

## CHAPTER 1

# Why Humans Would Not Have Survived Evolution without Music

**M**USIC IS DEEPLY ROOTED IN OUR SPECIES. While the oldest known musical instruments date back approximately 40,000 to 50,000 years, I shall argue in this chapter that the origins of music are likely to extend much further, probably coinciding with the emergence of *Homo sapiens*, if not predating it. So far, the oldest fossilized remains of *Homo sapiens* date back approximately 300,000 years; however, it is likely that the first humans existed well before this timeframe.<sup>1</sup> Whether the preceding human species made music remains an open question, but it is not improbable. About 1.5 million years ago, the early members of the genus *Homo* discovered the art of cooking. This breakthrough allowed them to get more calories in less time, laying the foundation for the enormous growth of a metabolically expensive organ that accounts for roughly 20 per cent of the resting metabolic rate in modern humans: the brain.<sup>2</sup> Brain evolution brought several new abilities, among them two music-specific skills: holding a pulse in a group and singing tones together.

Together with the birth of music in human evolution, the beneficial effects of music on health and social bonding emerged. Skills essential for communication and cooperation, along with novel forms of social organizations, developed rapidly. As humans started to live in larger

<sup>1</sup> Hublin, J. J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., and Gunz, P. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature*, 546(7657), 289.

<sup>2</sup> Attwell, D., and Laughlin, S. B. (2001). An energy budget for signalling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10), 1133–1145.

communities, this led to the emergence of more complex social structures. Whether music is a prerequisite or a concomitant of this development remains unknown. However, given its significant effects on social cohesion and health, I posit that humans would not have survived evolution without music.

Music represents a unique category within the realm of sound. More precisely, music is a succession of sounds in which we feel a pulse (usually a beat) and where sounds – if they have pitches – correspond to a scale. Around the globe, there are many scales: in addition to major and minor scales, one finds Gregorian modes, jazz scales, Indian ragas, the Indonesian pelog and slendro, as well as pentatonic and octatonic scales. Among the various scales, the pentatonic scale stands out for its simplicity. It comprises only five notes, and preschool children can easily sing it (as in the song ‘Old MacDonald Had a Farm’).

When humans produce sounds according to both a scale and a pulse, we recognize these sounds as *music*. With a few exceptions, the musical traditions of the *Homo sapiens* are based on these two characteristics: pulse and scale. They build the core of a universal grammar of music with two basic principles (technically called ‘rules’). These two core rules of a universal musical grammar are as follows: ‘the time intervals between sounds should be structured such that they fit recognizably into a pulse’, and ‘the pitches of sounds should be recognizable elements of a scale’. Strikingly, this simple universal grammar has led to the immense variety and diversity of musical systems, styles, and compositions.

However, not all music adheres to the principles of beat and scale. Drum music can get by without scales, and meditation music often has no recognizable beat, as with some pieces of modern art music (for example, Ligeti’s ‘Atmosphères’, which many know from Stanley Kubrick’s film *2001: A Space Odyssey*).

The immediate function of beat and scale lies in facilitating *collective* music making. We can best perform movements *together* simultaneously if they follow a beat. If we want to lift a heavy box together ‘on three’, it makes no sense if I first say ‘one’ slowly, then wait, and then abruptly and quickly say ‘two, three’. You count on *the beat*: ‘one – two – three!’ To clap, dance, stomp, or shout together, we need a beat. To sing together, a group must agree on which notes to sing, achieved using a

scale: a scale provides a set of pitches everyone can follow. Without a scale, there would be no coherent and harmonious blend of musical notes; without a tactus, the resulting sound would be chaotic and disorganized.

