Modulated neural processing of Western harmony in folk musicians. *Psychophysiology*, *50*(7), 653–663. https://doi.org/10.1111/psyp.12049

Brattico, E., Winkler, I., Näätänen, R., Paavilainen, P., & Tervaniemi, M. (2002). Simultaneous storage of two complex temporal sound patterns in auditory sensory memory. *Neuroreport*, *13*(14), 1747–1751. https://doi.org/10.1097/00001756-200210070-00011

Carrión, R. E., & Bly, B. M. (2008). The effects of learning on event-related potential correlates of musical expectancy. *Psychophysiology*, *45*(5), 759–775. https://doi.org/10.1111/j.1469-8986.2008.00687.x

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J. D., & Koelsch, S. (2019). Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity. *Current Biology*, *29*(23), 4084-4092.e4. https://doi.org/10.1016/J.CUB.2019.09.067

Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19858–19863. https://doi.org/10.1073/pnas.1207989109

Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, *117*, 188–198. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2018.06.005

Curtis, M. E., & Bharucha, J. J. (2009). Memory and musical expectation for tones in cultural context. *Music Perception*, *26*(4), 365–375. https://doi.org/10.1525/mp.2009.26.4.365

Demorest, S. M., & Osterhout, L. (2012). ERP responses to cross-cultural melodic expectancy violations. *Annals of the New York Academy of Sciences*, *1252*(1), 152–157. https://doi.org/10.1111/j.1749-6632.2012.06464.x

Deutsch, D. (1982). Psychology of Music. *Academic Press Series in Cognition and Perception.*

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–427. https://doi.org/10.1017/S0140525X00058027

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science*, *270*(5234), 305–307. https://doi.org/10.1126/SCIENCE.270.5234.305

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory and Cognition*, *37*(1), 1–9. https://doi.org/10.3758/MC.37.1.1/METRICS

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, *25*, 355–373. https://doi.org/10.1016/j.clinph.2007.04.019

Friston, K. (2005). A theory of cortical responses. *Philosophical Transcations of the Royal Society*, *360*, 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2010). The free-energy principle: A unified brain theory? In *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & Koelsch, S. (2009). Universal Recognition of Three Basic Emotions in Music. *Current Biology*, *19*(7), 573–576. https://doi.org/10.1016/j.cub.2009.02.058

Gaab, N., & Schlaug, G. (2003). The effect of musicianship on pitch memory in performance matched groups. *NeuroReport*, *14*(18), 342–354. https://doi.org/10.1097/00001756-200312190-00001

García-Larrea, L., Lukaszewicz, A. C., & Mauguiére, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, *30*(8), 723–741. https://doi.org/10.1016/0028-3932(92)90042-K

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PloS One*, *4*(12). https://doi.org/10.1371/JOURNAL.PONE.0008144

Gordon, J., & Gridley, M. C. (2013). Musical Preferences as a Function of Stimulus Complexity of Piano Jazz. *Creativity Research Journal*, *25*(1), 143–146. https://doi.org/10.1080/10400419.2013.752303

Guo, S., & Koelsch, S. (2015). The effects of supervised learning on event-related potential correlates of music-syntactic processing. *Brain Research*, *1626*, 232–246. https://doi.org/10.1016/j.brainres.2015.01.046

Guo, S., & Koelsch, S. (2016). Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Scientific Reports*, *6*(1), 1–11. https://doi.org/10.1038/srep19064

Hagenow, C. F. (2005). The Equal Tempered Musical Scale. *American Journal of Physics*, *2*(3), 81. https://doi.org/10.1119/1.1992883

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*(11), 466–472. https://doi.org/10.1016/j.tics.2007.08.008

Hund-Georgiadis, M., & Yves Von Cramon, D. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Experimental Brain Research*, *125*(4), 417–425. https://doi.org/10.1007/S002210050698

Huron, D. (2018). Sweet Anticipation. *Sweet Anticipation*. https://doi.org/10.7551/MITPRESS/6575.001.0001

Itoh, K., Suwazono, S., & Nakada, T. (2003). Cortical processing of musical consonance: an evoked potential study. *Neuroreport*, *14*(18), 2303–2306.https://doi.org/10.1097/00001756-200312190-00003

Itoh, K., Suwazono, S., & Nakada, T. (2010). Central auditory processing of noncontextual consonance in music: An evoked potential study. *The Journal of the Acoustical Society of America*, *128*(6), 3781–3787. https://doi.org/10.1121/1.3500685

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1995.7.2.153

Jentschke, S. (2007). *Neural Correlates of Processing Syntax in Music and Language-Influences of Development, Musical Training, and Language Impairment*.