The human ability to harmonize pitches and synchronize beats is not coincidental; it has been a vital factor in our evolutionary success. This unique musical ability provided humans with a significant evolutionary advantage – to live longer. This advantage includes the following:

- *Better cooperation and stronger social cohesion.* When people make music together, they engage in cooperative activities that foster a sense of unity and shared purpose. This collaborative spirit extends beyond the realm of music, leading to heightened cooperation and prosocial behaviours in various facets of life. For example, after engaging in joint music making, individuals are more inclined to help each other, enhancing the likelihood of achieving collective goals while minimizing the potential for conflicts. Humans were successful in evolution because they were more potent in groups than individually, and music's role in facilitating a sense of unity through coordinated movement often leads to selfish tendencies evolving into a commitment to the group. When singing or clapping with one voice, individuals transform from 'I' to 'we'. In the subsequent chapters, we will delve deeper into the profound impact of engaging in cooperation on both health and social relationships.
- *More positive emotions and promotion of health.* Music can evoke positive emotions and help regulate negative emotions. Due to this capacity, it can contribute to healing and enhance overall well-being. Whereas prolonged emotional stress has unhealthy consequences, facilitation of relaxation and joy with music has restorative effects. With music, we can relieve pain, and music may support us to persevere during difficult times. Music can invigorate the soul and thereby motivate individuals to persist, even potentially saving lives in extreme cases. (Shackleton's men in Antarctica persevered with music in the face of intense pain and hardships – see the Preface.) As we progress through this book, we will encounter many instances where music's ability to evoke positive emotions has therapeutic effects.

- *Conflict mitigation.* The decrease in physical confrontations within or between groups leads to a corresponding reduction in injuries and fatalities. Various hunter-gatherer cultures have customs where, instead of through duels using weapons, disputes are settled with the use of singing.<sup>3</sup> Such ‘song duels’ restore peaceful social relations and thus prevent violent confrontations, acts of revenge, or even murder. Because nomadic cultures around the globe practise such conflict-reducing musical customs, they seem to be inherent in human nature, probably dating back to the time the *Homo sapiens* came into existence. In the upcoming sections, we will come to understand how music’s social functions contribute to conflict resolution and foster peace.

### MUSIC AND LANGUAGE HAVE INTERTWINED EVOLUTIONARY ROOTS

Similarly to music, speech is also structured sound produced by humans. When we speak, we use melody to distinguish between questions and answers, rhythm to help conversation partners follow each other better, and timbre to convey the speaker’s mood. However, only *one* person can speak at a time; otherwise, speech sounds unpleasant and challenging to understand. In contrast, with music, several people can produce sounds simultaneously, and it still sounds good and is understandable. This capacity of music to facilitate collective expression surpasses that of language. Therefore, music is the language of the group, while language is the music of the individual.

The evolutionary advantage of language is that a single person can communicate their thoughts, intentions, desires, feelings, and so on. Language is, therefore, a special case of music: language comprises sounds, whereby the beat and the scale are considerably less clear than in music, and the sounds form words with specific meanings. Uli Reich, the linguist, once remarked to me that language is ‘music distorted by semantics’. The decisive difference between language and music is that

<sup>3</sup> Lehmann, C., Welker, L., and Schiefenhövel, U. W. (2008). The singing controversy in human ethological perspective. *Musicae Scientiae*, 12(1), 115–145.

to fulfil its function, language *does not need* pulse or scale, and music *does not need* sounds with specific meanings.

We can note that the *meaning* of words often relates to their *sounds*, even if it may seem random when we learn a new language. For instance, different languages use different words to express the same concept or property, such as ‘tiny’ in English, ‘*winzig*’ in German, ‘*bitte liten*’ in Norwegian, ‘*infima*’ in French, and ‘*piccolissimo*’ in Italian. However, a closer listen reveals a striking commonality: these words all contain two or more [i] sounds (English: ‘eee’). A scientific study compared a basic vocabulary of 100 words among 4,000 languages, about two-thirds of the known languages.<sup>4</sup> In many words, the investigation discovered systematic clusters of specific sounds, such as the sound ‘i’ in words for ‘tiny’, ‘r’ in words for ‘round’, or ‘n’ in words for ‘nose’. Because the study’s authors observed such similarities across different language families, they assumed these similarities arose independently and did not originate from a common original language. Therefore, the sounds of words are not as arbitrary as was long believed. As we appreciate this connection between the sounds of words and their meaning, we can see how music and language have deep and intertwined evolutionary roots.

### THE EMOTIONAL IMPACT OF THE VOICE

The close interconnection between music and language becomes even clearer when considering that musical features encode the emotional content of a voice. Why do we recognize that a voice sounds happy, sad, angry, surprised, or afraid? The music psychologists Patrik Juslin and Petri Laukka analysed data from about forty studies investigating which acoustic features of the human voice characterize certain emotions.<sup>5</sup> For this purpose, actors and actresses recorded words or sentences to express joy, sadness, anger, fear, or tenderness. Some studies also used recordings of real-life emotional voices, such as screams of fear in aircraft accidents. The voice recordings

<sup>4</sup> Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., and Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, 113(39), 10818–10823.

<sup>5</sup> Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770.

spanned different languages and cultures. Juslin and Laukka confirmed that specific acoustic features in the voice encode each emotion, enabling us to recognize these emotions in speech even without understanding the language. A cheerful voice, for instance, exhibits a faster-speaking tempo than a sad one, with a higher volume and greater variability in pitches, making the speech melody fluctuate and sound active. In contrast, a sad voice sounds darker, less bright, less melodious, with pitches tending to descend. By these acoustic characteristics, a Yanomami man from the Amazon rainforest, who has never had contact with Western culture, would recognize whether I feel happy, anxious, or sad. These acoustic features, encoding emotions in speech, are universal, transcending cultural barriers and highlighting a fundamental aspect of human communication.

The highlight of this study: Patrik and Petri also analysed a dozen studies investigating the acoustic characteristics through which musicians express emotions in music. For this purpose, musicians had played melodies to express joy, sadness, anger, fear, or tenderness. Besides classical music, these studies also used folk music, Indian ragas, jazz, rock, children's songs, and free improvisations played by musicians from different countries and cultures. Results showed that the acoustic features of emotional speech are predominantly the same ones that characterize the expression of emotions in various types of music. The beginning of the *Lacrimosa* ('weeping') from Mozart's Requiem, which sounds so sad, is also slow and quiet (*piano*), and the melody often descends (the so-called *pianto* motif). The fourth movement from Mozart's serenade 'Eine kleine Nachtmusik' sounds cheerful because it is relatively fast, has a high pitch variability (even just the first four notes played by the first violin span an octave), the melody often goes up, and the frequency register is relatively high.

Digital analysis allows for the precise measurement of these acoustic features from audio files. We took advantage of this for an experiment and selected happy- or scary-sounding music based on computer-calculated acoustic features.<sup>6</sup> We had taken the audio files of the scary

<sup>6</sup> Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Küssner, M. B., and Jacobs, A. M. (2013). The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *NeuroImage*, 81, 49–60.

music from soundtracks of thriller TV series and horror movies. The computer analysis showed that this type of music had many noisy, hissing, and percussive sounds, that is, sounds where the determination of the pitch was difficult – which caused uncertainty in the listener. In addition, assigning tones and chords to a key was often tricky, increasing this uncertainty. Finally, many chords were dissonant, which made them more uncomfortable to listen to (think of Bernard Herrmann’s music for Hitchcock’s *Psycho*, especially the shower scene).<sup>7</sup>

Western music often imitates emotional speech. Consistent with the findings that *emotional speech* is universally recognized, my research group has found that the expression of emotions through the *imitation of emotional speech* in Western music is also universally recognized, regardless of the listener’s cultural background. To explore this, Thomas Fritz, then a PhD student in my group, conducted an expedition to a remote region in northern Cameroon. There, he sought out participants from the Mafa people. These participants had never heard Western music before. He played them short piano pieces that sounded like joy, sadness, or fear. After each piece, the participants saw three photos: a happy face, a sad face, and a fearful face. They had the task of pointing to the face that best matched the music. The Mafa recognized all three emotions well above chance level, showing that the expression of happiness, sadness, and fear in Western music is universally recognized, that is, independent of cultural experience.<sup>8</sup>