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, *47*(2), 735–744. https://doi.org/10.1016/j.neuroimage.2009.04.090

Johnson-laird, P., Kang, O., & Leong, Y.C. (2012). On musical dissonance. *Music Perception,* *30*(1), 19–35. https://doi.org/10.1525/mp.2012.30.1.19

Justus, T. C., & Bharucha, J. J. (2001). Modularity in Musical Processing: The Automaticity of Harmonic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 1000–1011. https://doi.org/10.1037/0096-1523.27.4.1000

Kalda, T., & Minati, L. (2012). Detecting Scale Violations in Absence of Mismatch Requires Music-Syntactic Analysis: A Further Look at the Early Right Anterior Negativity (ERAN). *Brain Topography*, *25*, 285–292. https://doi.org/10.1007/s10548-011-0208-8

Koelsch, S. (2009a). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, *46*(1), 179–190. https://doi.org/10.1111/j.1469-8986.2008.00752.x

Koelsch, S. (2009b). Neural substrates of processing syntax and semantics in music. *Music That Works: Contributions of Biology, Neurophysiology, Psychology, Sociology, Medicine and Musicology*, 143–153. Springer Science + Business Media/SpringerWienNewYork. https://doi.org/10.1007/978-3-211-75121-3_9

Koelsch, S. (2011). Toward a Neural Basis of Music Perception – A Review and Updated Model. *Frontiers in Psychology*, *0*(110). https://doi.org/10.3389/FPSYG.2011.00110

Koelsch, S., & Friederici, A. D. (2003). Toward the neural basis of processing structure in music - comparative results of different neurophysiological investigation methods. *Annals of the New York Academy*. *999,* 15-28. https://doi.org/ 10.1196/annals.1284.002

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Trends in Cognitive Neuroscience,* 9 (12), 578-84. https://doi.org/10.1016/j.neuroimage.2004.12.050

Koelsch, S., Grossmann, T., Gunter, T. C., Hahne, A., Schröger, E., & Friederici, A. D. (2003). Children Processing Music: Electric Brain Responses Reveal Musical Competence and Gender Differences. *Journal of Cognitive Neuroscience*, *15*(5), 683–693. https://doi.org/10.1162/089892903322307401

Koelsch, S., Gunter, T. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*, *12*(7), 1385–1389. https://doi.org/10.1097/00001756-200105250-00019

Koelsch, S., Gunter, T. C., v. Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach Speaks: A Cortical "Language-Network" Serves the Processing of Music. Neuroimage, *17*(2), 956–966. https://doi.org/10.1006/NIMG.2002.1154

Koelsch, S., Gunter, T., Friederici, A. D., & Èger, E. S. (2000). *Brain Indices of Music Processing: ``Nonmusicians" are Musical.*

Koelsch, S., Gunter, T., Schröger, E., & Friederici, A. D. (2003). Processing Tonal Modulations: An ERP Study. *Journal of Cognitive Neuroscience*, *15*(8), 1149–1159. https://doi.org/10.1162/089892903322598111

Koelsch, S., & Jentschke, S. (2009). Differences in Electric Brain Responses to Melodies and Chords. *Frontiers in Human Neuroscience. 12*(425). https://doi.org/10.1162/0.3389/fnhum.2018.00425

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*(3), 476–490. https://doi.org/10.1111/j.1469-8986.2007.00517.x

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*(3), 302–307. https://doi.org/10.1038/nn1197

Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport*, *14*(5), 709–713. https://doi.org/10.1097/00001756-200304150-00010

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38),15443–15448. https://doi.org/10.1073/PNAS.1300272110