The Mafa recognized the emotions less successfully than Western listeners, but it is important to note that the concept of expressing emotions with music was entirely new for them. The music of the Mafa people always has a happy meaning, and thus they only know

<sup>7</sup> The music used in our experiment can be downloaded and listened to at [stefan-koelsch.de/stimulus\\_repository/joy\\_fear\\_neutral\\_music.zip](https://stefan-koelsch.de/stimulus_repository/joy_fear_neutral_music.zip). When I heard Beat Furrer’s piano concerto for the first time a few years ago, it enthralled me how it imitated precisely all these acoustic ‘thriller parameters’. Enthusiastically, I listened to the concerto repeatedly because it reminded me of suspenseful car-chasing scenes from older American crime films. When I spoke to Beat Furrer about my observations, he replied, ‘Interesting – I didn’t mean to express that at all!’ Anyway, I’m still a big fan of his concert.

<sup>8</sup> Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., and Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, 19(7), 573–576.

happy-sounding music. In addition, Western music sounds utterly different to the Mafa than it does to Western listeners. One Mafa man particularly enjoyed the music of Elvis Presley because he thought it sounded just like croaking frogs!

The results of this study reveal that even people unfamiliar with Western music recognize emotions expressed in it as long as the music sounds similar to an emotional voice. Such recognition occurs because recognizing *vocal* expressions of emotion is, to a considerable extent, biologically and genetically inherent to us.<sup>9</sup> The universal recognition of vocal expression of emotions also means that we can formulate universal definitions of how music *sounds* positive (for example, happy) or negative (for example, scary). Whether we *feel* it in that specific way is another question entirely. Sometimes, frightening music (i.e. negative-sounding) makes a thriller film particularly enjoyable for us, or positive-sounding country music gets on the nerves of a heavy metal enthusiast. Of course, music can sound happy to somebody even if it does not imitate or portray any emotions. For the Mafa, their music always has a ‘happy’ sound, but for listeners from Western cultures, it may sound very much like honking cars.

Thus, the sound of speech conveys the speaker’s emotional state, and the sounds of many words are related to their meaning. Since these *musical* aspects of language occur independently of culture, we can conclude that they are part of the essential biological endowment of humans. Babies can already recognize emotional signals that belong to this natural endowment. They are emotionally touched by music and by the sound of language. Therefore, it is important that one’s voice sounds warm and calm when interacting with babies and that it conveys security and protection. Singing lullabies to babies helps them calm down, decreasing their heart rate and making their movements and breathing slower and more regular. Such calming effects are particularly important

<sup>9</sup> Zimmermann, E., Leliveld, L. M. C., and Schekka, S. (2013). Toward the evolutionary roots of affective prosody in human acoustic communication: A comparative approach to mammalian voices. In Altenmüller, E., Schmidt, S., and Zimmermann, E. (eds). *Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man* (pp. 116–132). Oxford, Oxford University Press.

for preterm infants because agitation is dangerous for them. Moreover, music can help to relieve pain already in these infants.<sup>10</sup>

Interestingly, the acoustic and musical characteristics of lullabies are largely universal, sounding similar across diverse cultures around the globe. The melodies typically descend, are relatively simple in structure, and are repetitive (think of the English ‘Twinkle Twinkle Little Star’, the Japanese ‘*Yurikago No Uta*’, and the South African ‘*Thula Baba*’).<sup>11</sup> Therefore, today’s lullabies probably sound similar to those sung hundreds of thousands of years ago.

Unfortunately, some parents think of themselves as unmusical and do not sing songs to their babies. However, they forget that an infant does not yet have any opportunities for comparison, and thus no parent can embarrass him- or herself in front of their baby by singing. Moreover, the purpose of singing lullabies is not to pave the way for a baby’s future career in opera but to support social, emotional, and cognitive development in a playful and engaging manner. That is why singing is important, even if a parent thinks that s/he is unmusical: it promotes social bonding between parent and child, communication, and the learning of speech sounds. It also engages multisensory experiences, that is, experiences of several senses at the same time (hearing, seeing, feeling, sensing the own body moving when being cradled). Even singing these songs *before* birth is helpful because the baby will recognize them after birth, calming them. It can also be beneficial to occasionally place a music box on the pregnant woman’s belly because the baby will recognize this song after birth, providing a sense of security for the infant.