Koelsch, S., & Sammler, D. (2008). Cognitive components of regularity processing in the auditory domain. *PLoS ONE*, *3*(7). https://doi.org/10.1371/journal.pone.0002650

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*(5), 657–663. https://doi.org/10.1017/S0048577202010508

Koelsch, S., Schoger, E., & Gunter, T. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology*, *39*(1), 38–48. https://doi.org/10.1017/s0048577202000185

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*(12), 578–584. https://doi.org/10.1016/j.tics.2005.10.001

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, *23* (1), 63–77. Elsevier Ltd. https://doi.org/10.1016/j.tics.2018.10.006

Koelsch, S., Wittfoth, M., Wolf, A., Müller, J., & Hahne, A. (2004). Music perception in cochlear implant users: An event-related potential study. *Clinical Neurophysiology*, *115*(4), 966–972. https://doi.org/10.1016/j.clinph.2003.11.032

Krumhansl, C. L. (1979). The Psychological Representation of Musical Pitch in a Tonal Context. *Cognitive Psychology, 11*(3). 346–374. https://doi.org/10.1016/0010-0285(79)90016-1

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*(1), 159–179. https://doi.org/10.1037/0033-2909.126.1.159

Krumhansl, C. L., Bharucha, J., & Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics 1982 32:2*, *32*(2), 96–108. https://doi.org/10.3758/BF03204269

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368. https://doi.org/10.1037/0033-295X.89.4.334

Kung, C. C., Hsieh, T. H., Liou, J. Y., Lin, K. J., Shaw, F. Z., & Liang, S. F. (2014). Musicians and non-musicians' different reliance of features in consonance perception: A behavioral and ERP study. *Clinical Neurophysiology*, *125*(5), 971–978. https://doi.org/10.1016/j.clinph.2013.10.016

Lahdelma, I., & Eerola, T. (2016). Mild dissonance preferred over consonance in single chord perception. *I-Perception*, *7*(3). https://doi.org/10.1177/2041669516655812

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, *1142*(1), 169–177. https://doi.org/10.1016/j.brainres.2007.01.049

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, *17*(4), 481–509. https://doi.org/10.2307/40285830

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., & Stephan, K. E. (2013). Modelling Trial-by-Trial Changes in the Mismatch Negativity. *PLoS Comput Biol*, *9*(2), 1002911. https://doi.org/10.1371/journal.pcbi.1002911

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception & Psychophysics 2007 69:7*, *69*(7), 1084–1092. https://doi.org/10.3758/BF03193946

Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009). *A generalized mechanism for perception of pitch patterns*. https://doi.org/10.1523/JNEUROSCI.4503-08.2009

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception & Psychophysics*, *52*(6), 599–608. https://doi.org/10.3758/BF03211696

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, *4*(5), 540–545. https://doi.org/10.1038/87502

Maher, T. F. (1976). "Need for Resolution" Ratings for Harmonic Musical Intervals. *Journal of Cross-Cultural Psychology*, *7*(3), 259–276. https://doi.org/10.1177/002202217673001

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*(11), 1035–1041. https://doi.org/10.1016/j.cub.2010.04.019

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, *18*(4), 452–463. https://doi.org/10.1016/j.conb.2008.09.005.Music

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature 2016 535:7613*, *535*(7613), 547–550. https://doi.org/10.1038/nature18635

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V, Simson, J., Bainbridge, C. M., Pinker, S., O, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*, 1–17. https://doi.org/10.1126/science.aax0868

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

Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., Loveday, C., & Bruzzone, M. G. (2008). FMRI/ERP of musical syntax: Comparison of melodies and unstructured note sequences. *NeuroReport*, *19*(14), 1381–1385. https://doi.org/10.1097/WNR.0b013e32830c694b

Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., Loveday, C., & Bruzzone, M. G. (2009a). Functional MRI/Event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *NeuroReport*, *20*(1), 87–92. https://doi.org/10.1097/WNR.0b013e32831af235

Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., Loveday, C., & Bruzzone, M. G. (2009b). Functional MRI/Event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *NeuroReport*, *20*(1), 87–92. https://doi.org/10.1097/WNR.0b013e32831af235

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *NeuroImage*, *38*(2), 331–345. https://doi.org/10.1016/j.neuroimage.2007.07.034

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(40), 15894–15898. https://doi.org/10.1073/PNAS.0701498104

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology*, *115*(1), 140–144. https://doi.org/10.1016/j.clinph.2003.04.001

Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*, *24*(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Nager, W., Kohlmetz, C., Altenmuller, E., Rodriguez-Fornells, A., Thomas, ¨, & Munte, F. (2003). The fate of sounds in conductors' brains: an ERP study. *Cognitive Brain Research*, *17*, 83–93. https://doi.org/10.1016/S0926-6410(03)00083-1

Pagès-Portabella, C., Bertolo, M., & Toro, J. M. (2021). Neural correlates of acoustic dissonance in music: The role of musicianship, schematic and veridical expectations. *PLOS ONE*, *16*(12), e0260728. https://doi.org/10.1371/JOURNAL.PONE.0260728

Pagès-Portabella, C., & Toro, J. M. (2019). Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians. *Psychophysiology*, 57(2), 1–11. https://doi.org/10.1111/psyp.13476

Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., & Schulz, M. (2003). Music and Learning-Induced Cortical Plasticity. *Annals of the New York Academy of Sciences*, *999*(1), 438–450. https://doi.org/10.1196/ANNALS.1284.054

Park, J. M., Chung, C. K., Kim, J. S., Lee, K. M., Seol, J., & Yi, S. W. (2018). Musical Expectations Enhance Auditory Cortical Processing in Musicians: A Magnetoencephalography Study. *Neuroscience*, *369*, 325–335. https://doi.org/10.1016/j.neuroscience.2017.11.036

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach* (Springer S). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-74831-8

Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins?. *Musicae Scientiae, 18*(3), 324–353. https://doi.org/10.1177/1029864914542842

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. https://doi.org/10.1177/1029864914542842

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected Studies of the Event-Related Potential. *International Journal of Medical Sciences*, *2* (4). https://doi.org/10.7150/ijms.2.147

Pearce, M. T. (2005). The construction and evaluation of statistical models of melodic structure in music perception and composition. *City University, London*.

Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. *Ann. N.Y. Acad. Sci.* https://doi.org/10.1111/nyas.13654

Peretz, I., Aubé, W., & Armony, J. L. (2013). Toward a neurobiology of musical emotions. *Evolution of Emotional Communication*, 277–299. https://doi.org/10.1093/ACPROF:OSO/9780199583560.003.0017

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89–114. https://doi.org/10.1146/annurev.psych.56.091103.070225

Peterson, D. C., Reddy, V., & Hamel, R. N. (2021). Neuroanatomy, Auditory Pathway. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK532311/

Plantinga, J., & Trehub, S. E. (2014). Revisiting the innate preference for consonance. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 40–49. https://doi.org/10.1037/a0033471

Plomp, R., & Levelt, W. J. M. (1965). Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America, 38*(4), 548-60. https://doi.org/10.1121/1.1909741

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysioloigy, 118(*19), 2128-48. https://doi.org/10.1016/j.clinph.2007.04.019

Popescu, T., Neuser, M. P., Neuwirth, M., Bravo, F., Mende, W., Boneh, O., Moss, F. C., & Rohrmeier, M. (2019). The pleasantness of sensory dissonance is mediated by musical style and expertise. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-018-35873-8

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience*, *18*(9), 1545–1554. https://doi.org/10.1162/JOCN.2006.18.9.1545

Proverbio, A. M., Orlandi, A., & Pisanu, F. (2016). Brain processing of consonance/dissonance in musicians and controls: a hemispheric asymmetry revisited. *European Journal of Neuroscience*, *44*(6), 2340–2356. https://doi.org/10.1111/ejn.13330

Przysinda, E., Zeng, T., Maves, K., Arkin, C., & Loui, P. (2017). Jazz musicians reveal role of expectancy in human creativity. *Brain and*