#### SINGING, TALKING, AND DANCING WITH BABIES

Armed with the understanding of music’s profound impact on early development, you may be eager to bring its benefits into your

<sup>10</sup> Ilari, B. (2002). Music and babies: A review of research with implications for music educators. *Update: Applications of Research in Music Education*, 21(2), 17–26.

<sup>11</sup> Trehub, S. E., Unyk, A. M., and Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behaviour and Development*, 16(2), 193–211.

parenting routine. Below are some practical tips to make the most of music in bonding with, comforting, and engaging with your child.

- *The importance of tone in voice.* Infants are immediately influenced by the emotional tone of a voice. If a voice sounds angry, annoyed, or depressed, it inevitably triggers negative emotions, stress, and restlessness in the baby. Therefore, it's essential for one's own voice to sound warm and calm, conveying peacefulness, security, and safety. Singing lullabies and playing songs to an infant come naturally in this manner. Look into the baby's face with a welcoming expression and gently sway them to the rhythm of the music. Singing is particularly helpful during moments when you feel overwhelmed, anxious, or even depressed, as it is difficult to sound angry or upset while singing a gentle lullaby. Additionally, it results in slower and deeper breaths, which will help one become calmer and more relaxed. Thus, singing is *especially* effective when one least feels like it.
- *An alternative to singing – dancing.* If you cannot sing at the moment or it is not working, play some dance music (not too loud) and gently dance with your baby. For example, softly sway with your infant to the beat of the music.
- *Physical comfort through touch.* If the baby still is not calm, make skin contact and slowly stroke up and down one of the baby's arms (this activates nerve fibres that reduce pain).
- *Communication during care activities.* When changing nappies and clothes, calmly and warmly explain to the baby what *you* are doing (you would want the same courtesy if someone was doing things to you without asking). Avoid commenting on everything the *baby* does, as even babies find that annoying.
- *Structured routines and music.* Babies appreciate music, not only for its pleasant sound but also for its clear structure. This makes the world more predictable. Therefore, establishing a daily routine with a consistent rhythm that includes recurring times for sleeping, eating, playing, and singing can help the baby. This structure allows the infant to anticipate what's coming next and even develop a biological daily rhythm. (Of course, this rhythm should never be

enforced but should correspond to the baby's needs and continuously adapt to their development.)

- *Professional help for special cases.* In cases of premature birth or post-partum depression, it is highly recommended to seek help from a music therapist. For music therapists: such help should include tips on singing with the baby.

### THE HUMAN PREFERENCE FOR RHYTHM OVER CHAOS

Humans also have a biologically deep-rooted predisposition for rhythm. People generally prefer to hear rhythmic information over unstructured temporal information. On top of that, they involuntarily *produce* a rhythm or beat from random sound events. Andrea Ravignani, a mathematician and biologist, asked participants to imitate chaotic-sounding drum sequences generated quasi-randomly by a computer.<sup>12</sup> Then, he played the recordings of the imitated sequences to another group of participants, the 'next generation'. These participants were also asked to imitate the drum sequences in their recordings. This procedure was repeated several times, with each person unconsciously adding more rhythm to the drum sequences than was in the recordings they heard. Eventually, a beat emerged from what was originally a random drum sequence. Thus, the initially chaotic drum sequences gradually evolved into more rhythmic and musical sequences over multiple generations of participants.