Cognition, *119*, 45–53. https://doi.org/10.1016/j.bandc.2017.09.008

Reber, R., Wurtz, P., & Zimmermann, T. D. (2004). Exploring "fringe" consciousness: The subjective experience of perceptual fluency and its objective bases. *Consciousness and Cognition*, *13*(1), 47–60. https://doi.org/10.1016/S1053-8100(03)00049-7

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255. https://doi.org/10.1162/089892901564298

Roberts, L. A. (1983). Consonance and dissonance judgments of musical chords. *The Journal of the Acoustical Society of America*, *73*(S1), S85–S85. https://doi.org/10.1121/1.2020602

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–175. https://doi.org/10.1016/j.ijpsycho.2011.12.010

Saarinen, J., Paavilainen, P., Tervaniemi, M., & Naatanen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport*, *3*(12), 1149–1151. https://doi.org/10.1097/00001756-199212000-00030

Saffran, J. (2003). Musical learning and language development. *Ann. N.Y Acad. Sci.*, *999*, 397–401. https://doi.org/10.1196/annals.1284.050

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*(2), 293–304. https://doi.org/10.1111/j.1469-8986.2007.00497.x

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience 2002 5:7*, *5*(7), 688–694. https://doi.org/10.1038/nn871

Schön, D., Regnault, P., Ystad, S., & Besson, M. (2005). Sensory Consonance. *Music Perception*, *23*(2), 105–118. https://doi.org/10.1525/mp.2005.23.2.105

Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research*, *7*(1), 71–87. https://doi.org/10.1016/S0926-6410(98)00013-5

Seppänen, M., Pesonen, A.-K., & Tervaniemi, M. (2012). Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. *Attention, Perception, & Psychophysics*, *74*(3), 600–612. https://doi.org/10.3758/s13414-011-0257-9

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. https://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, *16*(2), 374–381. https://doi.org/10.3758/16.2.374

Sloboda, J. A. (1985). The musical mind: the cognitive psychology of music. *Oxford Press.*

Steinbeis, N., & Koelsch, S. (2008). Shared Neural Resources between Music and Language Indicate Semantic Processing of Musical Tension-Resolution Patterns. *Cerebral Cortex*, *18*(5), 1169–1178. https://doi.org/10.1093/cercor/bhm149

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2005). Emotional processing of harmonic expectancy violations. *Annals of the New York Academy of Sciences*, *1060*, 457–461. https://doi.org/10.1196/annals.1360.055

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380

Sussman, E., & Winkler, I. (2001). Dynamic sensory updating in the auditory system. In *Cognitive Brain Research* (Vol. 12). www.elsevier.com/locate/bres

Tekman, H. G., & Bharucha, J. J. (1998). Implicit Knowledge Versus Psychoacoustic Similarity in Priming of Chords. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 252–260. https://doi.org/10.1037/0096-1523.24.1.252

Terhardt, E. (1984). The Concept of Musical Consonance: A Link between Music and Psychoacoustics. *Music Perception*, *1*(3), 276–295. https://doi.org/10.2307/40285261

Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., & Näätänen, R. (1997). The musical brain: brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, *226*(1), 1–4. https://doi.org/10.1016/S0304-3940(97)00217-6

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10. https://doi.org/10.1007/s00221-004-2044-5

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). *Superior Formation of Cortical Memory Traces for Melodic Patterns in Musicians*. https://doi.org/10.1101/lm.39501

Tervaniemi, M., Tupala, T., & Brattico, E. (2012). Expertise in folk music alters the brain processing of Western harmony. *Annals of the New York Academy of Sciences*, *1252*(1), 147–151. https://doi.org/10.1111/j.1749-6632.2011.06428.x

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913. https://doi.org/10.1037/0033-295X.107.4.885

Tillmann, B., & Bigand, E. (2010). *Musical structure processing after repeated listening: Schematic expectations resist veridical expectations*. https://doi.org/DOI:10.1177/10298649100140S204

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception*, *16*(1), 99–117. https://doi.org/10.2307/40285780

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the Inferior Frontal Cortex in Musical Priming. *Annals of the New York Academy of Sciences*, *999*(1), 209–211. https://doi.org/10.1196/ANNALS.1284.031