Fascinatingly, the rhythmic sequences generated by the participants in Andrea's experiment exhibited musical properties corresponding to several statistical universals observed in music worldwide. Specifically, participants endowed the isochronous pulse with a 'metre', for example, grouping the underlying pulse either in twos (as in a march) or in threes (as in a waltz). Moreover, the resulting rhythmic sequences exhibited a structured pattern in which the bars contained a limited number of note

<sup>12</sup> Ravignani, A., Delgado, T., and Kirby, S. (2017). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, 1(1), 0007.

durations, typically no more than five, such as quarters, eighths, and dotted eighths. (Eighth notes have half the duration of a quarter note, and a dotted eighth note is three-quarters of a quarter note.) Finally, participants used these principles to create rhythmic figures, beats, and riffs. This structuring also made the drum sequences easier to learn, as they unconsciously aligned with human memory and cognitive abilities. The innate cognitive and biological properties of the human brain and body predispose musical rhythm to exhibit universal properties.

Many animals also communicate via sounds. Unlike natural sounds and sound textures (such as rain, the crackling of a fire, water splashing, or wind rustling), the sounds produced by animals are structured by living organisms and, therefore, often remind us of music: the songs of birds, whales, or gibbons, the synchronous chirping of cicadas, the drumming on tree roots and body parts by apes. However, as of the present stage of knowledge, there is no animal species where several individuals sing or drum in unison. *Homo sapiens*, by contrast, are the only vertebrate species in which several individuals can produce and keep a beat in a group, sometimes faster, sometimes slower, sometimes accelerating or decelerating. Although some studies reported that animals can synchronize their movements to a beat, their methods and conclusions have been contested by the scientific community, or the reported behaviours cannot be observed in the wild. Yes, whales do sing, but not in a choir. Gorillas pound on their chests together, but not in synchrony, let alone as a combo.

What no other species can achieve, humans are capable of just a few months after birth. The psychologist Marcel Zentner and the musicologist Tuomas Eerola played music to babies between five and ten months old. The pieces were lively classical music, such as the finale from Saint-Saëns' 'The Carnival of the Animals'.<sup>13</sup> As they listened, the babies started kicking so that the pulse of their movements matched the pulse of the music. This fascinating finding shows the innate tendency of humans to engage in music. In addition, the babies smiled when they synchronized their movements to the music – participating in music

<sup>13</sup> Zentner, M., and Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, 107(13), 5768–5773.

naturally gives us humans pleasure. The babies in this study were from Finland and Switzerland. A subsequent investigation obtained nearly identical results with babies from Brazil, except that the Brazilian infants moved significantly more to the same music.<sup>14</sup> They were probably already warming up for Carnival! These studies show that music stimulates a social function in us: moving together. The social effects of this function are prosocial behaviour and cooperation. When fourteen-month-old babies are bounced to music with a matching pulse, they tend to be more helpful afterwards than when bounced to a pulse that does not match the music (they are more likely to help the experimenter retrieve an ‘accidentally’ dropped pencil).<sup>15</sup> The anthropologists Sebastian Kirschner and Michael Tomasello observed a similar effect: after four-year-old children played music together, they cooperated more with each other and helped each other more.<sup>16</sup> Through this lens, the evolutionarily adaptive functions of music become apparent even in young children, who display increased prosocial behaviour after participating in musical activities.

Human musicality has its roots in the evolutionary development of auditory, vocal, and motor systems in mammals, a journey spanning tens of millions of years. Yet, it is only in humans that we find music characterized by pulse and scale, performed or sung collectively. Therefore, I propose that the simplest mental function distinguishing humans from animals is the ability to synchronize movements in a group to a pulse. This admittedly bold proposal means that music was the decisive evolutionary step of the *Homo sapiens* – possibly even of the genus *Homo*. Precisely this step brought several advantages to humans from which every individual can still benefit today, including positive social, emotional, and health effects. I will deal with these effects in more detail in the following chapters.

<sup>14</sup> Ilari, B. (2015). Rhythmic engagement with music in early childhood: A replication and extension. *Journal of Research in Music Education*, 62(4), 332–343.

<sup>15</sup> Cirelli, L. K., Einarson, K. M., and Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behaviour in infants. *Developmental Science*, 17(6), 1003–1011.

<sup>16</sup> Kirschner, S., and Tomasello, M. (2010). Joint music making promotes prosocial behaviour in 4-year-old children. *Evolution and Human Behavior*, 31(5), 354–364.