Tillmann, B., Poulin-Charronnat, B., & Bigand, E. (2014). The role of expectation in music: From the score to emotions and the brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(1), 105–113. https://doi.org/10.1002/wcs.1262

Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental Psychobiology*, *46*(3), 262–278. https://doi.org/10.1002/DEV.20059

Trainor, L. J., & Trehub, S. E. (1992). A Comparison of Infants' and Adults' Sensitivity to Western Musical Structure. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 394–402. https://doi.org/10.1037/0096-1523.18.2.394

Trainor, L. J., & Trehub, S. E. (1993). Musical Context Effects in Infants and Adults: Key Distance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(3), 615–626. https://doi.org/10.1037/0096-1523.19.3.615

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception & Psychophysics,* *56*(2), 125–132. https://doi.org/10.3758/BF03213891

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for Sensory Consonance in 2- and 4-Month-Old Infants. *Music Perception*, *20*(2), 187–194. https://doi.org/10.1525/MP.2002.20.2.187

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological Foundations for the Theory of Harmony in Western Tonal Music. *Annals of the New York Academy of Sciences*, *930*(1), 92–116. https://doi.org/10.1111/j.1749-6632.2001.tb05727.x

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience 2003 6:7*, *6*(7), 669–673. https://doi.org/10.1038/nn1084

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: the role of melodic contour. *Child Development*, *55*(3), 821–830. https://doi.org/10.1111/J.1467-8624.1984.TB03819.X

Trehub, S. E., Thorpe, L. A., & Trainor, L. J. (1990). Infants' perception of good and bad melodies. *Psychomusicology: A Journal of Research in Music Cognition*, *9*(1), 5–19. https://doi.org/10.1037/H0094162

Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., & Tervaniemi, M. (2011a). The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. *Neuroscience Letters*, *487*(3), 406–410. https://doi.org/10.1016/j.neulet.2010.10.066

Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., & Tervaniemi, M. (2011b). The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. *Neuroscience Letters*, *487*(3), 406–410. https://doi.org/10.1016/j.neulet.2010.10.066

Virtala, P., & Tervaniemi, M. (2017b). Neurocognition of Major-Minor and Consonance-Dissonance. *Music Perception: An Interdisciplinary Journal*, *34*(4), 387–404. https://doi.org/10.1525/mp.2017.34.4.387

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*(3), 311–317. https://doi.org/10.1525/MP.2012.29.3.311

Vuust, P., Brattico, E., Glerean, E., Seppä Nen, M., Pakarinen, S., Tervaniemi, M., Nä, R., & Tä Nen, ¨. (2011). Special section on music in the brain: Research report New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. *CORTEX*, *47*, 1091–1098. https://doi.org/10.1016/j.cortex.2011.04.026

Weiss, M. W., Cirelli, L. K., Mcdermott, J. H., & Trehub, S. E. (2020). Development of consonance pº in western listeners. *Journal of Experimental Psychology: General*, *149*(4), 634-649. https://doi.org/10.1037/xge0000680

Wessel, J. R., & Aron, A. R. (2017). On the Globality of Motor Suppression: Unexpected Events and Their Influence on Behavior and Cognition. *Neuron*, *93*(2), 259–280. https://doi.org/10.1016/j.neuron.2016.12.013

Woolhouse, M., & Cross, I. (2010). Using Interval Cycles to Model Krumhansl's Tonal Hierarchies. *Music Theory Spectrum*, *32*(1), 60–78. https://doi.org/10.1525/MTS.2010.32.1.60

Yu, X., Liu, T., & Gao, D. (2015). The Mismatch Negativity: An Indicator of Perception of Regularities in Music. *Behavioural Neurology*, *2015*. https://doi.org/10.1155/2015/469508

Zatorre, R. (2016). Amazon music. *Nature, 535,* 496–497. https://doi.org/10.1038/nature18913

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development*, *21*(3), 483–492.

Zhang, J., Zhou, X., Chang, R., & Yang, Y. (2018). Effects of global and local contexts on chord processing: An ERP study. *Neuropsychologia*, *109*, 149–154. https://doi.org/10.1016/j.neuropsychologia.2017.12.016